A Selective Literature Review on Digital Preservation Sustainability

Lorraine Eakin, Amy Friedlander, Roger Schonfeld with contributions by Sayeed Choudhury

1.0 Introduction

Our goal in providing this literature review is to provide a baseline understanding of the current state of research into and practice in the sustainability of digital preservation, particularly regarding the concrete components that drive costs in the area of digital preservation. Part of this endeavor includes determining whether any important gaps in the literature still exist and if so, to highlight those areas so that appropriate future work can be undertaken. Given this focus on costs, we excluded numerous excellent papers and studies that examine technical aspects of digital preservation, case studies of individual projects, economic sustainability more broadly, and other related topics. Where possible, some of the more well-developed bodies of literature (for example e-journals specifically or scholarly communication generally) are acknowledged by reference. But given the focus on the sustainability of digital preservation, some very good work focusing on broader sustainability issues may not be reflected here or it may receive only cursory notice. This is not meant to diminish the value of that work but rather to define a sufficiently narrow scope for fruitful discussion of costs and to offer a set of concrete cost elements that can frame future research questions.

Although maintaining a sustainable preservation initiative admittedly requires more than just understanding costs (Currall & McKinney, 2006), understanding one's preservation cost structure is nonetheless paramount for managing sustainability issues. Costs influence incentives, and incentives determine who will be willing to support preservation initiatives in both the short- and long-term. Likewise, gaining control over the structure of incentives can ensure more successful business models and funding structures. Few are willing to pay for a preservation initiative without knowing how much it costs and how the costs are distributed. Hence, costs are a necessary, if not sufficient, component of a viable sustainability plan. In the following discussion we consider the issues related to costs and incentives (2) and summarize the early (3) and more recent (4) literature on costs, followed by a detailed comparison of available data and a discussion of the constraints on any comparative analysis (5). Finally, several observations, including a discussion of gaps in prior work, are offered in the concluding section (6).

2.0 Costs, Incentives, and Who is Willing to Pay?

This section examines several studies that provide somewhat more general discussions either because of the strong impact they have had on current sustainability dialogue or because they point to concepts that fundamentally drive cost-related sustainability efforts. One of the earliest (and most influential) systematic treatments of digital sustainability issues occurs in Lavoie's (2003) white paper on incentives. In this paper, he examines the community objectives associated with digital preservation and the organizational incentives that might motivate their realization, concluding that objectives and incentives are often not well-aligned. This thoughtpiece grew out of the realization that there are three fundamental stakeholder positions in digital preservation: the rights-holder, the archive (or more generally the custodial or stewardship institution), and the beneficiary. These roles are frequently split, both organizationally as well as temporally, and the particular combinations of roles in any given situation will influence whether or not underinvestment in preservation efforts is likely to occur. This framework provides a coherent setting within which incentives might be designed that could motivate a level of investment in digital preservation that support shared objectives.

In 2005, the Australian Partnership for Sustainable Repositories released Kevin Bradley's economically focused sustainability discussion paper (Bradley, 2005), which called attention to the importance of the organizational authority structure – that is, to whom in the organization the repository initiative reports – and to the importance of the sustainability of that larger organization. These topics have been all too often overlooked or tacitly assumed. Bradley noted that although more costing models are being developed and more cost data being collected, the uncertainty of long-term costs remains problematic. Moreover, he also observed that although it has been (relatively) easy to obtain the one-time set-up costs associated with establishing a digital preservation solution, the ongoing operating costs critical for long-term sustainability typically have been harder to secure. He endorses Lavoie's analysis of organizational incentives and recommends significant additional research investments in this area. Finally, he recommends additional attention to the questions of valuation necessary for developing policies for selection and the preservation life-span for given objects or collections.

A critical element in understanding incentives is the notion of value, which is a substantial challenge for digital preservation, as it is more generally with valuation of all intangible assets. When unable to determine value with confidence, it becomes virtually impossible to analyze whether the costs associated with digital preservation are worthwhile expenditures in any given circumstance. Hunter (2006) called attention to the connection between digital preservation and valuation of intangible assets and recommends the Balanced Score Card methodology for communicating value strategically where it is not possible to do so in financial terms. The espida project (espida, 2007) used the Balanced Score Card methodology as a technique for communicating value to funding decision makers to justify investment in information technology projects from the perspective of an individual institution The developers of this methodology note that changes in value over time, away from the operational and towards the historical, often coincide with a shift of the principal value to outside the originator organization (Curall & McKinney, 2006). They offer a handbook that recommends detailed methods for evaluating and communicating the costs and benefits associated with information intangibles, including those intangibles associated with digital preservation projects.

A second analysis by Lavoie (2006) reflects the implicit relationship between the value of preservation and its costs, which can be usefully divided into "fixed costs," which reflect activities that take place upfront, and "operating costs," which recur; the ongoing costs of sustaining the repository to achieve long term preservation are subsumed into the latter. Lavoie identifies five basic preservation strategies: media refreshment/migration; format migration; normalization; emulation; and technology preservation. Each strategy is likely to have a different fixed/operating cost profile, and longer-term costs are difficult to determine because they may vary based on “the pace and direction of technological change as well as user expectations.” Thus it is impossible to develop realistic cost estimates until a strategy is chosen. Ultimately, this article recommends giving significant weight to what is termed the “other” cost – the cost of inaction, of losing culturally significant materials. In this respect, Lavoie focuses on how to determine the value of preservation. He surveys the non-market-valuation techniques that have been applied to library operations, most prominently to the British Library, and suggests that similar techniques might productively be utilized to develop valuations for digital preservation.

As the handful of studies discussed in this section illustrates, costs are a component of sustainability models that must also consider value and the context in which the understanding of value is reached. This has a cultural or institutional dimension, that is, the communities who assign value to objects. It also has a technical dimension, namely, the economic and financial techniques that might be employed to quantify a notion of value. As Bradley (2007) remarks: “Clearly it is not possible to preserve digital information without a sustainable organizational, economic, social, structural, and technical infrastructure, nor is it sensible to preserve material without sustained value” (p. 157).

