The National Archives

Guidance on Cloud Storage and Digital Preservation

How Cloud Storage can address the needs of public archives in the UK

Second Edition, March 2015 with updated case studies

Prepared by Charles Beagrie Ltd
Authors: Neil Beagrie, Andrew Charlesworth, and Paul Miller

Contents

Abstract 3

1 **Introduction** 3
- 1.1 Aims, Audience and Structure 3
- 1.2 What is Cloud Storage? 4
- 1.3 Cloud Storage and Digital Preservation 5
- 1.4 Security and the Cloud 7
- 1.5 Legal Issues 8
- 1.6 Costs 10
- 1.7 Summary of Key Issues 10

2 **Step by Step Guide** 12
- 2.1 Establishing Your Needs 12
- 2.2 Service Options 12
- 2.3 Service Providers 14
- 2.4 Procurement Options 19
- 2.5 Developing a Business Case 21

3 **Future Developments** 22
- 3.1 Overview 22
- 3.2 Keeping up to Date 23

4 **Current Best Practice** 24

5 **Case Studies** 26
- 5.1 Archives and Records Council Wales Digital Preservation Group 26
- 5.2 Dorset History Centre 26
- 5.3 Parliamentary Archives 26
- 5.4 Tate Gallery 26
- 5.5 University of Oxford 26
- 5.6 King's College London 27

6 **Sources of Further Advice and Guidance - annotated bibliography** 27
- 6.1 Cloud – General 27
- 6.2 Cloud and Digital Preservation 27
- 6.3 Cloud and Legal Issues 28
- 6.4 Relevant Standards and Good Practice 29

7 **Appendix** 31

Table 3 - Legal Issues 32

Abstract

The use of cloud storage in digital preservation is a rapidly evolving field and this guidance explores how it is developing, emerging options and good practice, together with requirements and standards that archives should consider. Five detailed case studies of UK archives that have implemented cloud storage solutions have been compiled as part of the Guidance and are available as standalone linked documents. Sources of further advice and guidance are also included.

1 Introduction

1.1 Aims, Audience and Structure

Digital preservation is a significant issue for almost all public archives. There is an increasing demand for storage of both born-digital archives and digitised material, and an expectation that public access to this content will continue to expand. At the same time the UK Government's recent adoption of a 'Cloud First' policy for public sector IT procurement is mandated to central government and strongly recommended to the wider public sector to achieve better value for money in IT services and data storage.

This Guidance is focussed on the cloud and its potential role in archival storage. It aims to help public archives in the UK develop an understanding of cloud storage and its potential contribution to their digital preservation activities, and to provide a balanced overview allowing archives to understand potential benefits and risks involved and the range of options available (including not using cloud if it does not meet your requirements).

Whilst primarily targeted at public archives, the aim is to provide information that will be useful within a range of organisational contexts, and overarching advice that can be translated into the private sector where relevant.

Its key audiences are archivists, records managers, and information management specialists in places of deposit and other public sector archives in the UK. The experience and the scale of digital preservation, or awareness and use of cloud storage, varies considerably across these archives but for the majority they are relatively new areas. This Guidance is therefore intended to be accessible to individuals with a range of previous knowledge and experience. It may also be of interest to other audiences, including private sector archives, cloud and digital preservation service providers, or IT and other professionals working with archives, and those outside the UK.

Although it is concise and accessible, it also has separate case studies (see section 5), sources of further advice and guidance (section 6) and an appendix on legal issues that can support more detailed analysis of options and requirements as needed.

The sector is very diverse and archives can be found in local and national government, the museums sector, and higher education. They have a wide range of governance, management and funding structures. The Guidance and case studies therefore have been prepared and selected to be applicable to a broad range of archives. The requirements and key needs in terms of the content of guidance for the sector have been collated via a series of interviews and a focus group with archive representatives.

The Guidance is structured into seven main sections:

1. Introduction – providing a general overview of key areas, definitions, and issues;

2. A Step by Step Guide – taking you through establishing a business case, requirements, services, providers, and procurement options;

3. Future Developments – a horizon scan suggesting likely developments in the field over the next one-two years;

4. Current Good Practice – a brief summary list of suggested good practice identified in compiling the Guidance;

5. Case studies of UK archives - that have implemented cloud solutions - case studies include detailed discussion of organisational context , nature of digital preservation requirements and approaches, use of cloud, technical infrastructure, business case and funding; and the key lessons they have learnt;

6. Sources of Further Advice and Guidance - an annotated bibliography providing short descriptions and key details to allow you to select relevant more detailed information and areas of interest to follow up;

7. A Table of Legal Issues – a concise summary of some of the key legal issues.

1.2 What is Cloud Storage?

Cloud Computing is a term that encompasses a wide range of use cases and implementation models. In essence, a computing 'cloud' is a large shared pool of computing resources including data storage. When someone needs additional computing power, they are simply able to check this out of the pool without much (often any) manual effort on the part of the IT team, which reduces costs and significantly shortens the time needed to start using new computing resources. Most of these 'clouds' are run on the public Internet by well-known companies like Amazon and Google. Some larger organisations have also found value in running private clouds inside their own data centres, where similar economies of scale begin to apply. Whilst we are specifically concerned with the potential utility of cloud-based archival storage solutions, it is useful to briefly consider the generally accepted characteristics of a typical cloud service. These characteristics, enumerated in a broadly accepted piece of work[1] by the United States' National Institute of Standards & Technology (NIST), may be paraphrased for the purposes of this Guidance as defining computers and data storage which are:

- Available when required ('on demand'), without the need for lengthy procurement and configuration processes;

- Available on standard networks such as the Internet, without special requirements for obscure or proprietary networking, protocols, or hardware;

- Able to offer additional capacity as demand increases, and less as demand falls ('elastic');

- Capable of only billing customers for the storage they use.

However, often the use of the term cloud and the range of its features deployed by providers or institutions can be variable, reflecting all or just some of the characteristics defined by NIST.

[1] See http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

It is often assumed that these cloud solutions only run in large data centres accessed by customers over the public internet. This is one model (called a public cloud), but there are others which allow installation locally and access over private networks (private clouds), or even the flexibility to combine public cloud solutions with local organisational storage, creating a hybrid cloud.

All providers of cloud infrastructure services are able to offer cloud-based storage solutions. Archives, however, typically have additional requirements beyond the simple availability of a place to store data files. These requirements may include specific concerns around data protection and the processing of personally identifiable information, or attitudes to risk and data loss which are both more conservative and longer-term than those displayed by social media companies, computer games creators, and some of the other domains already gravitating to the cloud. Generic providers of cloud services, such as Amazon, Google, Microsoft and others, do not typically address specific archival considerations within their basic offerings. However, a number of specialist providers have also emerged to offer value-added services. By using them it is possible for archives to layer additional processes and procedures on top of generic cloud services in order to build the systems that they require.

These options are explored in more detail in Section 2.3.

1.3 Cloud Storage and Digital Preservation

Digital preservation concerns the management of digital content over time to ensure ongoing access. It can be defined as: 'the series of managed activities necessary to ensure continued access to digital materials for as long as necessary, beyond the limits of media failure or technological and organisational change'[2].

This definition emphasises both the technical and organisational challenges involved in maintaining digital materials over time. It is important to recognise that the challenges are urgent but can be taken one step at a time: addressing current technology and organisations but ensuring you are in a position to pass on to the next generation of technology or staff when needed. That definition and the explicit approaches that follow from it in terms of being prepared for managing transition are particularly important when thinking about cloud services that are developing, evolving and contracted for short time horizons.

There is significant experience of digital preservation building up in archives and other memory organisations that is shared across the community and available to you via organisations such as The National Archives (TNA) and the Digital Preservation Coalition (DPC)[3]. This can assist you in acting on advice in this guidance and starting to work on digital preservation as a major part of your activities. Although much of that experience in archives is of relatively small-scale digital preservation, the tools, procedures and workflows are available and/or being developed that can support preservation of much larger volumes of digital material across the sector. Guidance on policies and documentation of procedures are also available[4].

[2] Definition adapted and updated from Beagrie, N. and Jones, M. 2001, *Preservation Management of Digital Materials: A Handbook* (British Library: London) p 10.

[3] See http://www.nationalarchives.gov.uk/information-management/projects-and-work/digital-preservation.htm; and http://www.dpconline.org

[4] See TNA Digital preservation policies: guidance for archives http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf

The growing volumes of digital materials requiring preservation in archives come from many different sources including those created or acquired in digital form from parent organisations and donors; or via digitisation of existing physical collections. These types of digital material may have different characteristics and preservation needs.

Digital archives are becoming more widespread, and the Open Archival Information System (OAIS) Reference Model (ISO 14721) provides a common set of concepts and definitions that can assist discussion across sectors and professional groups and facilitate the specification of archives and digital preservation systems. The OAIS model defines a broad range of digital preservation functions including ingest, access, archival storage, preservation planning, data management and administration. Digital preservation functions other than archival storage can be provided via the cloud although these are not the specific focus of this guidance.

There are also emerging systems for certification of digital archives such as ISO 16363[5] and ISO 16919[6] and the Data Seal of Approval[7] that in time could provide formal standards for accreditation of digital archives, and extend the Archive Service Accreditation Standard to digital only archives.

Cloud storage can offer those involved in digital preservation several areas of potential and promise:

- The flexibility of the cloud allows relatively rapid and low-cost testing and piloting of emerging service providers. There are already some pilot activities with these cloud services and opportunities for shared learning across the community;

- There is now much greater flexibility and more options in deployment of cloud storage services and therefore greater relevance to archives compared to earlier years (see Public, Community, Private and Hybrid clouds);

- There are potential cost savings from easier procurement and economies of scale, particularly for smaller archives. These are important at a time of financial pressures;

- Cloud services can provide easy, automated replication to multiple locations and access to professionally managed digital storage; in addition, the specialists can add access to other dedicated tools, procedures, workflow and service agreements, tailored for digital preservation requirements.

As business processes in organisations and use of archives increasingly become digital, digital preservation is a strategic current and future interest for the sector and individual archives. Cloud may be a component of required solutions and enable wider participation and collaboration.

Balanced against these areas of potential and promise however, there are areas where significant issues need to be understood by archives and addressed, particularly in terms of information security and potential legal requirements.

[5] ISO 16363: 2012, Space data and information transfer systems – Audit and certification of trustworthy digital repositories. Geneva: International Organization for Standardisation

[6] ISO 16919:2011, Space data and information transfer systems - Requirements for bodies providing audit and certification of candidate trustworthy digital repositories

[7] http://www.datasealofapproval.org/

1.4 Security and the Cloud

Security is frequently cited as a significant concern for those considering use of the cloud, particularly for sensitive, commercial, or personally identifiable information.

We should be concerned about the security of data, wherever it is stored, but it would be unrealistic to suggest that most cloud services are inherently less secure than most local data centres. The larger public cloud service providers, including Amazon, Microsoft, Rackspace, Google and others, invest significant sums in ensuring the physical security of their data centre buildings. They also employ teams of dedicated IT security staff, trained and concerned with ensuring that their systems are as secure as possible. Generic standards such as ISO 27001[8] and 27002[9] describe the steps to be taken in maintaining physical and online security, and also detail the steps to be taken in responding to breaches. Domain-specific standards such as ISO 27799[10] provide additional specificity in areas such as the storing of patient data. Operators of public data centres typically adhere to the guidance enshrined in these standards, and can reasonably be asked to describe the nature and extent of their compliance by sharing risk management plans, or completing regular external audit, etc.

These companies are high profile targets, and their servers are almost certainly under near-constant attack. Occasionally those attacks succeed, and services are adversely affected. But the security policies and procedures are, in all likelihood, at least as good as those employed in the local data centres typically used by archives.

Recent guidance[11] from bodies such as the European Network and Information Security Agency (ENISA) goes some way towards describing the manner in which cloud providers should inform their customers of data breaches.

Public sector systems can frequently store sensitive, confidential, or personally identifiable information, and the process of assessing the associated risks is comprehensively documented in Cabinet Office guidance on Technical Risk Assessment[12]. The result of this risk assessment will often be summarised in a simple statement that a product or service 'is accredited to IL2' or equivalent. Archives should normally have a clear understanding of their own statutory and policy requirements in this area, and can certainly expect prospective cloud providers to be in a position to demonstrate their own level of likely or actual accreditation.

[8] ISO 27001:2013, Information technology - Security techniques - Information security management systems - Requirements. Geneva: International Organization for Standardization

[9] ISO 27002:2013, Information technology – Security techniques – Code of practice for information security controls. Geneva: International Organization for Standardization

[10] ISO 27799:2008, Health informatics – Information security management in health using ISO/IEC 27002. Geneva: International Organization for Standardization

[11] See https://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/incident-reporting-for-cloud-computing/

[12] See http://www.cesg.gov.uk/publications/Documents/is1_risk_assessment.pdf

1.5 Legal Issues

Adoption of a digital preservation strategy utilising cloud computing inevitably brings with it a range of legal questions. Some of these may be familiar from the contexts of outsourcing or outstoring, whilst others are unique to the technologies involved in cloud storage. When considering the legal issues involved, it is perhaps helpful to begin by dividing them into three key categories:

- Any legal requirements in terms of management, preservation, and access placed upon archives and their parent organisations, by their donors and funders via contracts and agreements or via legislation by Government (e.g. accessibility, availability, information security, retention, audit and compliance, Public Records Act, etc.);

- Those legal obligations relating to third party rights in, or over, the data to be stored (e.g. copyright, data protection); and

- The legal elements of the relationship between an archive and a cloud service provider or providers (e.g. terms of service contracts and service level agreements).

The first category of issues will be largely familiar, although strategies for ensuring those obligations can continue to be met after a move to 'The Cloud' will need to be determined and adopted.

The second category will require consideration of those particular attributes of cloud computing that may hamper an archive's ability to meet its legal obligations in regard to particular types of data. These issues too will be broadly familiar including copyright-related questions, such as:

- Who currently owns the copyright in works to be stored in the cloud;

- Whether additional licence permissions may be required to address the technical aspects of such storage (e.g. the making/holding of additional transitory or permanent copies of a work);

- What permissions the cloud provider will need in order to provide the service without infringing (e.g. to reproduce and provide access to the material that is uploaded); and,

- Whether the cloud provider is able to use the data in the works for their own purposes (e.g. sub-licensing content to others for commercial purposes);

- Which party will own the rights in any data or works created from the original data (e.g. metadata or works generated from metadata).

In addition there are data protection issues, where the relevant data is, by itself, or if combined with other data accessible to the archive, 'personal data' (e.g. data that permits the identification of a living individual). Risks include those relating to:

- Outsourcing - placing data in the cloud is to effectively outsource an archive's data processing and this will raise governance and accountability questions (e.g. which party is responsible (statutorily or contractually) for ensuring legal requirements for data protection are observed, or appropriate data han dling standards are set and followed?);

- Off-shoring - allowing data processing to take place outside the UK increases risk factors and legal complexity, as issues of jurisdiction (whose courts can/will hear a case), choice of law (whose law applies) and enforcement (whether a legal remedy can be effectively applied) must be considered;

- Virtualisation - there are security risks in sharing machines, e.g. loss of control over data location, and who has access to it, through technical attacks that breach the isolation between 'virtual machines' on a cloud service allowing extraction of data;

- Autonomic technology - if technological processes are granted a degree of autonomy in decision-making, e.g. automatically adapting services to meet changing needs of customers and service providers, this may make it difficult to maintain consistent security standards, and to provide appropriate business continuity and back-up, and further obscure where data processing will take place within the cloud[13].

It is likely that the third category will be most problematic. The relationship between archive and cloud service provider needs therefore not only to reflect the requirements derived from consideration of the first two categories, but to do so in the context of a fast-moving service provision environment in which cloud service providers, the nature of services they provide, and the technologies utilised will both evolve, and inevitably be overtaken, at a rapid rate. When considering that environment, it is important to be mindful of three key elements:

- First, data held in archives must be expected to be both preserved and accessible beyond the commercial lifespan of any current technology or service provider;

- Second, an approach to addressing serious risks, such as loss, destruction or corruption of data that is based purely on financial compensation will not be acceptable, as this takes no meaningful account of the preservation and custodial role of archives; and,

- Third, in order to reinforce the criticality of the first two elements, explicit provision must be made for pre-defined exit strategies (e.g. synchronising content across two cloud service providers or an external cloud with local internal storage; or agreeing an escrow copy[14]), and effective monitoring and audit procedures.

These elements should be made clear at an early stage to both cloud service providers, and to legal advisors who, unless they are already familiar with digital preservation, are likely to approach negotiations with rather different perceptions of requirements of service provision and acceptable forms of risk amelioration.

Table 3 provided in section 7 as an appendix to the Guidance, lists legal points in greater detail for each of the three key categories above.

[13] See Pearson, S. & Charlesworth, A. 2009. Accountability as a Way Forward for Privacy Protection in the Cloud, HP Laboratories Technical Report (HPL-2009-178), p.2: **http://www.hpl.hp.com/techreports/2009/HPL-2009-178.pdf**

[14] Escrows are arrangements in contracts, whereby an independent trusted third party receives and in certain circumstances disburses, money, documents, software, digital publications, or other digital content for the transacting parties. Access to the escrow is triggered when pre-defined conditions in the contract are met, e.g. bankruptcy and loss of service of a provider.

1.6 Costs

Cloud storage services can achieve significant economies of scale and offer costs and features that can be attractive to organisations. Addressing legal and preservation concerns may inevitably have the effect of increasing some of the costs of achieving effective digital preservation via cloud services. A counter-balance to such increased costs may be found via collaboration on requirements, standards, support, and purchasing, making it a more cost effective proposition for cloud providers and intermediaries or consortia to offer services tailored more appropriately to the archive marketplace.

Cloud services are typically considered to be operational rather than capital expenditure: instead of buying computer equipment up-front, and writing its cost down over several years, users of cloud services are often billed retrospectively for the computing power, network bandwidth and storage space that they actually consume. This can be cheaper, especially for short or fluctuating workloads, but it requires organisations to think differently about the way their budgets are managed.

Generalist cloud services tend to bill each month for capacity that has actually been consumed. As a result it can be difficult to budget ahead, or to accurately predict the amount of data likely to be uploaded, stored, or downloaded; each of which can incur a separate cost. As a result, third party services such as Cloudyn and Cloudability have emerged, specifically catering to organisations which need to track, control, and predict cloud spending[15].

Some specialist cloud providers and intermediaries operate on longer subscription periods: months or years, rather than the minutes or hours of the generalist cloud services. As a result, their pricing and billing processes can be more amenable to some archives and their digital preservation budgets and funding requirements.

1.7 Summary of Key Issues

The Positives

- Cloud services can provide easy, automated replication to multiple locations and access to professionally managed digital storage and integrity checking. As a result bit preservation (durability) of digital information can be at least as good (or better) than can be achieved locally;
- Archives can add access to dedicated tools, procedures, workflow and service agreements, tailored for digital preservation requirements via specialist vendors;
- There are potential cost savings from easier procurement and economies of scale, particularly for smaller archives;
- The flexibility of the cloud allows relatively rapid and low-cost testing and piloting of providers;
- There is much greater flexibility and more options in deployment of cloud services and therefore greater relevance to archives compared to earlier years. In particular private cloud or hybrid cloud implementations can address security concerns over storage of more sensitive material perhaps considered unsuitable for public cloud;
- Exit strategies can be put in place to address archival concerns over provider stability and longevity or other change risks. For example synchronising content across two cloud service providers or an external cloud with local internal storage; or agreeing an escrow copy held independently by a trusted third-party;
- There are already some pilot activities with these cloud services and opportunities for shared learning across the community.

[15] Cloudyn: http://www.cloudyn.com; Cloudability: https://cloudability.com

The Negatives

- The Cloud is designed for flexibility and rapid change. Archives however are long-term. Cloud storage and service contracts need careful management through time to meet archive needs. Data held in archives must be expected to be both preserved and accessible beyond the commercial lifespan of any current technology or service provider;

- Cloud can be cheaper, but it often requires organisations to think differently about the way their budgets are managed. There are also different skills to IT service vendor and contract management that may involve re-training or recruitment costs;

- Public cloud services tend to bill each month for capacity that has actually been consumed. As a result it can be difficult to budget ahead, or to accurately predict the amount of data likely to be uploaded, stored, or downloaded (however some vendors can invoice you for an annual subscription based on volume);

- With cloud storage as in any form of outsourcing, it is important that archives exercise due diligence in assessing and controlling the risks. You need to ensure any legal requirements in terms of management, preservation, and access placed upon archives and their parent organisations, by their donors and funders via contracts and agreements or via legislation by Government; and obligations relating to third party rights in, or over, the data to be stored will be met;

- Use of cloud services will require archives to consider copyright-related questions including: who currently owns the copyright; whether additional licence permissions may be required; what permissions the cloud provider will need to provide the service; whether the cloud provider is able to use the data for their own purposes; and which party will own the rights in any data or works created from the original data;

- Use of cloud services may raise data security issues, where the relevant data is 'personal data' (e.g. data that permits the identification of a living individual), these include determining responsibility for securing data and audit of providers, as well as about location of processing and the extent to which risks incurred by automation of service provision can be addressed by contract;

- The legal elements of the relationship between an archive and a cloud service provider or providers (e.g. terms of service contracts and service level agreements) must be well defined and meet your requirements. This can be challenging as many cloud providers have standard SLAs and contracts to achieve commodity pricing and have limited flexibility on negotiating terms;

- Explicit provision must be made for pre-defined exit strategies and effective testing, monitoring and audit procedures.