3.0 Early Work on Costs: It’s More than Technology

Beginning at least in the mid-1990’s, much of the attention focused on economic issues within the digital environment revolves around the rising costs of scholarly communication (Baumol & Marcus, 1973; Cummings, Witte, Bowen, Lazarus, & Ekman, 1992; DeGennaro, 1987; Economic Consulting Services, 1989; Getz, 1992; Peek & Pomerantz, 1998) or the economics of digital libraries (Connaway & Lawrence, 2003; Galvin, 1996; King, Boyce, Montgomery, & Tenopir, 2003). Researchers began commenting on the economic factors of digital preservation around 1998, with Beagrie and Greenstein’s JISC-funded addition to the Electronic Libraries Studies Programme, *A Strategic Framework for Creating and Preserving Digital Resources : A JISC/NPO Study Within the Electronic Libraries (elib) Programme on the Preservation of Electronic Materials* (1998) and Hendley’s *Comparison of Methods and Costs of Digital Preservation* (1998), also funded by JISC. The former document, although a broad framework for creating and preserving digital resources, contains a number of references to the cost factors that exist within the lifecycle of digital resources and which a digital preservation policy-maker should consider to ensure cost-effective preservation of these resources. The latter document was the result of the JISC-funded workshop on the *Long Term Preservation of Electronic Materials* held at Warwick in November, 1995. It deals specifically with assessing the relative costs of various technical strategies then being discussed to handle the long-term preservation of digital resources, namely, technology preservation, technology emulation, and digital information migration.

Early attempts did not provide explicit cost calculations but did begin to assess the functional phases in the preservation lifecycle (beginning with data creation), and to delineate the particular cost factors associated with each step in the lifecycle. For example, Hendley outlines seven areas which must be assessed in order to come up with a clear picture of preservation costs: data creation; data selection and evaluation; data management; resource disclosure; data use; data preservation; and rights management. These early efforts also noted that a key difficulty in assessing digital preservation costs lies in distinguishing between the preservation stage and other stages, given that the preservation costs may be impacted by decisions made much earlier in the lifecycle of the digital resource, perhaps even at the time of its creation. Thus, Russell and Weinberger (2000) point out that distinguishing between the costs of preservation and the costs of access can be problematic. This distinction has proved valuable: On the one hand, preservation and access represent different technological and management processes and therefore have different costs attached to the activities. They are linked, however, by the fact that access to the material, even under highly restricted conditions (such as the need to protect confidentiality or rights holder interests), confers value on the stored information.

Ashley (2000) outlines the cost factors in digital preservation, providing guidance to archivists interested in assessing their own digital preservation costs. Likewise, Sanett (2002) presents a framework for an organization interested in developing its own cost model, highlighting and elaborating various preservation-specific factors. She reviews the literature noted above, as well as cost discussions on topics like the costs of reformatting (Kenney & Rieger, 2000), archival repository design (Crespo & Garcia-Molina, 2001), and the costs of emulation versus migration (Granger, ND). However, none of the projects she cites provides detailed, concrete cost estimates for the preservation of digital resources. In fact, most early literature related to the costs of digital preservation focuses exclusively on comparing technical methods, such as emulation vs. migration, and tend not to identify the costs specifically attributable to preservation (Dollar, 1999; Holdsworth, 2001; Lorie, 2001; Rothenberg, 1999; Rothenberg & Bikson, 1999; Wheatley, 2001). Cumulatively, these early studies reflect growing awareness of the complexities of parsing cost into its components. They also begin to show awareness of the relationships between technologies and the full gamut of activities required to achieve long term preservation.

4.0 More Recent Efforts: Lifecycle, Initial Costs, and Costs over Time

This section concentrates on the specific cost elements and calculations provided by newer studies that orient themselves directly to digital sustainability. These are summarized in Table 1 of this review. For the most part, when seeking to develop detailed cost assessments, organizations have had to fall back on their own data. This is reflected in the literature by a series of cost models and assessments that are largely atomistic. Definitions of terminology reflect the conditions of the specific projects. The authors, who wish to publish project updates, typically do not attempt to create economic "crosswalks" between their and others' frameworks. Even for those projects that explicitly build from earlier work, within any given project the nature of costing activities is generally focused upon only a small subset of activities within the digital preservation lifecycle (for example, storage costs). Consequently, these analyses are isolated, with differences between them stemming largely from disparities in context, terminology, and choice of inputs.

Two important exceptions to this generalization are (i) the cost model developed by the LIFE (Lifecycle Information for E-Literature) project (http://www.life.ac.uk/) (McLeod, Wheatley, & Ayris, 2006) and (ii) the recently published model developed by Beagrie, Chruszcz, and Lavoie

(2008). The LIFE project conducted an extensive review of prior work to isolate the various components that prior studies had identified; existing studies are described and analyzed in detail in Watson (2005). The project, described in more detail in section 5.6 of this document, has developed a comprehensive model, geared toward library oriented operations, that provides a list of elements for costing a digital library project or initiative. It has recently completed its second phase (Ayris et al., 2008), during which the model was used to assess the costs of a series of digital publication projects (http://www.life.ac.uk/2/).

DESCRIPTION	COST
2001	
Roquade Project *Dekker et al.*	**Basis for assessment**: Experiential estimates, published literature reports **Unit of measurement**: Cost per information item • Personnel costs of assigning metadata: approximately 10 euros • Processing SIP's: approximately 10 euros per information item • 5,000 items per year added: 6 PC's with a network card and AV facilities: 1500 euros each + professional serer: $5000 euros • Total hardware costs: approximately 32,000 euros, depreciated over 4 years • Software and licensing fees: 15,000 euros per year using proprietary software • Maintenance support costs: 2,000 euros per year • Technical support: 0.2 FEs = 9,000 euros per year • Data refresh every 5 years @ 1 euro per MB; if DIPs are kept for 20 years and DIP is about 500 kB, cost - about 2 euros per information item, that is, 10,000 euros per year for all information items • Total per information item costs: 29 euros per item
2003	
Harvard Depository **Chapman** *Format:* Microfilm, Book	**Basis for assessment**: Billing model *Excludes ingest costs, excludes access costs* **Unit of measurement**: Billable square feet • $0.08 per 332-page (microfilm) volume per year in the standard vault • $0.19 per 332-page (microfilm) volume per year in the film vault • $0.31 per 332-page (book) volume in the standard vault
OCLC, Inc. **Chapman** *Format:* ASCII text, 600-dpi 1-bit page images	**Basis for assessment**: Billing model *Excludes ingest costs, includes access costs* **Unit of measurement**: Total GB of data deposited • $0.01-0.06 per 332-page ASCII text • $0.47/$1.01/$1.89 per 332-page 600-dpi 1-bit page image (variable rate, based upon total amount of data deposited per account)

<table>
<tr><th colspan="2">2005</th></tr>
<tr><td>Digital Preservation Testbed, Nationaal
Testbed Digitale Bewaring Archief of the Netherlands

Format:
Email, Text, Spreadsheet, Database</td><td>Basis for assessment: Literature review, testbed experience, and external project cost information
Includes an estimate of 20% overhead
Unit of measurement: Annual costs (total), Email batch costs