Outcomes

The term 'cloud' can encompass a wide range of implementation models for archival storage. There is much that can be learnt from archives who have already piloted or moved to use of cloud storage. Several archives have been able to address the most widely held concerns over cloud services and find ways to successfully integrate cloud storage into their digital preservation activities. We profile a number of different approaches in the case studies listed in Section 5. Cloud storage is also being adopted successfully by a number of other sectors such as the legal profession, who have needs for maintaining confidentiality, data protection, and security[16].

[16] Silver Linings:cloud computing, law firms and risk Solicitors Regulation Authority November 2013 http://www.sra.org.uk/documents/solicitors/freedom-in-practice/cloud-computing-law-firms-risk.pdf

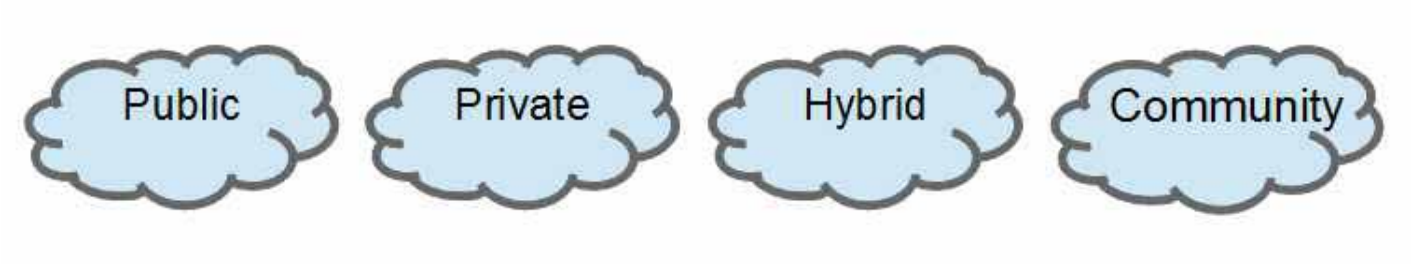
2 Step by Step Guide

2.1 Establishing Your Needs

Initially think about the capabilities you require rather than a specific technology, implementation, or product. Identify what are the 'must have' needs and what are the 'wants'. This may be an iterative process particularly if the archive is relatively inexperienced in digital preservation. Remember one of the strengths of cloud computing is its flexibility and rapid deployment at low cost so you can also pilot new procedures and cloud services to refine your requirements. Liaise closely with other staff and departments to draw on their expertise and to validate capabilities required and likely budget.

Establishing requirements can be a lengthy process, particularly for larger systems so do not underestimate time required for this. You may find some of the presentations from the DPC briefing day helpful in this process: Procuring Preservation Writing and Understanding Requirements in Digital Preservation[17]. Section 4 (Current Best Practices) in this guidance may also be helpful in developing thinking about your requirements.

2.2 Service Options



The concept of cloud computing encompasses a range of different scenarios, from providing scalable IT infrastructure that may be rented on demand (often called Infrastructure as a Service) through to the delivery of full-blown suites of software productivity tools via the web (often called Software – or Applications – as a Service).

It is often assumed that these cloud solutions are accessed by customers over the public internet. This is one model, but there are others which are also worthy of consideration.

2.2.1 Public Cloud

The public cloud is the most widely recognised manner in which cloud computing is used, with commercial services hosted in large data centres around the world, accessible over public networks to anyone with the means to pay.

Public cloud services typically realise economies of scale by sharing expensive hardware amongst customers (a 'multi-tenancy' arrangement, in which more than one secure virtual machine may run on the same physical server) and by refusing to negotiate non-standard service level agreements or sets of terms and conditions. These cloud providers offer a (often extensive) menu of configurations, and customers are required to select from the available set.

[17] See presentations linked from the agenda at http://www.dpconline.org/events/details/72-ProcuringDP?xref=78, particularly *Specifying Requirements: A technologists' view* (Angela Dappert, Digital Preservation Coalition) and *Procuring Preservation: hoops, hurdles and processes* (Susan Corrigall National Records of Scotland).

Public cloud infrastructure tends to be most cost-effectively used for relatively short-term activities; scaling services to meet peaks in demand, or running tasks of finite duration. With appropriate management and application, it can also be used for long-term activities such as storage.

For case studies of archives using public cloud see documents linked from summaries in section 5.1 Archives and Records Council Wales Digital Preservation Consortium, and section 5.2 Dorset History Centre.

2.2.2 Private Cloud

Large organisations, in particular, are successfully taking ideas from the public cloud and applying them inside their own data centres to create private clouds. By virtualising large sets of physical servers, and implementing processes to give employees the ability to requisition computing resources on demand, these organisations are able to replicate many of the public cloud's advantages whilst retaining direct control over hardware, data, and spend. Scale tends to matter in these use cases, as the organisation needs to have a large enough IT estate in order to support elastic scaling on demand, as well as ensuring that there is excess capacity from which employees can quickly request new virtual machines. Constrained IT capacity, or a traditional procurement process requiring a large administrative overhead and long lead times to process new requests, removes much of the value of a cloud-like approach.

The private cloud tends to be most cost-effectively used by organisations that already have significant investment in data centre space, equipment, and personnel. The effectiveness of the approach typically increases as the size of their existing data centre investment grows.

There is a growing trend amongst the specialist service providers to begin offering cloud-hosted or on-premise installation of their software to support archival storage and other digital preservation functions. In some cases (such as DuraCloud's proposed use of OpenStack) on-premise installation may be considered a private cloud. In others, it is simply installation locally within the organisational premises and network; strictly there may be no cloud attributes to this even if a superficially similar product is also offered via the cloud.

For case studies of archives using private cloud/local installation see documents linked from summaries in section 5.4 Tate Gallery and section 5.5 University of Oxford.

2.2.3 Hybrid Cloud

A hybrid cloud seeks to combine aspects of public and private, creating a fusion of the two. An organisation which typically runs its ecommerce system in-house, for example, might choose to temporarily add additional web servers from the public cloud in the run-up to the busy Christmas trading period. Core customer databases and billing systems remain in-house and under their direct control.

Hybrid cloud solutions tend to be most applicable to applications that anticipate significant but reasonably predictable fluctuations in load. For greatest benefit, applications should be designed from the outset with the expectation that they will need to run in this hybrid configuration. In an archival context, the scale and elasticity of a public cloud might be exploited to store large volumes of public data, or to run batch conversions of photo archives from one file format to another. The easier realisation of greater control and security in the private cloud might be used to store and process more sensitive personally identifiable material. An overarching system could be designed to leverage both of these capabilities within an apparently seamless whole.

For case studies of archives using hybrid cloud see documents linked from summaries in section 5.3 Parliamentary Archives and section 5.6 King's College London.

2.2.4 Community Cloud

A community cloud is a special instance of a cloud, optimised for or only available to a particular group of users. Amazon, for example, offers its cloud-based services in various regions around the world. Most of those are available to everyone, but the company also runs a version of their cloud that is only available to Federal and State governments in the United States. Architecturally, it is effectively the same as Amazon's public cloud services in that country, but access is restricted to a particular set of customers.

A public, private or hybrid cloud storage resource, procured on behalf of the UK archival sector and only contributed to by appropriate archives, might be considered an Archive community cloud.

There are no dedicated Archive community clouds currently in the UK. However aspects of the Archives and Records Council Wales Digital Preservation Consortium pilot could evolve towards this. See document linked from summary in section 5.1 Archives and Records Council Wales Digital Preservation Consortium for the case study.

2.3 Service Providers

There are a variety of ways in which archives might choose to deploy cloud storage capabilities, from simply using an existing provider of cloud-based storage as a direct replacement for local physical disk drives through to more complex arrangements which implement fault-tolerant archival workflows and policies (often in partnership with some third-party technology provider). In this Guidance, we consider two classes of cloud storage service provider: generalists (Amazon, Rackspace, Google, etc), and specialists (companies that address specific archival requirements, often by adding value on top of technology from the generalists).

There is no single provider in either of these two categories that clearly represents the best solution to all of the UK archival sector's requirements. Individual implementation choices will typically depend upon a wide range of factors beyond the scope of this Guidance, including such issues as budget, existing expertise and technical infrastructure, the principal purpose of an implementation project, etc.

We do not present an exhaustive list of every company and product feature in each category. We discuss those providers we believe are likely to be the most relevant to archives in the UK and selected generic features that illustrate some of the key issues that archives need to consider during any procurement process. The latest and most detailed information on products and services can be obtained from the web sites of individual service providers. We anticipate archives will identify their own requirements for generic and detailed features and select potential providers that may meet those requirements in a procurement process.

The US-based POWRR project has worked on a table (Tool Grid) arranged according to the OAIS Reference Model, in which various detailed digital preservation features are assessed for selected tools. This includes some cloud services. POWRR have since thrown their support behind COPTR, a Community Owned digital Preservation Tool Registry. This combines the form and function of the original POWRR grid with the far greater coverage of tools and sustainability provided by the COPTR data feed[18].

[18] http://www.digipres.org/tools/about/ For the Community Owned digital Preservation Tool Registry (COPTR), see http://coptr.digipres.org/Main_Page

2.3.1 Generalists

Generalist cloud storage providers include the oft-cited giants of cloud computing more broadly: companies such as Amazon Web Services (AWS), Google, IBM, Microsoft, and Rackspace. These global companies support a wide range of cloud computing use cases, and their cloud storage products typically permit customers to store data in particular geographies such as the European Economic Area (EEA), US, and UK.

The market for general cloud computing services is far larger than just the big names, though, and the EU and EEA are home to a growing number of European companies that cater to businesses cautious about trusting US-headquartered companies. GreenQloud in Iceland and CloudSigma in Switzerland, for example, both offer cloud storage solutions that are technically and commercially competitive with the products offered by their better known US competitors.