• Creation of a batch of 1000 records (assuming 50kb per email, 100 kb per text document, 250 kb per spreadsheet, and 2 Mb per database): 333 euros
• "Repair" of a batch of 1000 records (assuming 50kb per email, 100 kb per text document, 250 kb per spreadsheet, and 2 Mb per database): 10,000 euros
• Acquisition and input of metadata for "normal" email: 1.41 euros
• Acquisition and input of metadata for XML email: 0.06 euros</td></tr>
<tr><th colspan="2">2006</th></tr>
<tr><td>Riksarkivet
National Archives of Sweden
Palm

Format:
1-bit 600-dpi files in A4 format, 8-bit grey-scale at 297 dpi, Audiovisual</td><td>Basis for assessment: Audiovisual digitization costs
Includes storage system, staff operations, staff data input, service/support, and premises - averaged over 5 years
Intertemporal adjustment: 3% interest on hardware included in intertemporal calculations
Unit of measurement: Cost per year per 1 Gb stored; Total costs per year

• 1 Hierarchical Storage Management System (i.e., HSM) (2003 price + 3% interest per year): 449,694 euros over five years
Storage medium for additional 40 Tb/year: 43648 euros over five years
• Staff operations costs: 132240 euros over five years (0.6 FTE)
• Staff ongoing data input: 88160 euros over five years (0.4 FTE)
• Total annual input cost: 131808 euros over five years (staff & storage medium included)
• Facilities ("Premises") (100 square meters): 66228 euros over five years
• Service/support: 138300 euros over five years
• Digitization of paper materials (1-bit 600 dpi files in A4 format): 0.10 euro per file, with 5 million images scanned annually
• Scanning of large-format drawings and maps (8-bit grey-scale at 297 dpi, in manually fed scanners): 0.61 euro per file, with 1,321,000 image files created annually
• Production costs for 1 Gb 1-bit digitized information: approximately 17 euros per Gb
• Production costs for 1 Gb 8-bit digitized information: approximately 30 euros per Gb
• Production costs for Audiovisual information: approximately 11 euros per Gb</td></tr>
</table>

<table>
<tr><td>LIFE
Ayris et al.</td><td>Includes full life cycle costs
Intertemporal adjustment: 7% inflation factor for materials costs, 3.5% cost of living increase for staff costs
First year of e-monograph's life: lifecycle cost = 19 pounds (English)
Tenth year of hand-held e-monograph's life, lifecycle cost = 48 pounds (English) predicted
First year of hand-held serial's life, lifecycle cost = 19 nth year of hand-held serials life, lifecycle cost = 14 pounds (Engligh) pounds (English) predicted
First year of non hand-held e-monograph's life, lifecycle cost = 15 pounds (English)
Tenth year of a non hand-held e-monograph's life, lifecycle cost = 30 pounds (English) predicted
First year of non hand-held e-serial's life, lifecycle cost = 22 pounds (English)
Tenth year of a non hand-held e-serials life, lifecycle cost = 18 pounds (English) predicted
First year cost for new website = 21 pounds (English)
Tenth year cost for new website = 6,800 pounds (English) predicted
First year of e-journal's life, lifecycle cost = 206 pounds (English)
Tenth year of an e-journal's life, lifecycle cost = 3,000 pounds (English) predicted</td></tr>
<tr><td colspan="2">2007</td></tr>
<tr><td>Academy of Motion Picture Arts and Sciences
AMPAS Science and Technology Council

Format:
"all film" production; film-captured, digitally finished production at 4K; digitally captured, digitally finished production using HDCAM SR at 1920 x1080; digitally captured, digitally finished production using uncompressed system at 2K; digitally captured,</td><td>Basis for assessment: Data storage costs as reported by the San Diego Supercomputer Center; Annual (total) storage costs
Excludes initial inspection and access costs for both film storage and digital storage
Intertemporal adjustment: Amortized cost of YCM separation master manufacture ($800/year)
Unit of measurement: Total Annual Storage Costs
"All film" production generating no digital assets, annual storage costs for archival master: $1059
A film-captured, digital finished production at 4K, annual storage costs for archival master: $12,514
Digitally captured, digitally finished production using HDCAM SR videotape as the capture medium at 1920 x 1080, annual storage costs for archival master: $1,830
Digital captured, digitally finished production using an uncompressed digital data capture system at 2K, annual storage costs for archival master: $1,955
Digitally captured, digital finished production using an uncompressed digital data capture system at 4K, annual storage costs for archival master: $12,514</td></tr>
</table>

digitally finished production using uncompressed system at 4K	
Cambridge University *in Beagrie, Chruszcz, and Lavoie*	**Basis for assessment**: Literature review, survey questionnaire-based interviews, three in-depth case studies; Archaeology Data Service Charging Policy **Unit of measurement**: Staff time, Total cost in pounds or pence • Project initiation costs: 2-5 days per project for 2 FTEs at Grade 8 • Creation costs: 1-5 days at Grade 8 and above • Metadata creation: 5% of Grade 6 post (recurrent) • Acquisition: 5 days for one-off standard terms plus additional recurring negotiations • Outreach support: 20% of ongoing effort for repository manager (Grade 8) and Support and liaison officer (Grade 6) • Ingest: 10% of ongoing effort (recurrent) • Metadata upload and integration: 20% of two Grade 8 posts for 3 months • Integration with other campus systems: 300,000 English pounds over 3 years • Preservation planning: 0.5 FTE at Grade 8 (recurring)\ • First Mover Innovation: up to 5% of two Grade 8 personnel (recurring) • Sun Fire X4500 x64 Server: about 69,200 English pounds • DELL/EMC CX3-20c FC4 SPE DAE4P-OS for CX3-20: about 107,100 English pounds • Initial staff costs for 2 FTE graduate students Grade 6: 16,656 English pounds • Production staff costs for 1 10% FTE Graduate student Grade 6: 3331 English pounds • Production staff costs for 1 25% FTE Computer Officer Grade 8: 12,339 English pounds
King's College London *in Beagrie, Chruszcz, and Lavoie*	**Basis for assessment**: Literature review, survey questionnaire-based interviews, three in-depth case studies; Archaeology Data Service Charging Policy **Unit of measurement**: Staff salaries • Archive manager: 45,000 English pounds • Half-time system administrator: 24,000 English pounds • Collections officer: 35,000 English pounds
Archaeology Data Service (ADS) *in Beagrie, Chruszcz, and Lavoie*	**Basis for assessment**: Literature review, survey questionnaire-based interviews, three in-depth case studies; Archaeology Data Service Charging Policy **Unit of measurement**: Staff days, Total cost in pounds or pence, Per megabyte (Mb) cost in pence • Text and image file deposits: 1-10 files = minimum of 1 day • Text and image file deposits: 11-100 files = minimum of 2 days • Text and image file deposits: 100+ files = minimum of 4 days • Mixed files including GIS, CAD, Geophysics, Databases, etc.: 1-10 files

	= minimum of 2 days • Mixed files including GIS, CAD, Geophysics, Databases, etc.: 11-100 files = minimum of 3 days • Mixed files including GIS, CAD, Geophysics, Databases, etc.: 100+ files = minimum of 6 days • Queriable database: about 1000-5000 English pounds • Fully-functional GIS interface: as much as 10,000 English pounds • Cost of a gigabyte of disc storage: as low as 7 English pence; in 5 years, as low as 1 pence; approaching zero cost • Refreshment costs: for five year retention period: about 13 pence per Mb • Refreshment costs: for ten year retention period: about 22 pence per Mb • Refreshment costs: for fifteen year retention period: about 227 pence per Mb • Refreshment costs: for twenty year retention period: about 28 pence per Mb • Ongoing refreshment costs: about 30 pence per Mb
University of Southampton *in Beagrie, Chruszcz, and Lavoie*	**Basis for assessment**: Literature review, survey questionnaire-based interviews, three in-depth case studies; Archaeology Data Service Charging Policy • Staffing: 4 RA's = 332,000 English pounds • Staffing: Department Service experimental officer = 90,000 English pounds • Staffing: Department Self Service RA = 83,000 English pounds • Staffing: 3 PhD Research students = 90,000 English pounds • Lab instrumentation capital cost 45,000 English pounds (@ 10%) • Maintenance: 2,000 English pounds • Repair (averaged over 10 years): 10,000 English pounds • Raw data storage: 1,200 English pounds • Consumables: 4,000 English pounds • Assuming 2,000 datasets collected per annum, cost per crystal structure = 328.60 English pounds
2008	
SHERPA-DP IR *Ayris et al. (LIFE^2)*	**Basis for assessment**: Case study / workflow analysis *Includes full life cycle costs; Excludes interest rate, depreciation* Unit of measurement: Costs measured at the unit for which metadata is created (e.g., per object cost for analogue, per page cost for digital) • Year 1: 18.40 English pounds per year total cost • Year 5: 9.70 English pounds per year total cost • Year 10: 8.10 English pounds per year total cost
SHERPA-LEAP IR *Ayris et al. (LIFE^2)* ~Goldsmiths ~Royal Holloway ~UCL	**Basis for assessment**: Case study / workflow analysis *Includes full life cycle costs; Excludes interest rate, depreciation* **Unit of measurement**: Costs measured at the unit for which metadata is created (e.g., per object cost for analogue, per page cost for digital) Goldsmiths: • Year 1: 31.50 English pounds per year total cost

	• Year 5: 32.00 English pounds per year total cost • Year 10: 32.20 English pounds per year total cost Royal Holloway: • Year 1: 23.10 English pounds per year total cost • Year 5: 23.60 English pounds per year total cost • Year 10: 23.90 English pounds per year total cost UCL: • Year 1: 15.00 English pounds per year total cost • Year 5: 16.50 English pounds per year total cost • Year 10: 16.70 English pounds per year total cost
British Library Newspapers Digitization Project *Ayris et al. (LIFE[2])*	**Basis for assessment**: Case study / workflow analysis *Includes full life cycle costs; Excludes interest rate, depreciation* **Unit of measurement**: Costs measured at the unit for which metadata is created (e.g., per object cost for analogue, per page cost for digital) • Digital: 1,045,587 English pounds total project cost • Analogue: 1,820,702 English pounds total project cost

The model developed by Beagrie, Chruszcz, and Lavoie (2008), on the other hand, is oriented toward enabling institutions of higher education to develop a digital preservation cost model. This model builds upon the LIFE and other cost models, and generalizes the findings from a variety of initiatives to come up with a comprehensive, concrete model that accounts for the range of functional phases within the preservation environment, enables context-specific adjustments to occur (such as accounting for inflation or depreciation of capital assets), and maps to the OAIS and TRAC[1] models. Both the LIFE (5.6) and the Beagrie-Chruszcz-Lavoie (5.7) cost models are addressed in detail, highlighting their cost categories and elements together with other available cost information, including important studies conducted by universities, archives and libraries in the Netherlands (5.1, 5.3), U.S. (5.2), and Sweden (5.4) and by the motion picture industry (5.5).

5.0 Model Development – Digital Object Lifecycle and Preservation Needs

This section describes seven major studies that reflect the increasing sophistication of the cost models together with their major findings. The conclusions are summarized in Table 1.

5.1 Roquade Project

In 2001, researchers from the Library of Utrecht University, the Library of Delft University of Technology and the Netherlands Institute for Scientific Information Services reported on the Roquade project, an initiative involving the "the setting-up of an infrastructure for organising, coordinating, supporting and facilitating the digital publishing process, including an electronic

[1] In this context the acronym "TRAC" refers to the "Transparent Approach to Costing," a standard approach to costing in Higher Education in the UK (Joint Costing & Pricing Group, 2005), and *not* to the Trusted Repositories Audit and Certification Checklist, as one steeped in the archival literature might at first assume.

archiving facility" for scholarly scientific communication (Dekker, Dürr, Slabbertje, & van der Mee, 2001, p. 1) The authors acknowledged the deficiency of cost models at the time of this article's publication but nonetheless attempted to estimate various costs of preservation for items such as metadata generation, administrative and quality control efforts per item, and technical infrastructure costs for a flow of about 5,000 items per year for three broad preservation cost categories – metadata assignment, administrative and quality control activities, and technical infrastructure. Detailed information regarding the assumptions they used to come up with specific costs is not available. In addition, although offering preservation-related services, the project itself was an attempt to provide a framework for scholarly communication based on models suggested by Harnad (1999). Thus, their endeavor encompassed far more than "mere" preservation services. They recognized that costing digital preservation requires considerations beyond "pure" technology decisions and offered one set of rubrics for parsing the component costs appropriately.

5.2 Harvard Depository and OCLC, Inc. Digital Archive

In a now-classic piece on costs, Stephen Chapman (2003) cited the groundbreaking work of Kevin Ashley in identifying the difficulties of coming up with a comprehensive cost model for digital preservation services. Ashley (2000) had pointed out in "Digital Archive Costs: Facts and Fallacies" that the primary factor influencing the magnitude of preservation costs is not the quantity of objects preserved. Rather, it is the range of services offered by the archival provider. Chapman attempted to examine one particular aspect of this service provision: archival storage costs across microfilm, books and digital data. He offered a comparison of the services provided by Harvard University's Depository and OCLC, Inc.'s Digital Archive, highlighting the type and amounts of storage provided by each, and tried to provide relative costs of similar storage media and sizes across these two archival entities. The Harvard Depository offers storage for microfilm and books, and Chapman computed comparable digital costs of archival activities for OCLC, Inc., assuming several possible text formatting schemes.