Generalist cloud providers tend to appear cheaper than the more specialist offerings discussed below. They also tend to support a much broader ecosystem of third party developers, consultants, and tool builders, making it likely that custom support or development effort will be more readily available, if needed. However, archival requirements do not always align well with the broader business requirements and technical choices of these companies. They are cheaper, they are more broadly available, and they are far more broadly used. But they may not always be able to meet a very particular archival requirement related to ingest, processing, long-term storage or maintenance of data.

Table 1: Examples of General Providers

Provider / Product	Choice of Locations	Speed of Access	Degree of Adoption	Costs	Security	Data Migration Out
Amazon Web Services (AWS) Glacier	EEA (Ireland and Germany) and Global	Typically within 3-5 hours	High	No initial costs. Billed for usage by the hour	Comprehensive accreditations[19]	Download standard formats by API, and move large data volumes on disk[20]
Amazon Web Services (AWS) S3	EEA (Ireland and Germany) and Global	Immediate, by widely adopted API	Very high	No initial costs. Billed for usage by the hour	Comprehensive accreditations[21]	Download standard formats by API, and move large data volumes on disk[22]
CloudSigma	Switzerland (EEA equivalent) and USA	Immediate, by API	High	No initial costs. Billed for usage in 5 minute increments	Suitable accreditations, some through hosting provider Interaxion[23]	Download standard formats by API
GreenQloud	EEA (Iceland) and USA	Immediate, by AWS-compatible API	High	No initial costs. Billed for usage by the hour	Suitable accreditations, some through hosting partner Verne Global[24]	Download standard formats by API
Microsoft Windows Azure	EEA and Global	Immediate, by API	High	No initial costs. Billed for usage by the minute	Comprehensive accreditations[25]	Download standard formats by API. US option to move large data volumes on disk not yet available in Europe[26]
Rackspace	UK and Global	Immediate, by Open-Stack API	High	No initial costs. Billed for usage by the hour	Comprehensive accreditations[27]	Download standard formats by API

[19] http://aws.amazon.com/compliance/
[20] http://aws.amazon.com/importexport/
[21] http://aws.amazon.com/compliance/
[22] http://aws.amazon.com/importexport/
[23] http://www.interxion.com/Documents/Case%20Studies/English/CloudSigma_Online.pdf
[24] http://www.verneglobal.com/news/corporate-news/verne-global-receives-industry-certification-iso-27001-information-security
[25] http://www.windowsazure.com/en-us/support/trust-center/compliance/
[26] http://www.windowsazure.com/en-us/documentation/articles/storage-import-export-service/
[27] http://www.rackspace.co.uk/certifications

2.3.2 Specialists

Although the generalist cloud providers discussed above do not, themselves, offer storage services tailored to the particular needs of the archival community, the nature of their basic cloud infrastructure does lend itself to use and reuse by a growing group of intermediaries. These specialist services tend to be slightly more expensive than accessing the generalist cloud storage directly. They also tend to operate on longer subscription periods; months or years, rather than the minutes or hours upon which Amazon, Google and others bill. Services like Arkivum, DuraCloud, and Preservica[28] are amongst the best-known of these specialist providers. These specialist intermediaries are able to take generalist infrastructure (often, but not always, from Amazon) and to layer archival workflows and processes on top in order to create something more recognisable to the archival sector. It is worth noting that whilst DuraCloud relies upon Amazon's cloud to deliver its core service, it currently only operates from Amazon's US-based data centres (but a future service from Amazon's Dublin (Ireland)-based data centre is also under consideration).

Specialists may offer a range of cloud-based digital preservation functions although we focus here specifically on that of archival storage. They can also offer cloud-hosted or on-premise installation of their software to support these archival storage and other digital preservation functions. Our case studies (see section 5) illustrate a range of archival storage types and software installation options that archives are testing and implementing to support digital preservation.

Some specialists also target their offerings at very tightly scoped use cases. The US-based Internet Archive, for example, offers a well-regarded service called Archive-It. This is explicitly designed to support the archiving of websites.

[28] Preservica is a fully owned subsidiary company of Tessella. The Preservica company name and re-branding of its products was adopted in April 2014.

Table 2: Specialist Providers

Provider / Product	Choice of Locations	Speed of Access	Degree of Adoption	Costs	Security	Data Migration Out / Exit Strategy
Arkivum 100 and 1+1	UK data centres	Access to tape-based storage, typically within 5 minutes of request by file system, API, or GUI	Moderate	Annual subscription, or paid-up fixed term contacts, based upon volume and duration	Certified to ISO27001 and audited on a 6 monthly basis	Company offers comprehensive escrow arrangement for its 1+1 product
Arkivum Onsite	Installed locally	Access to tape-based storage, typically within 5 minutes of request by file system, API, or GUI	Moderate	Annual subscription, or paid-up fixed term contacts, + hardware	Depends upon accreditations at host institution's data centre	Can have two Pods and an offline copy. Pods can be managed locally (or remotely by Arkivum)
DuraSpace DuraCloud	AWS data centres in USA	Immediate, by AWS API	Moderate	Annual subscription, based upon volume	As AWS - Comprehensive accreditations	Source records available for retrieval by API. Client can opt to sync to Rackspace as 2nd cloud service to AWS
Internet Archive Archive-It	US data centres	Immediate access by web User Interface	Moderate-High. Over 300 partners in 16 countries	Annual subscription, based upon volume	No formal accreditations?	Partner institutions can receive a copy on a hard drive or download their files directly from servers
Preservica Cloud Edition	AWS data centres in EEA (Ireland) and US	Immediate, by AWS API	Moderate	Annual subscription, based upon volume	As AWS - Comprehensive accreditations	Source records available for retrieval by API or user can copy home to their internal servers
Archivematica	Open source preservation software – cloud hosting of the software - is being tested in Wales in combination with cloud-based archival storage. Cloud hosting of the software is operational in Canada and in a joint ArchivesDirect service with DuraCloud					
Ex-Libris Rosetta	Rosetta digital preservation system not currently available as a cloud installation but cloud product release is under review					

2.4 Procurement Options

Owing to the variety of ways a cloud service may be purchased, this Guidance does not prescribe how an organisation should tender and contract for cloud storage. However the following provides general advice to assist you. Individual organisations must identify and follow their statutory and regulatory purchasing policies to ensure that the services are purchased using the correct procedures. Failure to purchase under the specific guidelines could lead to a serious issue possibly involving compensation to other potential contractors disadvantaged by incorrect purchasing processes.

2.4.1 Introduction to Framework Agreements and Contracts

Framework agreements are a type of agreement negotiated for the supply of common goods and services. It is essentially a contract, under which one or more suppliers have been selected to provide a particular set of goods or services following standard terms and conditions. Once awarded a framework agreement is made available for organisations to purchase from, this is often known as a 'call-off'. They are designed to:

- Speed up the buying process: by using a framework agreement purchasing organisations do not have to go through an often lengthy tendering process, e.g. full Official Journal of the European Union (OJEU) tendering, every time the requirements arise. This reduces the time required for tendering and associated costs;

- Improve prices by combining demand from a number of organisations to get a better discount from suppliers.

If the framework agreement is awarded to one supplier, then the purchasing organisation can simply call-off the requirement from the successful supplier as and when it is needed. Where the framework is awarded to several suppliers, there are two ways in which call-offs might be made:

- Where the terms laid out in the framework agreement are detailed enough for the purchasing organisation to be able to identify the best supplier for that particular requirement, the organisation can award the contract without re-opening competition.

- If not, a further mini-competition would be held between all the suppliers on the framework agreement who are capable of meeting that requirement, where framework buying rules allow. Each framework will usually have a buying process to ensure procurements are legal.

A potential disadvantage of a framework agreement for a purchasing organisation is that they can be relatively unresponsive to change. In a fast developing market such as cloud services there may be new suppliers and/or new solutions within the market that were not included when the framework agreement was initially set up. For this reason, some public sector frameworks such as G-Cloud are of relatively short durations to allow the successor frameworks to include new suppliers, new requirements, and reflect past experience of purchasers. Others such as the Janet Cloud Services may adopt longer timeframes to allow for the provider recouping the cost of a Janet connection or for the specific data needs and timeframes of educational and research institutions.

Framework agreements and contracts can be of shorter duration and successive in nature compared to traditional IT procurement. The procurement itself may not be as intensive initially but will be spread over a longer period. Purchasing cloud services via frameworks therefore can involve a change in process and mindset. Procurement may involve timely review and pro-actively watching the market and suppliers. Internal processes may need to evolve to keep evaluations as living documents, retaining, revising, and adding to previous knowledge. In addition there may be greater opportunity for archives and other purchasing organisations to be pro-active in encouraging suppliers and frameworks to reflect their emerging needs and experience as successive framework agreements develop.

There are two Frameworks currently that are likely to be particularly relevant to public sector archives: G-Cloud and Janet Cloud Services, these are described below.

Each Framework will have a pre-defined scope in terms of eligible institutions and each organisation will need to consider carefully the frameworks open to them.

If their requirement doesn't fit into an available framework agreement, organisations of course also have the option of traditional OJEU and local procurement processes.

2.4.2 CloudStore

CloudStore is an online marketplace where suppliers offer their services to the public sector via the G-Cloud framework. G-Cloud is the only UK public sector framework dedicated to Cloud services. It is a set of rolling annual frameworks that overlap by around 6 months, follow OJEU regulations, and allows the public sector to buy cloud-based services through a marketplace called the CloudStore. Most public sector archives or their parent organisations are eligible to use G-Cloud and only a few Places of Deposit in third sector organisations (e.g. heritage trusts or cathedral archives) are ineligible.

The CloudStore contains all the services currently on the G-Cloud frameworks. It is a searchable database of over 13,000 services split into 4 areas or Lots – Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) and other Specialist Cloud Services (SCS). Contracts under the current framework (at time of writing January 2014) have no minimum duration but a maximum of 24 months. Contracts can extend beyond the termination date of the Framework used, as long as the contract was taken out prior to the termination date, up to the 24 months contract limit.

Of the examples of service providers listed in section 2.3, Amazon, Arkivum, CloudSigma, Microsoft, Rackspace, and Preservica, are available directly and/or via other vendors under the current framework (at time of writing January 2015).

Further information on past and current Frameworks and service providers are available on the CloudStore website[29].

2.4.3 Janet Cloud Services (HE/FE)

Janet Cloud Services is maintained by Jisc and assists organisations in the education/research sector moving to cloud services by providing guidance, collaborative purchasing power, and due diligence with terms and conditions.

Janet offers Frameworks to access a range of IT services including cloud services. Typically they are available to UK universities and colleges and often to other institutions such as the research councils connected to the Janet network. Cloud services offered that are potentially relevant to archives and digital preservation in the sector (either directly or more likely indirectly through their institutions' use of them) include: the Cloud and Data Centre Framework (eight providers offered), the Shared Data Centre Space, and the Data Archiving Framework with Arkivum (a 'Data Archive to Tape as a service'). The Brokerage is also peered with Amazon Web Services providing managed bandwidth for Janet members connecting to the AWS cloud, upon which institutions can leverage the pay as you go AWS Direct Connect service. This gives them a high capacity, low cost network to connect to the technology resources and storage provided from Amazon servers in Dublin, Ireland.

The duration of the Frameworks are typically quite long-term – 4 years for the Cloud and Data Centre

[29] https://www.gov.uk/how-to-use-cloudstore

Framework, and 10 years for the Data Archiving Framework. The range of providers available is currently quite restricted compared to G-Cloud but they can provide sector specific features such as service providers with connections to the Janet network.

Further information on current Frameworks and service providers are available on the Janet Cloud Services website[30].

2.5 Developing a Business Case

Many archives will find that their organisation has an institutional template to be followed for presenting an internal business case. Understanding the processes within your own organisation is essential as this will not only determine which templates are relevant for you but will also inform how you proceed with your business case for digital preservation and any subsequent cloud procurement. The generic guidance and links below together with the case studies (see Section 5) will also be helpful, particularly if you are completing a business case for the first time.

The process of putting a business case together may at first appear daunting but it needn't be. The archive and the wider organisation in which it fits will already have in existence a lot of documentation, strategies, and policies that can feed directly into relevant parts of the document. The broader issues and benefits of the Cloud are outlined in this Guidance and in the sources listed in section 6. You may also have IT colleagues and business analysts who can assist you.

In addition, there are other support materials and approaches to assist in the development of the business case that can be drawn upon. For example, the Jisc -funded SPRUCE Project has produced a comprehensive toolkit to help practitioners build business cases to fund digital preservation activities[31]. Support is also available in other areas: the KRDS benefits toolkit offers useful ways to identify and present the benefits of digital preservation[32], and there are several tools that can help to identify and describe relevant risks and risk mitigation[33].

[30] https://www.ja.net/products-services/janet-cloud-services

[31] Digital Preservation Business Case Toolkit: http://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit

[32] KRDS Digital Preservation Benefits Analysis Toolkit: http://beagrie.com/krds-i2s2.php

[33] e.g. DRAMBORA (Digital Repository Audit Method Based on Risk Assessment): http://www.dcc.ac.uk/resources/repository-audit-and-assessment/drambora or Preservica's Digital Value at Risk calculator: http://preservica.com/resource/digital-value-at-risk-dvar-calculator

3 Future Developments

3.1 Overview

The public sector continues to find new ways in which cloud-based tools can cut costs, increase agility, or open up new opportunities. Cloud-based email, calendaring and collaboration tools continue to spread. The UK Government's adoption of a 'Cloud First' policy for public sector IT in 2013 will strengthen this trend. In future, when procuring new or existing services, public sector organisations will need to consider and fully evaluate potential cloud solutions first – before they consider any other option. This approach is mandated to central government and strongly recommended to the wider public sector[34].

The cloud is not – and will not become – the only manner in which public sector IT is implemented, but it continues to offer a number of advantages which buyers and sellers of IT will seek to exploit.

Archives have worked to develop digital preservation policies and procedures to grapple with the move of their core records from paper to various born-digital forms. Increasingly many of the digital preservation solutions and tools available to them will also be cloud-based and the number of archives using them can be expected to grow.

Specifically in the area of cloud-based storage of archival data, today's suppliers continue to iterate and evolve their products. More suppliers may move to offer versions of their product hosted in UK or European clouds. Suppliers are also looking at ways to meet demand for locally installed versions of their cloud-based offerings. In addition, cloud providers increasingly recognise the opportunity to tailor their offerings in order to meet the particular needs of specific sectors or industries. While many of the public cloud's strengths lie in offering essentially the same product to all customers, there are situations in which a group has particular requirements, the scale to make specialisation cost-effective, and the budget to pay for special treatment. The best-known example of this is currently Amazon's GovCloud[35], a special region of Amazon's cloud offering, reserved for authorised (US) Government use. Similar arrangements may prove feasible in Europe, and there is also interest in providing equivalents for heavily regulated industries such as healthcare. The UK market place for cloud-based storage for digital preservation therefore is likely to expand further in terms of products and providers.

Archives can play an important part in shaping this emerging marketplace by sharing experience across the community and with cloud service providers. Collaboration on requirements, standards, support, and purchasing will also benefit the wider UK archives sector. We hope this Guidance and future updates and initiatives will help support that work.

When looking to the future development of legal issues relating to cloud-based technologies two interlinked issues likely to play a significant role are data sharing and data protection.

In the former case, there is significant pressure for existing data to be made accessible to and usable by 'big data' technologies. In the public sector, existing data sharing initiatives are likely to accelerate, and there will be an increasing expectation that, in the absence of overriding commercial or social imperatives, data should be accessible, extractable and capable of reuse. In the private sector, there may also be pressure to permit access to data for hitherto unforeseen purposes – developments in the EU/UK copyright regime, both in terms of mandating accessibility to orphan works and making the regime more friendly to data mining indicate that there are clear incentives to seek more interactive engagement with digital archives.

[34] https://www.gov.uk/government/news/government-adopts-cloud-first-policy-for-public-sector-it
[35] http://aws.amazon.com/govcloud-us/

Expansion of data will, however, require a careful navigation of its interface with data protection, an area of law likely to be in flux into the near future. It is clear that the EU wishes to create a more harmonised EU-wide data protection regime. Key elements of that reform include requiring greater 'up-front' accountability to regulators and data subjects; and broadening accountability beyond data controller to data processors – a group hitherto largely ignored by the legislation. Both of these elements would inevitably impact the liability dynamic between cloud service providers and purchasers of those services, requiring a greater focus on the contractual and service agreements between the parties.

Additionally, the ease with which home and overseas national security and law enforcement agencies can access and process stored digital data will add further complications to the privacy and security debate.

3.2 Keeping up to Date

A second edition of this guidance will be published in Spring 2015 to disseminate anticipated updates to a number of the case studies and in the field of cloud storage and provider offerings.

Alongside updates to this guidance on cloud storage, The National Archives is funding the Digital Preservation Coalition to produce a revised online 2nd edition of the Digital Preservation Handbook[36]. It is anticipated this work will be completed in phases and released in modular sections over a period of 24 months. After its next revision, it is planned that this guidance will be integrated into and maintained with the Handbook and its wider treatment of digital preservation issues.

To keep up to date with these and other relevant developments, readers are encourage to subscribe to the digital preservation and NRA-Archives announcement email lists on JiscMail, and the The National Archives blog via its RSS feed[37].

[36] First published in a print edition: Beagrie, N and Jones, M 2001, *Preservation Management of Digital Materials: a Handbook* (British Library, London).

[37] To subscribe to the digital preservation email list see https://www.jiscmail.ac.uk/cgi-bin/webadmin?SUBED1=digital-preservation&A=1; for NRA Archives see https://www.jiscmail.ac.uk/cgi-bin/webadmin?SUBED1=archives-nra&A=1; for The National Archives blog see http://www.nationalarchives.gov.uk/rss/

4 Current Best Practice

Defining your Requirements

- Initially think about the capabilities you require rather than a specific technology, implementation, or product;
- Identify the organisational rationale and drivers for your digital preservation requirements e.g., is it public access, regulatory compliance, or for another purpose? Once you can articulate 'why', you can address the 'what' and 'how';
- Think about the range of digital preservation functions you need and how these might be provided and integrated with the archival storage you select;
- Your requirements may involve access to people as well as infrastructure. Specialists providers can be as much about using people with understanding of your institution/sector and accessing relevant specialist expertise as well as relevant infrastructure;
- Check the network(s) that can be offered to connect between the provider and you. What network connectivity exists, who pays, what security does it offer, is there enough bandwidth, is there redundancy?

Implementation

- Take baby steps at first. Start with pilots, proof of concepts, non-critical content, peripheral parts of the business, etc. Try out the cloud small-scale and see if it works in practice for you. Build experience and confidence. Scale up. Move to more critical areas of content as familiarity, understanding and trust in a solution(s) grows;
- Remember there are several installation models (public, private, hybrid or community) for cloud and flexibility for integration with local storage systems. Also bear in mind your digital archives may be very diverse and have different requirements e.g. impact levels, within them;

Risk Management

- Exercise due diligence in assessing and controlling the risks. You need to ensure any legal requirements in terms of management, preservation, and access placed upon archives and their parent organisations, by their donors and funders via contracts and agreements or via legislation by Government; and obligations relating to third party rights in, or over, the data to be stored (e.g. copyright, data protection) will be met;
- Conduct a Technical Risk Assessment (this can incorporate a Privacy Impact Assessment) and allocate confidential or sensitive material to storage with appropriate security levels;
- Test that the chosen solution(s) work: Test the exit plan; test speeds of access; test rates of getting data in or out; test availability of support. Effectively do the virtual equivalents of what you would do for physical storage e.g., fire drills, setting off the smoke alarm, testing what would really happen if it all goes pear-shaped;
- Make explicit provision for a pre-defined exit strategy should you need to move to another provider. Consider synchronising content across two cloud service providers or an external cloud with local internal storage; or agree an escrow copy held independently by a trusted third-party;
- Take references from other similar archives or organisations using the proposed provider(s).

Data Security

- Where appropriate, use software to automatically encrypt material during uploading to the cloud, using security keys that are under the user's control and not known to the provider to ensure higher security. If it is also necessary, because of its sensitivity and security requirements, to encrypt any material whilst it is stored in the cloud, ensure your careful management of the keys and be very aware of digital preservation risks involved if this is mismanaged;

- When considering the storage or processing of personally identifiable information, use only providers that are based in EEA countries or countries offering equivalent or greater data protection laws, and that can guarantee that data will not be held in jurisdictions that do not offer such protections;

- If personally identifiable information is to be stored outside the EEA, ensure that the proposed service can offer at least full Safe Harbour compliance;

Contracts

- Ensure the legal elements of the relationship between an archive and a cloud service provider(s) (e.g. terms of service contracts and service level agreements) are well defined and meet your requirements;

- The contracts show what commitments the service providers are really able or willing to make. If it's not in the contract or SLA then why not? Beware terms like 'designed for', 'aims to provide', etc.