Chapman assumed that prices equal costs, offering the observation that most content owners will be consumers of repository services rather than developers and managers of them. (p. 2) In his analysis, he therefore utilized the pricing structures of his two exemplars, recognizing that for the content owners, the price they pay is their cost. While a reasonable assumption, it is important to recognize that the price of services can easily diverge from the cost structure of the archive service provider, even a service provider that actively attempts to provide services at cost. This is because a single organization has no benchmark for assessing whether its cost structure is efficient with respect to the overall set of organizations within which it resides. Thus, any individual organization's reports on cost are potentially skewed. Nonetheless, Chapman's article provided a crucial introduction to cost assessment comparisons and helped to spur further discussion of this topic within the digital preservation environment, especially within that body of literature that compares hard copy archival costs to digital archiving costs and for a number of the works that discuss digital imaging and scanning (for example, Puglia, 1999). This latter literature relates to a much wider set of digital library development issues (as opposed to digital preservation issues per se) and is not discussed here.

5.3 Sweden’s Riksarkivet/National Archives

Early decisions made at Sweden’s Riksarkivet regarding digital preservation were inspired by the Chapman and Puglia articles and informed by a recognition that although storage capacity had been increasing rapidly, “the real cost of storage is management.” (Palm, 2006, p. 5) Palm focuses primarily on the digitization costs for audiovisual materials and finds that although storage media are continually becoming cheaper, the quantity of material to be stored has increased. Moreover, the capacity required for processing ever larger files has also increased. In addition, the notion of storage requires both the primary storage system and the surrounding hardware and software required to manage the stored material and to provide access to it or to usable copies of the stored digital masters. Indeed, he finds that the latter costs predominate, even in the face of cheaper storage media, with the primary costs being staff costs. In other words, although storage and support costs remain relatively static with scale, the overwhelming influence of the resulting staff (and facility) costs may lead to diseconomies of scale over time. This poses interesting trade offs with respect to decisions to convert analog materials, at least in terms of preservation if not access. Such trade offs are out of scope for this literature review, however. Nonetheless, nuanced characterizations of storage systems, the implications of scale, and the differentiation between preservation and access are relevant and important distinctions.

The information from Palm is shown in Table 1; it represents production costs for 1 gigabyte (Gb) of digitized information, and compares audiovisual to 1-bit and 8-bit image files. For a large audiovisual preservation project, he assesses that the lion’s share of costs will go to the conversion of the materials. Palm also describes the changing storage costs across time using a five to ten year timeframe, noting the annual changes in cost per newly added Gb of information and stressing the importance of planning for the changing cost structure of preservation activities across time.

5.4 Digital Preservation Testbed, Nationaal Archief of the Netherlands

The Digital Preservation Testbed, which operates as the digital preservation system for the Nationaal Archief of the Netherlands, has documented cost factors and calculations derived from its own work and that of others in order to provide a guide for larger archives, providing a highly detailed taxonomy of operational functions and cost factors for five high level categories of digital preservation costs: the cost of the archival system and “functionality for the long-term preservation of the digital records;” personnel costs; development of software and technical methods of preservation; the “cost of the actual storage of digital records;” and “other factors.” The drilldown into subcategories is quite comprehensive. For example, one component of the cost of the archival system is “physical space,” which can be further broken down into server room, with air-conditioning; sufficient office space; conference room; toilets and kitchen; and security (Testbed Digitale Bewaring, 2005, p. 5). This taxonomy obviously includes far more than the preservation costs alone, potentially offering a clear understanding of both the overall cost structure and the “pure” preservation component of all activities. An Excel spreadsheet cost modeling tool to allow a preservation facility to calculate its own preservation costs is provided: http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=185&categorie=6.

The Testbed has made a series of assumptions regarding the necessary full time staff (in FTEs), hardware, software, and facilities requirements, and the costs of these resources. These assumptions are derived largely from published literature and other information gained from a variety of institutions or projects such as Delft University of Technology (Utrecht University), the Annual Report of the National Archives of Canada, the UK National Archives, the National Archives of Australia, the U.S. National Archives and Records Administration (NARA), and the Amsterdam Municipal Archives. The results of their assumptions provide the cost assessments given here in Table 1 and are meant to represent the costs of a facility that preserves and manages 1 terabyte (Tb) of information, with the capacity to expand to at least 1 petabyte (Pb) of information stored. Taking a "records continuum" approach to digital preservation, they consider the entire lifecycle of management to be a component of digital preservation. (pp. 4, 15) They also estimate the cost for the "repair" of objects that were not appropriately captured or preserved at the "right" time. Their primary findings are consistent with the widespread belief among digital preservationists that archiving begins at creation. Thus, creating records properly at the beginning of the lifecycle results in considerably lower preservation costs over the long-run than will occur if object repair is needed.

5.5 Digital Motion Pictures

An important addition to the preservation literature is a study released in 2007 by the Science and Technology Council of the Academy of Motion Picture Arts and Sciences (A.M.P.A.S.), which reviews strategic issues in archiving digital motion picture materials (Science and Technology Council). The document discusses the changing nature of archiving within the digital motion picture industry and provides an assessment of digital storage economics in general, and costs in particular, for that industry. Noting that the annual cost of preserving digital materials per title is approximately eleven times greater than the cost of preserving film material per title (pp. 1-2) and that the born-digital masters nonetheless need to be kept for at least 50-100 years just as the original film masters, the authors stress the urgency of ensuring that the challenges of digital preservation in this industry are faced and resolved.

The Motion Picture Archiving document explicitly notes the tendency for many vendors (and researchers) to provide incomplete cost analyses by "simplifying the cost components to just the storage media and devices…, ignoring other costs that the user will inevitably face." (p. 40) It offers the results of two economic case studies, one drawn from a motion picture originally captured on film and then digitally finished to distribution, the other drawn from a motion picture that was born digital and also finished digitally through distribution. The authors limit their assessment to picture and sound elements. They assess the annual storage costs, "exclusive of ingest, inspection and access costs" (p. 43) for five scenarios ranging from a production that generates no digital assets to digitally finished productions either captured on film or digitally and finished using various media and compression schemes. Table 1 shows the total costs of each of the five scenarios.

5.6 LIFE Project and LIFE^2

Using data from UCL's (University College London) e-journals, the British Library Web Archiving activities, and the British Library VDEP digital collections, the LIFE project developed a lifecycle framework for testing their methodology around managing and costing the preservation of digital material. The authors published their Phase 1 project report in 2006 and a second, Phase 2 report, in 2008. Their goal was to develop a framework that would allow them to determine the long term cost of digital preservation and the long term costs of library partnerships within the higher education community, to provide comparisons of paper versus digital preservation for any given publication, and to assess the relative risks of digital versus paper archiving. They also wanted to be able to use their framework to determine who would engage in the preservation and when it would be appropriate to switch from paper to digital formats for long-term preservation. (McLeod et al., p. 6) Because the case studies provided did not provide sufficient evidence for the production of digital preservation costs, they developed a separate model for LIFE preservation.