- Check that the provider can offer a level of guaranteed uptime, data integrity, and continuity protection that is acceptable to the archive;

Standards

- Check that the provider can offer regularly audited information security that at a minimum is compliant with procedures such as those documented in ISO 27001:2013, Information technology - Security techniques - Information security management systems - Requirements;

- Use the Open Archival Information System Reference Model (ISO 14721) to provide a common set of concepts and definitions that can assist discussion across sectors and professional groups and facilitate the specification of archives and digital preservation systems. Do not treat it as a standard for preservation services;

- Encourage take-up of emerging certification systems such as ISO 16363 (Audit and certification of trustworthy digital repositories) and ISO 16919 (Requirements for bodies providing audit and certification of candidate trustworthy digital repositories), and the Data Seal of Approval that could provide formal standards for accreditation of digital archives;

Community Engagement

- Share experience across the community and with cloud service providers. This should be done on a number of different levels for greatest impact: within user groups for the specific services, with other archives, and with the broader digital preservation community;

- Collaborate on requirements, standards, support, and purchasing to make it a more cost effective proposition for cloud providers and intermediaries or consortia to offer services tailored to the archive market place;

5 Case Studies

This section provides brief summaries of case studies. They illustrate implementations by archives in a range of: sectors (local authorities, universities, and museums); different cloud deployment options (public, private, hybrid, and community); and service providers. The first five case studies have been compiled as part of the guidance and are **available on The National Archives website**. Information was provided by interviewees and approved by them for public release. Details are correct as of January 2015. The sixth (King's College London) points to pre-existing independently compiled open-access documents on the Web.

5.1 Archives and Records Council Wales Digital Preservation Group [PDF]

This case study discusses the experience of a cross-sectoral group of Welsh archives as they cooperated to test a range of systems and service deployments in a proof of concept for cloud-archiving. The proof of concept has examined the open source Archivematica software with Microsoft's Windows Azure; Archivematica with CloudSigma; Preservica Cloud Edition; and has begun testing Archivematica with Arkivum 100. It explains the organisational context of the consortium, the varied nature of their digital preservation requirements and approaches, and their experience with selecting, deploying and testing digital preservation in the cloud. It concludes with the key lessons they learned, and discusses current proposals to secure funding in order to move this pilot into operation.

5.2 Dorset History Centre [PDF]

This case study covers the Dorset History Centre, a local government archive service, and its procurement via G-Cloud and use of Preservica Cloud Edition. It explains the organisational context of the archive, the nature of its digital preservation requirements and approaches, its two-year pilot project using Preservica Cloud Edition, the archive's technical infrastructure, and the business case and funding for the pilot. It concludes with the key lessons they have learnt and future plans.

5.3 Parliamentary Archives [PDF]

This case study covers the Parliamentary Archives and their experience of procuring via the G-Cloud framework and running public cloud storage as part of their digital preservation infrastructure. For extra resilience/an exit strategy they have selected two cloud service providers with different underlying storage infrastructures. The archive is not storing sensitive material in the cloud and has a locally installed preservation system (Preservica Enterprise Edition) for this. As such it is an example of an archive using a hybrid set of solutions part-cloud and part-locally installed for digital preservation.

5.4 Tate Gallery [PDF]

This case study discusses the experience of developing a shared digital archive for the Tate's four physical locations (Liverpool, St. Ives, and two in London), powered by a commercial storage system from Arkivum. It explains the organisational context of the Gallery, the nature of their digital preservation requirements and approaches, and their rationale for selecting Arkivum's on-premise solution (Arkivum OnSite), in preference to cloud-based offerings from Arkivum and others. It concludes with the key lessons learned, and discusses plans for future development.

5.5 University of Oxford [PDF]

This case study covers the Bodleian Library and the University of Oxford, and their provision of a 'private cloud' local infrastructure for its digital collections including digitised books, images and multimedia, research data, and catalogues. It explains the organisational context, the nature of its digital preservation requirements and approaches, its storage services, the technical infrastructure, and the business case and funding. It concludes with the key lessons they have learnt and future plans.

5.6 King's College London

The Kindura project led by King's College London and funded by Jisc, sought to pilot the use of a hybrid cloud for research data management. It used DuraCloud to broker between storage or compute resources supplied by external cloud services, shared services, or in-house services. There is an earlier Jisc prepared case study[38] and a more recent open-access article on the project[39].

6 Sources of Further Advice and Guidance - Annotated Bibliography

6.1 Cloud – General

NIST. 2011, *The NIST Definition of Cloud Computing* (7 pages)
http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

NIST's work to define the scope of cloud computing remains widely cited, and offers a useful baseline against which to measure emerging services.

UCISA, 2011, *Cloud computing briefing paper* (7 pages)
http://www.ucisa.ac.uk/publications/cloud.aspx

This briefing paper highlights the reasons for considering cloud provision. Its main purpose is to assist members of IT/IS departments in discussions with senior management in Higher Education who may be less familiar with the concept of cloud provision.

6.2 Cloud and Digital Preservation

Aitken, B, McCann, P, McHugh, A and Miller, K, 2012, *Digital Curation and the Cloud, DCC* (30 pages)
http://www.jisc.ac.uk/media/7/C/1/%7B7C1A1FD7-44B4-4951-85A8-FC2C4CEB1564%7DCuration-in-the-Cloud_master_final.pdf

This 2012 report focused on the use of cloud services for research data curation. It provides some definitions of Cloud computing and examined a number of cloud approaches open to HE institutions in 2012.

Anderson. S, 2014, *Feet On The Ground: A Practical Approach To The Cloud Nine Things To Consider When Assessing Cloud Storage*, AV Preserve (7 pages)
http://www.avpreserve.com/wp-content/uploads/2014/02/AssessingCloudStorage.pdf

A white paper on cloud services, divided into nine topics and questions to ask. Vendor profiles against these nine topics will be available at a later date.

Convery, N, 2010, *Storing Information in the Cloud*, ARA and Aberystwyth University. (38 pages)
http://www.archives.org.uk/images/documents/Cloud_computing_report_final-1.pdf

This 2010 report presents an overview of cloud computing uses and challenges in relation to common recordkeeping practices, plus guidance for assessing risks and opportunities when outsourcing to the cloud. The authors also produced in 2010 a Cloud Computing Toolkit (83 pages), available at
http://www.archives.org.uk/images/documents/Cloud_Computing_Toolkit-2.pdf

[38] https://jiscinfonetcasestudies.pbworks.com/w/page/45197715/Kindura
[39] http://www.journalofcloudcomputing.com/content/2/1/13

Dionne, M, 2013, *Digital Preservation, Records Management in the Cloud: Challenges & Opportunities*, SAA
http://www.cmswire.com/cms/information-management/digital-preservation-records-management-in-the-cloud-challenges-opportunities-saa13-022147.php

This is a short conference report of discussion at SAA 2013. Archivists with recent experience of moving large collections to the cloud offer advice on what to ensure is in the SLA.

Instrumental, 2013, *Report on Digital Preservation and Cloud Services prepared for Minnesota Historical Society April 1, 2013*. (24 pages)
https://wiki.duraspace.org/download/attachments/34636606/Instrumental_MHS_Report_Final.pdf?version=1&modificationDate=1366054137863

This 2013 report reviewed and compared a number of cloud providers and services available in the US. It looked at security, data integrity monitoring and correction, cost, and preservation facilities. They also discuss a hybrid approach for the different types of material held, using cheaper solutions for the less critical data.

Mediasmiths International ,2013, *The Coming Storm? A report on the impact of cloud on broadcast Digital Production Partnership* (20 pages)
http://dpp-assets.s3.amazonaws.com/wp-content/uploads/2013/09/The-Coming-Storm.pdf

The Digital Production Partnership (DPP) is an initiative formed by the UK's public service broadcasters including the BBC, ITV and Channel 4. This report addresses the question of whether cloud technology and services (including those for archival storage and preservation) could benefit the production community.

POWRR *-Preserving [Digital] Objects With Restricted Resources*
http://digitalpowrr.niu.edu

A US based project testing digital preservation tools. The project concluded in December 2014. The evaluation included some Cloud based tools/services that are likely to be relevant to a UK audience. Results are now integrated with COPTR.

COPTR -Community Owned digital Preservation Tool Registry
http://coptr.digipres.org/Main_Page

COPTR is primarily a finding and evaluation tool collating in one place the knowledge of the digital preservation community on preservation tools. Currently there are 389 different tools described in COPTR.

6.3 Cloud and Legal Issues

McDonald, S. 2010, *Legal and Quasi-Legal Issues in Cloud Computing Contracts* (4 pages)
http://net.educause.edu/section_params/conf/CCW10/issues.pdf

This short paper published in 2010 covers the legal issues of cloud contracts from the perspective of an educational institution in the USA, so there is emphasis on managing the expectations of the vendor on the institution to control or be responsible for the actions of student users. The author notes that indemnification is critical in at least two areas: infringement of third-party intellectual property rights and inappropriate disclosure or data breach.

Trappler, T. 2010, *If It's in the Cloud, Get It on Paper: Cloud Computing Contract Issues*, Educause.
http://www.educause.edu/ero/article/if-its-cloud-get-it-paper-cloud-computing-contract-issues

This paper provides a good overview of the key issues to cover in a contract with a cloud service provider. It also contains example wording of clauses. However, as it is premised on US practice and law, it should not be solely relied on for guidance as regards to either the legal issues that may arise, or appropriate contractual solutions, in the UK/EU legal environment.

Marchini, R. 2011, *Cloud computing: a practical introduction to the legal issues*, London: BSI.

This book, published by the British Standards Institute in 2011, is targeted primarily at those who wish to provide or acquire cloud services. It aims to be a practical and accessible introduction to the legal issues for non-lawyers, and for lawyers who are unfamiliar with the technology and its specific issues. As such, the discussion of the technology is relatively light touch, and the discussion of the law not excessively encumbered with the minutiae of statute and case law.

Millard, C. (ed.) 2013, *Cloud Computing Law*, Oxford, Oxford University Press.

This book is an academic text developed out of a Cloud Law Project at Queen Mary University of London. It is designed primarily for postgraduate and practitioner readers who have a good understanding of the technology behind cloud computing, and who are familiar with key legal concepts.

Oppenheim, C. & Korn, N. 2012, *The No-nonsense Guide to Legal Issues in Web 2.0 and Cloud Computing*, London: Facet Publishing.

This book is aimed at information professionals working in public, academic or special libraries, archives or museums, who are working with, using or managing Web 2.0 or cloud computing applications. It is relatively basic, and targets those seeking to engage with cloud computing at the application rather than services level.

Cloud Standards Customer Council. 2012, *Practical Guide to Cloud Service Level Agreements*
http://www.cloudstandardscustomercouncil.org/2012_Practical_Guide_to_Cloud_SLAs.pdf

This document aims to provide a practical reference to help information professionals and IT analyse service level agreements (SLAs) from different cloud service providers.

6.4 Relevant Standards and Good Practice

6.4.1 Cloud Security

ENISA. 2013, *Cloud Security Incident Reporting*.
https://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/incident-reporting-for-cloud-computing/

The EU's Agency for Network & Information Security offers recommendations on the ways in which cloud providers and their customers should respond to – and report – security breaches.

ISO 27001:2013, *Information technology - Security techniques - Information security management systems - Requirements. Geneva: International Organization for Standardization*
http://www.iso.org/iso/catalogue_detail?csnumber=54534

ISO 27001 describes the manner in which security procedures can be codified and monitored. Conforming organisations – including most providers of cloud storage services – can be externally accredited and validated.

ISO 27002:2013, ***Information technology – Security techniques – Code of practice for information security controls. Geneva: International Organization for Standardization***
http://www.iso.org/iso/catalogue_detail?csnumber=54533

ISO 27002 provides guidelines on the implementation of ISO 27001-compliant security procedures.

ISO 27799:2008, ***Health informatics – Information security management in health using ISO/IEC 27002. Geneva: International Organization for Standardization***
http://www.iso.org/iso/catalogue_detail?csnumber=41298

ISO 27799 provides specific advice on implementing ISO 27002 and 27001 in the healthcare sector.

CESG Cloud Security Guidance
https://www.gov.uk/government/publications/cloud-security-guidance-introduction

This guidance provides advice to UK public sector organisations who are considering the security aspects of cloud services. CESG is the UK government's National Technical Authority for Information Assurance and advises organisations on how to protect their information and information systems against today's threats.

Cabinet Office, 2009, HMG IA Standard No. 1 – Technical Risk Assessment **(114 pages)**
http://www.cesg.gov.uk/publications/Documents/is1_risk_assessment.pdf

A detailed discussion and standard intended for Risk Managers and IA Practitioners who are responsible for identifying, assessing and treating the technical risks to ICT systems and services that handle, store and process government information.

6.4.2 Audit and Certification of Digital Repositories

APARSEN 2012, ***Report on Peer Review of Digital Repositories, APARSEN-REP-D33_1B-01-1_0.***
http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D33_1B-01-1_0.pdf

Lessons learnt to date from the process of repository certification have been usefully summarized by the APARSEN project in this report. It suggests although there has been considerable progress, arguably audit procedures are not yet fully bedded down and some issues remain for both auditors and repositories.

ISO 14721:2012, ***Space Data and Information Transfer Systems – Open Archival Information System (OAIS) – Reference Model, 2nd edn. Geneva: International Organization for Standardization***
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57284

The OAIS (Open Archival Information System) reference model is a conceptual framework describing the environment, functional components, and information objects associated with a system responsible for the long-term preservation of digital materials. As a reference model, its primary purpose is to provide a common set of concepts and definitions that can assist discussion across sectors and professional groups and facilitate the specification of archives and digital preservation systems. It has a very basic set of conformance requirements that should be seen as minimalist. OAIS was first approved as ISO Standard 14721 in 2002 and a 2nd edition was published in 2012. Although produced under the leadership of the Consultative Committee for Space Data Systems (CCSDS), it had major input from libraries and archives and has wide recognition across the archive sector. The term 'open' in OAIS is used to imply that the standard has been developed in open forums, and it does not imply that access to the archive is unrestricted.

ISO 16363: 2012, *Space data and information transfer systems – Audit and certification of trustworthy digital repositories. Geneva: International Organization for Standardization*
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56510

The Trusted Digital Repository Checklist (TDR) is ISO Standard 16363 and was published in February 2012. It is a revision of the well-known Trusted Repository Audit Checklist (TRAC). Many of the changes were structural, and it continues to address the same core areas.

ISO 16919:2011, *Space data and information transfer systems - Requirements for bodies providing audit and certification of candidate trustworthy digital repositories*
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57950

A supplementary standard (ISO 16919) to ISO 16363 is in preparation, on requirements for bodies providing audit and certification of candidate trustworthy digital repositories. Its preparation is led by a Consultative Committee for Space Data Systems (CCSDS) working group.

Data Seal of Approval
http://www.datasealofapproval.org/

DIN 2012, DIN 31644 *Information and documentation – Criteria for Trusted Digital Repositories*
http://www.nabd.din.de/cmd?level=tpl-art-detailansicht&committeeid=54738855&artid=147058907&languageid=de&bcrumblevel=3&subcommitteeid=112656173

In addition to the ISO standards developed by CCSDS, other formal initiatives in this area of archive certification have been the Data Seal of Approval (DSA), and the German Standard on Trustworthy Archive Certification DIN 31644.

European Framework for Audit and Certification of Digital Repositories
http://www.trusteddigitalrepository.eu/Site/Welcome.html

In 2010, the European Framework for Audit and Certification of Digital Repositories was established as a collaboration between the Data Seal of Approval (DSA) certification, the Repository Audit and Certification Working Group of the CCSDS, and the German Standards (DIN 31644) Working Group on Trustworthy Archives Certification. It aims to support an integrated framework for auditing and certifying digital repositories consisting of a sequence of three levels, in increasing trustworthiness.

7 Appendix

Please note, the information in this table is provided solely as general guidance on the legal issues arising from various aspects of digital preservation and cloud services and is not legal advice. An adviser-client relationship is not created by the information provided. If you need specific details pertaining to your rights and obligations, contract agreements, or legal advice about what action to take, please contact a legal adviser or solicitor.

Table 3 - Legal Issues

3.1: Legal requirements relating to management, preservation, and access of archival data in a cloud computing service (CCS)

Core Issues		Requirements	Legal Basis	Cloud-Specific Issues
Safekeeping of archival data	Management and maintenance	Data Ingest	For appropriate bodies, the Public Records Acts 1958 and 1967 or 2011 (Scotland); the s.46 Code of Practice under the Freedom of Information Act 2000; BIP 0008-1:2008 Evidential weight and legal admissibility of information stored electronically.	Standard IT management and maintenance requirements should be capable of delivery by a CCS. However, a CCS may not deliver other digital preservation and archive requirements as a standard offering. An institution should have clearly identified its essential, desirable and optional requirements prior to opening negotiations with a CCS to ensure that it can meet its operating objectives.
		Authenticity of data		
		Logical integrity of data		
		Reliability of data		
		Usability of data		
		Ability to refresh data		
		Ability to migrate data		
		Preservation of significant properties of data		
		Preservation of resource management metadata		
		Preservation of fixity metadata		
		Preservation of resource discovery metadata		
		Preservation of resource use metadata		
	Data Security	Authorised access, amendment and deletion powers	For appropriate bodies, the s.46 Code of Practice under the Freedom of Information Act 2000, the Data Protection Act 1998.	Data security is a common concern with CCS provision. The nature and scope of security required will depend heavily upon the legal obligations that the institution has under relevant legislation, agreements entered into by the institution with depositors, or other institutional warranties. An institution will need to assess the risks attendant upon its archival holdings, perhaps via a Privacy Impact Assessment (PIA)/ Technical Risk Assessment (TRA), before entering into negotiations with a CCS.
		Audit of access, amendment and deletion of data		
		Appropriate levels of encryption for dataset(s)		
		Prevention of accidental or unauthorised destruction, deletion or amendment of dataset (s)		
	Audit/ compliance	Audit of access, amendment and deletion of data	For appropriate bodies, the s.46 Code of Practice under the Freedom of Information Act 2000, the Data Protection Act 1998.	For legal responsibilities, such as data protection, it's often incumbent upon an institution not just to contract for data security, but also to audit the CCS for compliance with its contractual obligations. As CCS services are both outsourced and multi-user, providers may be unwilling to permit direct independent audit by each of its customers. Equally, an institution may not have the necessary in-house expertise to carry out a technical audit. It may, however, be possible to obtain a third party audit for compliance to ISO27001, or conduct a joint audit with other users of the service.
		Audit of compliance with institutional standards		

Continue on next page ▸

3.1: Legal requirements relating to management, preservation, and access of archival data in a cloud computing service (CCS)

Core Issues		Requirements	Legal Basis	Cloud-Specific Issues
Access to archival data	Mandatory data access rights	Access to personal data on request by a data subject within 40 days	Data Protection Act 1998	Where data is held in a low-usage CCS, then the need to service frequent statutory requests for information may prove expensive. If availability or performance criteria have not been established with CCS then access to data may be delayed.
		Access to information held by or behalf of a public authority on request by a 3rd party within 20 working days	For appropriate bodies, Freedom of Information Act 2000	
		Access to environmental information held by or behalf of a public authority on request by a 3rd party within 20 working days	For appropriate bodies, Environmental Information Regs 2004	
Transfer or Disposal of archival data	Destruction or transfer	Secure transfer of data to a third party for continued preservation	For appropriate bodies, the Public Records Act 1958, or the s.46 Code of Practice under the Freedom of Information Act 2000, the Data Protection Act 1998.	Transfer of archived data from the CCS to a third party organisation for ongoing preservation, may be hampered if the CCS cannot or will not provide the data in a viable format within a suitable timeframe and to a location other than that of the institution. The institution should have an exit strategy which is resilient in the event of bankruptcy or failure of the CCS.
		Verified destruction of data		If the institution does not have unfettered ability to access, edit and delete its data on an ongoing basis then granular control of data deletion will be problematic. The institution should be able to audit/verify actual destruction of data.

3.2: Legal requirements arising from obligations to or from third parties that may be affected by use of a cloud computing service (CCS)

Core Issues		Considerations	Approaches	Cloud-Specific Issues and questions
Intellectual property rights	Depositor rights	There may be various intellectual property rights in archived data, including copyright and related rights, trade marks and design rights owned by the archive, by depositors, or by other 3rd parties. While ownership of such rights works to be stored in the cloud would usually continue to rest with the original owner, identification of ownership at ingest and prior to cloud storage will be important, as will establishing the extent of the liability for breaches to be accepted by both CCS and archive via warranties and indemnities.	Depositor assignments, licences and waivers upon data ingest	Does the archive's contract with the CCS provide adequate provisions for security, access restrictions and audit to meet its stated obligations to depositors? If there is a CCS contract breach that results in infringement of depositor IPRs, who is liable?
	Third Party rights		Depositor warranties of non-infringement, and grant of 3rd party licences upon data ingest	If the CCS is notified by a third party that it is hosting IPR-infringing material on its service, what are its obligations to the archive? Can the CCS remove material from its service without notification to the archive, or only after notification?
	CCS rights	Possible accrual of rights either in archived data or related metadata from processing or other use of archived data	Agreement that the CCS acquires no rights to any IPR in the archived data, or in metadata or other outputs obtained by processing or other use of archived data	A clear statement to that effect should be incorporated in the CCS contract and/or set out in a specific licensing agreement.