The original LIFE model examines all the elements of maintaining an object over its lifecycle from acquisition through to preservation (that is, acquisition, ingest, metadata, access, storage, and preservation.) It details what is meant by each of these stages and what types of activities fall within them. The authors of the LIFE model recognize that the upfront (i.e., "one-time") costs of a project are often distinct in structure from the recurring maintenance aspects of the same project, and have built a model that allows one to track these costs separately.

For each type of activity, a case study example is provided for illumination. Where available, research review evidence is also provided. Although none of the case studies provides cost information for the preservation activities, the authors do attempt to make realistic estimates that allow them to draw certain conclusions about management. For example, web archiving clearly needs additional preservation tools to reduce costs to institutions and intra-institutional cost sharing may be a wise service strategy. (McLeod et al., p. 62) They note that "real life cost data" is needed to fine tune and test the preservation component of the model. The elements within the preservation phase are technology watch, preservation tool cost, preservation metadata, preservation action, quality assurance, and "other." (p. 15) In addition to their model, they provide a comprehensive literature review of cost modeling both within and outside digital preservation projects (Watson, 2005).

Phase 2 of the LIFE project (LIFE^2) refined the original model and added three new case studies. They updated their overall cost and digital preservation cost models and validated their approach by offering it for analysis to an independent economic review. Within a "British Library newspapers" case study, they compared the lifecycle costs of both paper and digital collections. They also provide a series of spreadsheets and offer advice on how an institution can apply the LIFE model to their own circumstances. Overall, they fine tuned their original model and offered additional empirical support for their costing approach.

5.7 *Beagrie-Chruszcz-Lavoie (BCL) Cost Model*

The most recent new digital preservation cost model is that developed by Beagrie, Chruszcz, and Lavoie (2008), henceforth referred to as the BCL model.[2] Like the LIFE Project model, the BCL model focuses upon long-term preservation costs within higher education institutions. Also like the LIFE model, the BCL model recognizes that one-time and recurring cost structures may differ, and enables one to track these two types of costs separately. In addition, however, the BCL model includes what they call "First-Mover Innovation" costs (p. 6), or the (sometimes substantial) costs involved in being the originator or developer of a new tool, technology, or method that moves the state of practice forward.

This cost model reflects the lifecycle of digital objects by breaking the costs into three primary stages: pre-archive, archive, and support services. Each of these phases encompasses a number of cost creating activities, as shown below. (pp. 36-46)

Phase	Function	Activity
Pre-Archive Phase	Initiation	Project Design
		Data management plan
		Funding application
		Project implementation
	Creation	Negotiate IPR/Licensing/ethics
		Generate research data
		Generate descriptive metadata
		Generate user documentation
		Generate customised software
		Data management
		Create submission package for archive
Archive Phase	Acquisition	Selection
		Negotiate submission agreement
		Outreach and depositor support
	Disposal	Transfer to another archive
		Destroy
	Ingest	Receive submission
		Quality assurance
		Generate Information Package for archive
		Generate administrative metadata
		Generate/upgrade descriptive metadata and user documentation
		Co-ordinate updates

[2] Although LIFE^2 has in fact been published since the BCL model, it is interpreted here as an extension of the original LIFE model, rather than an entirely new model.

Phase	Function	Activity
		Reference linking
	Archive Storage	Receive data from ingest
		Manage storage hierarchy
		Replace media
		Disaster recovery
		Error checking
		Provide copies to access
	Preservation planning	Monitor designated user community
		Monitor technology
		Develop preservation strategies and standards
		Develop packaging designs and migration plans
		Develop and monitor SLAs for outsourced preservation
		Preservation action
		Generate preservation metadata
	First Mover Innovation	Develop community data standards and best practice
		Share development of preservation systems and tools
		Engage with vendors
	Data management	Administer database
		Perform queries
		Generate report
		Receive database updates
	Access	Search and ordering
		Generate information package for dissemination to user
		Deliver response
		User support
		New product generation
Support Services	Administration	General management
		Customer accounts
		Administrative support
	Common services	Operating system services
		Network services
		Network security services
		Software licences and hardware maintenance
		Physical security
		Utilities
		Supplies inventory and logistics
	Estates	Estates management activities

A key finding of the BCL case studies is that the costs of preservation do increase but do so at a decreasing rate as the retention period increases. That is, the *pace* of the increase tends to fall over time after an initial period of increase. For example, for the ADS repository, the per megabyte (Mb) refreshment costs of a collection with a five year retention period are significantly higher than the refreshment costs of a collection with a fifteen year retention period. The longer the retention period, the slower the rate of increase in refreshment costs, suggesting that "preservation costs become negligible after 20 years" and that "a charge of 30p per megabyte would cover ongoing preservation beyond 20 years." (p. 92)

The BCL model is both sophisticated and comprehensive. A key benefit it provides is the mapping of the model to the OAIS requirements. Nonetheless, its comprehensiveness relies on the use of an activity-based costing approach. This approach is consistent with the British TRAC system, but may be more difficult to map to other settings (notably in the U.S.), where activity-based models are often not consistent with higher education accounting practices and employee culture (Courant & Knepp, 2002; Ellis-Newman, 2003). Nonetheless, this type of costing has been accepted for several decades as a superior mechanism for capturing all relevant costs of providing services, ever since it was first lauded by Robin Cooper and Robert Kaplan in the *Harvard Business Review* (Cooper & Kaplan, 1988). Because it is especially helpful in allocating indirect costs to services based on the factors that most influence them, this accounting model can provide key insights into cost behavior in a service-oriented organization, by forcing the organization to determine what activities drive the costs of service provision and what activities do *not* add value to the provision of those services (Ellis-Newman, 2003).

6.0 Discussion: Observations and Gaps

The project reports and publications presented here and summarized in Table 1 illustrate the conceptual and practical changes that have taken place in the discussion of the costs of digital preservation over the last decade. However, as Table 1 highlights, it is difficult to take full advantage of individual project findings because these cost models remain largely incommensurable with each other. They use different models; they define costs differently and assign different units of measurement; different formats are captured; and decisions regarding which costs and cost adjustments to include and exclude vary from project to project. Earliest attempts at ascertaining costs generally focused upon a very small range of preservation costs, such as storage costs alone or the cost of metadata and technical infrastructure. With time, the discussion has become more sophisticated, and explicitly recognizes that costs are embedded in a larger framework that includes stakeholder interests, organizational structure, and cultural milieu. Interests have widened to include the costs of preserving an object across its entire lifecycle, as exemplified in the very comprehensive cost models of the LIFE project and the Beagrie, Chruszcz, and Lavoie project. These models not only attempt to determine the individual costs for the various lifecycle activities; they also include intertemporal considerations (for example, inflation/deflation, interest rates, etc.) and potential cost dependencies, where "changes in one affect the other." (Beagrie et al., 2008, p. 21)

This increasing sophistication has led to a series of discoveries about the costs of digital preservation. For example, by focusing upon the lifecycle of digital objects, various projects have determined that costs may be unevenly distributed over the lifecycle of digital objects. Beagrie and colleagues noted that the costs of ingest carry especially heavy weight, leading to a recognition that if done right, the costs of preservation will become negligible in the long run. Of course, if the preservation activities were not done right, then the costs of repair are exceedingly expensive, as pointed out by the National Archives of the Netherlands. The Riksarkivet of Sweden likewise points to the weight of people costs in the preservation of audio-visual materials. Even in the face of declining storage media costs, more time in storage management is required, and this is due to the rapid increase in manually intensive storage management routines.

Palm's (2006) paper, together with studies on the growth of the information universe (Gantz, 2008) and the storage industry itself (Peterson, Zasman, Mojica, & Porter, 2007) suggest that storage is a more textured component than early studies may have indicated. Not only is equipment swapped out on a 3- or 5-year cycle, a step which Beagrie et. al. (2008) and others build into their models, but storage itself becomes more complex with scale. An organization that manages 1 Pb today will need to migrate 2.25 Pb in 5 years. This conforms to industry best practice of data migration in a 3-5 year cycle. "How do organizations expect to do that and keep up with the growth, cost, and complexity?" the Storage Networking Industry Association *100 Year Archive* Task Force asked rhetorically. "The answer is they can not. They will not" (Peterson et al., 2007, p. 2). In addition, power costs are continuing to rise and to rise more quickly than the costs of new servers. Power consumption, estimated at 1kW per server rack in 2000 is estimated at 20kW per rack in 2007 and project to rise to 20kW per rack in the next few years (Gantz, 2008, p. 4).[3] Greater capacity will not necessarily solve the problem for a custodial institution since the creation of new information is accelerating faster than our capacity to store it (Gantz, 2008, p. 2) and the engineering of cooling large, complex servers is complicated; hence, power requirements to operate and maintain such systems are considerable and the costs not only of upgrading but also of operating these systems require serious deliberation.

There is a growing body of anecdotal evidence that the scale and complexity of large-scale data sets could push our current IT systems beyond a tipping point. The growing sustainability movement adds further incentive to reduce the overall environmental impact of IT infrastructure. Many IT organizations such as Educause have produced briefing papers or position statements in this context (http://connect.educause.edu/Library/EDUCAUSE+Review/OnBeingGreen/46617). The impact of power and cooling requirements has become an overarching and critical consideration for IT service providers, resulting in a "green IT" track at the Interop conference (http://www.interop.com/lasvegas/education/green-it.php). Both for-profit and not-for-profit institutions are adopting strategies such as data center consolidation, use of more efficient equipment, or virtualization. However, at the Spring 2008 Preservation and Archiving Special Interest (PASIG) meeting, there was discussion about the need to reconsider storage architecture

[3] At a recent (May 2008) meeting of Preservation and Archiving Special Interest Group (PASIG), one of Sun's chief engineers offered a slide with back of the envelope calculations that providing power and cooling for a petabyte of data could cost $1 million a year based on California utility rates.

in a fundamental manner. At the least, it is essential to consider the costs of power and cooling requirements, especially given that the cost of energy will continue to rise in the future.

Review of the cost literature to-date points out what may now seem obvious, but was not always recognized: Format matters; scale matters. Image files are larger than text or numeric files so the costs associated with storing, managing, and accessing image files are greater. As the authors of *The Diverse and Exploding Digital Universe* point out (Gantz, 2008), the financial industry employs some of the most sophisticated computing systems using enormous quantities of data, generates 6 percent of the world's gross output, and buys 20 percent of its computers. Yet the multibillion dollar sector only accounts for *6 percent* of the total information universe, and its share is projected to fall to 3 percent by 2011 (p. 6). Numerical data, largely stored in complex databases, is extremely efficient in its ability to represent information – a lot of meaning can be captured in very concise notation. As the authors put it, "There is simply not enough imaging going on" (p. 6).

The A.M.P.A.S. Science and Technology Council provide a particularly clear example of the importance of both scale and format and calls attention to the potential variable nature of costs across different sectors. It also highlights the preservation environment in a number of different industries. Digital movie productions differ in preservation cost depending upon whether they originated on film or were born-digital. The council's report provides particularly compelling evidence for the immensity of the cost issue to this industry in light of the unheard of growth in materials to be preserved in the future. Lavoie's broader discussion of the impact of different preservation strategies (2006) supports these findings and Chapman's early study (2003) also supports the idea that management decisions regarding preservation strategy strongly impact the costs. Finally, and as pointed out by others, the LIFE project and the research by Beagrie, Chruszcz, and Lavoie (2008) stem from the higher education sector in the U.K. and its accounting framework. While this framework greatly assists analysis of indirect costs, it is not obvious that the cultural assumptions that underpin the U.K. higher education sector will map easily to different nations with different academic traditions or to other sectors. Thus, as was pointed out by early researchers, institutional culture and their impacts on the mission of the entity in which preservation activities are conducted matter.

References

Ashley, K. (2000). *Digital Archive Costs: Facts and Fallacies*. Paper presented at the DLM Forum '99. from http://ec.europa.eu/archives/ISPO/dlm/fulltext/full_ashl_en.htm.

Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H., & Wheatley, P. (2008). *The LIFE^2 Final Project Report*: JISC.

Baumol, W. J., & Marcus, M. (1973). *Economics of Academic Libraries*. Washington, D.C.: American Council on Education.

Beagrie, N., Chruszcz, J., & Lavoie, B. (2008). *Keeping Research Data Safe: A Cost Model and Guidance for UK Universities*. London: JISC.

Beagrie, N., & Greenstein, D. (1998). *A Strategic Framework for Creating and Preserving Digital Resources: A JISC/NPO Study Within the Electronic Libraries (elib) Programme on the Preservation of Electronic Materials*. London: Library Information Technology Centre.

Bradley, K. (2005). *APSR Sustainability Issues Discussion Paper*: Australian Partnership for Sustainable Repositories.

Bradley, K. (2007). Defining Digital Sustainability. *Library Trends, 56*(1), 148-163.

Chapman, S. (2003). Counting the Costs of Digital Preservation: Is Repository Storage Affordable? *Journal of Digital Information, 4*(2).

Connaway, L. S., & Lawrence, S. R. (2003). Comparing Library Resource Allocations for the Paper and the Digital Library. *D-Lib, 9*(12).