Continue on next page ▸

3.2: Legal requirements arising from obligations to or from third parties that may be affected by use of a cloud computing service (CCS)

Core Issues		Considerations	Approaches	Cloud-Specific Issues and questions
Data protection rights	Jurisdiction	Personal data may not be transferred out of the European Economic Area unless data subjects have consented, or the country to which the data is sent has 'adequate' data protection laws, or there is a suitable contract between the data controller and a third party outside the EEA who is receiving the data	Some CCSs can offer services which ensure that any data stored on their servers does not leave the EEA.	A clear geographic limitation should be included in the CCS contract, and where the CCS leases capacity from other CSSs on an ad hoc basis, the original CCS should remain directly responsible for all terms of the contract with the archive, including the geographic limitation clause.
	Protection of data subject rights	Data subjects must be able to effectively exercise their rights under the Data Protection Act 1998, including data subject access.	Prior to choosing a CCS, the archive should evaluate whether it will be storing personal data, i.e. data relating to an identifiable, living individual.	If personal data is held, the archive should ensure that it is able to access individual personal data within the 40 day time limit, and that it can produce the data in a format meaningful to the data subject. The archive should also ensure that where data is held in a low-usage CCS, the likely cost of access to fulfil statutory requests for personal data is not prohibitive.
	Access for exempted uses	Where a disclosure is required by, or under, any enactment, by any rule of law or by the order of a court other organisations may have the right to access personal data under the Data Protection Act 1998, i.e. for law enforcement purposes.		If personal data is held, the archive may wish to ensure that it is able to access individual personal data, and that it can produce the data in a format meaningful to the requestor. The archive should also ensure that where data is held in a low-usage CCS, the likely cost of access to fulfil legal requests for personal data is not prohibitive.
	Breaches of data protection law	Where there is a breach of data protection law, the data controller remains liable for the breach, even where a data processor is holding the data on their behalf.		Does the archive's contract with the CCS provide adequate provisions for security, access restrictions and audit to meet its stated legal obligations to data subjects? If there is a breach of DP law, is the CCS obliged to notify the archive? Under what circumstances will the CCS be contractually liable to the archive for costs relating to liability for breach of data protection law?

Continue on next page ▸

3.2: Legal requirements arising from obligations to or from third parties that may be affected by use of a cloud computing service (CCS)

Core Issues		Considerations	Approaches	Cloud-Specific Issues and questions
Defamatory or illegal content	Defamation - Jurisdiction	Under UK law material is defamatory if it is untrue and has caused or is likely to cause serious harm to the reputation of the claimant. For an action to be brought in the UK there must be a clear link with the UK jurisdiction e.g. material was viewed by a significant number of persons in the UK, and the claimant had a reputation capable of harm in the UK.	Depositor warranties and indemnities, upon data ingest. Prior to choosing a CCS, the archive should evaluate the risk that defamatory or illegal content may be included in archived data.	In which jurisdiction(s) will data in the CCS be accessible? To whom will the archived data be accessible, e.g. will the staff of the CCS have access to the data in unencrypted form?
	Liability in the EU	Under UK law, rapid removal or redaction of defamatory material will reduce the extent of the liability, as there is less damage to the claimant's reputation. In addition, under the e-Commerce Regulations 2002, where a CCS stores information provided by an archive, it will not be liable for damages or any criminal sanction arising from that storage if it does not have actual knowledge of unlawful information; or is not aware of facts or circumstances from which it would have been apparent that the information was unlawful; or once it has such knowledge or awareness, it acts expeditiously to remove or to disable access to the information.		Can defamatory or illegal material be easily removed or redacted by either the archive or the CCS to limit potential liability? If notified of defamatory or illegal material, what are the CCS's obligations to the archive? In the event that defamatory or illegal material is not removed or redacted by the CCS, or the archive, under what circumstances will the CCS be contractually liable to the archive for costs relating to liability?

3.3: Key Contractual and Service Agreement issues relevant to use of a cloud computing service (CSS)

Issues	Key elements	Archive Questions	Possible CCS questions
Availability of service	Target uptime of service	What target uptime meets your institution's specific needs? Can CCS claims about service levels be validated against reports of problems from other customers?	How do you define 'uptime'? What are the specific mechanisms for calculating compliance with service level agreements? What are the penalties open to institution if the CCS fails to meet service level parameters – credit, financial penalties, termination of contract?
Performance	Latency / speed of response of service	What is the maximum response time that meets your institution's specific needs?	How do you define 'response time'? When does failure to meet performance criteria become 'downtime'?
Functionality	Sustainability of essential architecture, features or services	Are there elements of a CCS that are essential to your institution's specific needs?	Are existing features or services guaranteed for the lifetime of the service? How much notice will be given of changes to features or services? Will the institution have the ability to comment in advance on changes to architecture, services or functions?
Vendor Outsourcing	CCS may lease processing and storage capacity from other CCSs	Do your institutional requirements (location of data, audit etc.) place a restriction on vendor outsourcing?	Does the CCS lease capacity from other CCSs? What criteria are used to determine suitable external providers? Are these criteria audited? Does the CCS remain directly responsible for all terms of the contract, regardless of outsourced functions?
Data protection	Access to data, encryption	Has your institution undertaken a Privacy Impact Assessment (PIA)/ Technical Risk Assessment (TRA)? What type of encryption process is appropriate for your data?	Is access to the institution's data limited to CCS's authorized employees? What encryption standards are supported by the CCS? At what points will the data be encrypted/ unencrypted? Are encryption keys controlled by the institution or CCS? What is the process for handling legal requests for access to data by third parties?
Security	Access control and replication of data	How important is the level of data security at the CCS to your institution? Does your institution have the capacity to evaluate CCS security on an ongoing basis?	What authentication and access controls exist within media, applications, operating systems and equipment? Can the CCS replicate and continuously update the institution's data at multiple locations? Can the CCS provide real-time data streams from intrusion detection systems?
Monitoring and Audit	Third-party audits and/or certifications, right to periodic on-site inspection	Does your institution require outsourced services to conform to particular standards, e.g. ISO 27001 or 27002?	What are the CCS's infrastructure and security specifications? Can these form part of the contract as the minimum infrastructure and security requirements?
Ownership of data and metadata	Ensuring ownership of original data and data generated by processing etc. is clear	What data will be stored in the CCS? Will additional data or metadata be generated in or by the CCS?	Is ownership of stored data clearly stated in the CCS contract? Does the CCS claim rights in the results of any data processing that occurs on its system?

Continue on next page ▸

3.3: Key Contractual and Service Agreement issues relevant to use of a cloud computing service (CSS)

Issues	Key elements	Archive Questions	Possible CCS questions
Geographic location of data	Legislative, regulatory, sectoral or institutional requirements	Does your institution process data that needs to be restricted to a particular geographic location?	Can data be processed in specific geographic locations? If other CCSs are used to provide capacity, can they meet this criterion?
Disposition of data	Customer's right to access, edit and delete its data on an ongoing basis	Are there situations that require immediate access to your data? Does your institution have specific disposition requirements e.g. deletion under a retention schedule?	Can the CCS provide procedures and timelines for time-sensitive access that meet specific institutional needs?
Portability of data	Ability to transfer data to a different CCS and remove from existing CCS	In what format and time-frame does your institution need the data in order to move between CCSs?	What are the CCS's common formats for data return/retrieval? Can the CCS provide the data in a specified format within a specified timeframe to a specified location?
Change Management	Process for updates or new services	How does your institution wish to be notified of updates or new services? How much notice of changes is required?	Can the CCS provide notification in the form required, and in the timeframe specified? Is the institution able to request or require that the CCS provide particular updates?
Changes to terms and Conditions	CCS terms and conditions may vary over time	What parts of the terms and conditions are to your institution's specific requirements?	Are the CCS's active terms and conditions at the time of contract signature locked in for the duration of the contract?
Problem Identification and Resolution	CCS technical support provision, e.g. call centre	What level of support does your institution require? Within what timeframe must problems be resolved?	What contractual assurances relating to the quality and continuity of key provider personnel can the CCS offer?
Data Breach	CCS obligations if the customer data is accessed inappropriately	What data will be stored in the CCS? Does this carry particular risks if accessed without authorisation?	What constitutes a 'data breach'? What are the CCS's data breach obligations (e.g. requirement to notify, timeframe, and provision of details)? What indemnification is provided by the CCS?
Disaster Recovery	Worse case recovery commitment		What processes and safeguards does the CCS have in place to protect customer data and services in the event of system failure? Can we test recovery procedures?
Dispute Mediation	Escalation process, consequences		

Continue on next page ▸

3.3: Key Contractual and Service Agreement issues relevant to use of a cloud computing service (CSS)

Exit Strategy/ Termination	Notice, consequences, data access, data deletion (see Access and Portability, above)	How much time and what information will your institution need to in-source/re-source the service, regardless of the reason for termination? For what reasons might your institution wish to terminate the agreement?	What is the minimum notice period that the CCS will accept? What grounds for termination are available under the CCS contract? Are there termination charges? How long after termination does a customer have to recover data before deletion? Can the data be auditably deleted from the CCS after termination of the contract? Can we test exit procedures?
Change in status of vendor	CCS goes into liquidation, or becomes involved in a merger or buy-out	What provision will you make for resilience to other events such as bankruptcy of a CCS?	In such circumstances, what are the responsibilities of the CCS and transferability of contracts or contract terms? Does the CCS offer third party escrow as part of its service provision? If not, can the customer sync data to a local server or a second CCS?
Contract Renewal	Timing of renewal, renegotiation of terms	How has the market and procurement frameworks and contracts evolved since? How has the CCS performed in the existing contract?	Does the CCS require notice of non-renewal within a set period before contract expiry? Is there an opportunity to renegotiate terms prior to renewal?
Cost	Initial or upfront costs, maintenance and continuation costs, renewal costs, volume commitments, caps on the increases in costs permitted over time	Is pricing and value still competitive? Can we obtain better value via new frameworks or new collaborations?	
Legal	Law governing contract	Does the institution have the capacity to engage in litigation outside its own jurisdiction, if necessary?	How and where will any legal disputes be settled? Can the CCS comply with laws/regulations of importance to the institution, even if these are not those of the CCS's jurisdiction?