Cooper, R., & Kaplan, R. S. (1988). Measure Costs Right: Make the Right Decisions. *Harvard Business Review*, 96-103.

Courant, P. N., & Knepp, M. (2002). Activity Based Budgeting at the University of Michigan: A Case Study. In D. M. Priest, W. Becker, D. Hossler & E. St. John (Eds.), *Incentive-Based Budgeting Systems in Public Universities* (pp. 137-160). Camberley Surrey, UK: Edward Elgar, Ltd.

Crespo, A., & Garcia-Molina, H. (2001). *Cost-Driven Design for Archival Repositories*. Paper presented at the First ACM/IEEE-CS Joint Conference on Digital Libraries. from http://www-db.stanford.edu/~crespo/publications/cost.pdf.

Cummings, A. M., Witte, M. L., Bowen, W. G., Lazarus, L. O., & Ekman, R. H. (1992). *University Libraries and Scholarly Communication: A Study Prepared for The Andrew W. Mellon Foundation*: Association of Research Libraries.

Curall, J., & McKinney, P. (2006). Investing in Value: A Perspective on Digital Preservation. *D-Lib Magazine, 12*(4).

Currall, J., & McKinney, P. (2006). Investing in Value: A Perspective on Digital Preservation. *D-Lib Magazine, 12*(4).

DeGennaro, R. (1987). *Libraries, Technology and the Information Marketplace: Selected Papers*. Boston, MA: G. K. hall and co.

Dekker, R., Dürr, E., Slabbertje, M., & van der Mee, K. (2001, April, 2002). *An Electronic Archive for Academic Communities.* Paper presented at the ICEIS 2001/NDDL Workshop.

Dollar, C. M. (1999). *Authentic Electronic Records: Strategies for Long-term Access*. Chicago, IL: Cohasset Associates.

Economic Consulting Services. (1989). A Study of Trends in Average Prices and Costs of Certain Serials Over Time. In *Report of the ARL Serials Prices Project*. Washington, D.C.: Association of Research Libraries.

Ellis-Newman, J. (2003). Activity-Based Costing in User Services of an Academic Library. *Library Trends, 51*(3), 333-350.

espida. (2007). *espida Handbook: Expressing Project Costs and Benefits in a Systematic Way for Investment in Information and IT*. Glasgow.

Galvin, T. (1996). Economic Considerations for Digital Libraries. In M. A. Butler & B. R. Kingma (Eds.), *The Economics of Information in the Networked Environment*. Washington, D. C.: Association of Research Libraries.

Gantz, J. F. (2008). *The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information Growth Through 2011*: International Data Corporation (IDC).

Getz, M. (1992). Electronic Publishing: An Economic View. *Serials Review, 18*, 25-31.

Granger, S. (ND). Digital Preservation & the CAMiLEON Project. Retrieved May 21, 2008, 2008

Harnad, S. (1999). Report on Santa Fe Initiative (Interoperable Eprint Archives). *AMERICAN-SCIENTIST-OPEN-ACCESS-FORUM archives* 2008, from http://listserver.sigmaxi.org/sc/wa.exe?A2=ind99&L=american-scientist-open-access-forum&D=1&O=D&F=l&S=&P=30213

Hendley, T. (1998). *Comparison of Methods and Costs of Digital Preservation*. London: British Library Research and Innovation Centre.

Holdsworth, D. (2001). Emulation, Preservation, and Abstraction. *RLG DigiNews, 5*(4).

Hunter, L. (2006). Investment in an Intangible Asset. In S. Ross & M. Day (Eds.), *DCC Digital Curation Manual*: HATII, University of Glasgow

University of Edinburgh

UKOLN, University of Bath

Council for the Central Laboratory of the Research Councils.

Joint Costing & Pricing Group. (2005, May 25, 2008). Transparent Approach to Costing (TRAC) Manual. from http://www.jcpsg.ac.uk/guidance/about.htm

Kenney, A. R., & Rieger, O. Y. (2000). *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. Mountain View, CA: Research Libraries Group.

King, D. W., Boyce, P. B., Montgomery, C. H., & Tenopir, C. (2003). Library Economic Metrics: Examples of the Comparison of Electronic and Print Journal Collections and Collection Services. *Library Trends, 51*(3).

Lavoie, B. F. (2003). *The Incentives to Preserve Digital Materials: Roles, Scenarios and Economic Decision Making*: OCLC Research.

Lavoie, B. F. (2006). The Costs of Digital Preservation. In M. Deegan & S. Tanner (Eds.), *Digital Preservation* (pp. 106-132). London, UK: Facet Publishing.

Lorie, R. A. (2001, June 24-28, 2001). *Long Term Preservation of Digital Information.* Paper presented at the First ACM/IEEE-CS Joint Conference on Digital Libraries, Roanoke, VA.

McLeod, R., Wheatley, P., & Ayris, P. (2006). *Lifecycle Information for E-literature: Full Report from the LIFE Project*. London, U.K.: University College London and the British Library.

Palm, J. (2006). *The Digital Black Hole*. Stockholm, Sweden: Riksarkivet/National Archives.

Peek, R. P., & Pomerantz, J. P. (1998). Electronic Scholarly Journal Publishing. *Annual Review of Information Science and Technology, 33*, 321-356.

Peterson, M., Zasman, G., Mojica, P., & Porter, J. (2007). *100 Year Archive Requirements Survey*: Storage Networking Industry Association (SNIA).

Puglia, S. (1999). The Costs of Digital Imaging Projects. *RLG DigiNews*, *3*(5).

Rothenberg, J. (1999). *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. Washington, D.C.: Council on Library and Information Resources.

Rothenberg, J., & Bikson, T. K. (1999). *Carrying Authentic, Understandable and Usable Digital Records Through Time*: RAND Europe.

Russell, K., & Weinberger, E. (2000). *Cost elements of digital Preservation*. Leeds: Cedars Project.

Sanett, S. (2002). Toward Developing a Framework of Cost Elements for Preserving Authentic Electronic Records into Perpetuity. *College & Research Libraries, 63*, 388-404.

Science and Technology Council. (2007). *The Digital Dilemma: Strategic Issues in Archiving and Accessing Digital Motion Picture Materials*: Academy of Motion Picture Arts and Sciences (A.M.P.A.S.).

Testbed Digitale Bewaring. (2005). *Costs of Digital Preservation*. The Hague, Netherlands: Nationaal Archief of the Netherlands.

Watson, J. (2005). *The LIFE Project Research Review: Mapping the Landscape, Riding a Life Cycle*. London, UK: University College London and the British Library.

Wheatley, P. (2001). Migration - a CAMiLEON Discussion Paper. *Ariadne, 29*.