

# Planning an effective digital preservation strategy for a research organization

Adèle van der Merwe  
CSIR Information Services  
[avdmerwe@csir.co.za](mailto:avdmerwe@csir.co.za)

Dr Martie van Deventer  
CSIR Information Services  
[mvandeve@csir.co.za](mailto:mvandeve@csir.co.za)

## 1. INTRODUCTION

Internationally organisations are struggling with a general disregard for the importance of taking proper care of digital artefacts. The purpose of this paper is to provide insight into the background for developing a digital preservation strategy for a research organization such as the Council for Scientific and Industrial Research (CSIR). The CSIR's shift to a predominantly digital environment, started in the early 1990s. The international focus on digital preservation has created awareness for us that our average creators of research artefacts are also ignorant of the threats facing digital records. Believing that a regular backup process is more than sufficient for the safekeeping of their information and data, several researchers have of late been confronted by the fact that especially their old research data have become difficult to access. The proposed strategy will highlight the shortcomings in the existing culture and will present a framework for a formal organizational strategy. Fortunately, much work has been done internationally and the CSIR can learn from these initiatives.

## 2. Ten laws of preservation

Preservation does not however exist in a vacuum and neither should it be a purpose on its own. The following ten 'laws' of preservation are based on Lavoie and Dempsey's (2004) thirteen ways of looking at preservation. They address the essential elements and serve as good guidance for the CSIR, and other organizations, that need to develop a sustainable digital preservation.

- Digital preservation is an ongoing activity and therefore procrastination can no longer be allowed. Furthermore, the preservation activities should co-exist harmoniously with other operating procedures and standards.

- Digital preservation is a set of agreed outcomes and all of the role-players must reach consensus in terms of the 'what, why, when and how' of preservation activities and methods.

- Digital preservation is an understood responsibility and the responsibility for preservation does not/should not rest with the library or the archive exclusively but is also a responsibility of the creators of the artefacts. Preservation starts with the creation and continues thereafter.

- Digital preservation is also a selection process. Although most organizations have been in the 'save now, preserve later' mode for many years, this is no longer a viable option. However, it is also unrealistic to try and preserve all digital material in existence within the organization. Therefore realistic selection criteria, based on the long term value of the objects, should be developed and maintained in order to keep track with new developments and research.

- Digital preservation must be seen a sustainable activity. To date, preservation activities have been characterised by high implementation costs. It is anticipated however, that with maturity, the costs of preservation will continue to decline and that the benefits will outweigh the costs.

- Digital preservation is a cooperative/joint effort. The cooperation between role-players will ensure cost effectiveness, system efficiency and applicable selection criteria. Approaching digital preservation blindly and on a piecemeal basis will be impractical as the 'bigger picture' will remain obscure and this could lead to wrong and impractical decisions. However, that is not to say that

everything must be done simultaneously or immediately. A holistic approach will assist with the prioritising process and the selection of a suitable preservation method.

- Digital preservation will not harm or threaten intellectual property rights. Rather than threaten IP rights, effective preservation will ensure the safekeeping thereof. Preservation does not imply free access. It rather strives to guarantee the long-term security of and access to digital content.

- Digital preservation is a disaggregated service. Various units/departments have responsibilities in terms of preservation. For example in the CSIR ICT is responsible for the management and maintenance of the infrastructure, hardware and technology. The archivist is responsible for the metadata and validation of the data's integrity and authenticity and the repository manager is responsible for ensuring effective accessing and retrieval.

- Digital preservation complements existing record management and library services. Currently, at the CSIR, the preservation activities form part of the library services.

- Digital preservation must be a well understood process. In order to achieve success, the preservation strategy should be communicated to the whole organization. Standards should be in place regarding the selection of digital objects to be preserved, the activities involved and the minimum acceptable standards.

Understanding and applying these 'laws' obviously requires an understanding of the context within which they need to be adopted. The next section will therefore provide some insight into the CSIR's work environment.

### **3. DEFINING THE CONTEXT**

The CSIR is one of the leading scientific and technology research, development and implementation organisations in Africa. As such, the CSIR undertakes directed and multidisciplinary research and technological innovation to improve the quality of life of South Africa's people (CSIR 2008a). Following a records management drive in 2008, the preservation of the CSIR's research related artefacts has been identified as an organizational imperative activity. Preservation, in this context includes, for example, the obvious activities that would ensure that the digital formats remain usable and accessible but it also includes the assignment of metadata and the safe storage of the actual digital objects.

The CSIR has a very well defined risk management strategy, policy and procedures. The procedures include and guide the way in which our Information and Communication Technology (ICT) department needs to do backups. While backups do address the fundamental safekeeping of digital records, preservation far exceeds these activities. As a first step it was therefore essential to understand and explain the basic differences between making backups and doing preservation. The backup activity in reality only copies files for use in case of data loss or hardware failure. By implication, backups are made of active files and perform an essential role in the daily activities of an organization's basic risk management strategy. Digital preservation activities refer to the processes and actions that will help to ensure the continued and indefinite access to information and records that exists in a digital format even when the digital files are taken out of their creation context.

Preservation is not a new phenomenon. It is rather just a neglected and under valued activity in the majority of organizations. The way and format in which digital information is stored are important in ensuring the artefact's ongoing and continued use. Rather than focusing on active files only (as is the case with creating backup files), preservation focuses on inactive files, as well as on those finalized or completed records that have been closed to prevent any future changes or manipulations.

For the backup activity the focus is on reproducibility. For preservation the focus is on long term access and use.

Preservation is a very important component of a more comprehensive digital curation lifecycle. The Digital Curation Centre (DCC) curation lifecycle model (Figure 1) clearly illustrates that preservation action relates to those activities where artefacts are identified and 'adopted' for preservation, where very specific activities are performed to preserve the artefacts, where the artefacts are then deliberately stored so that they could be accessed, used and reused.

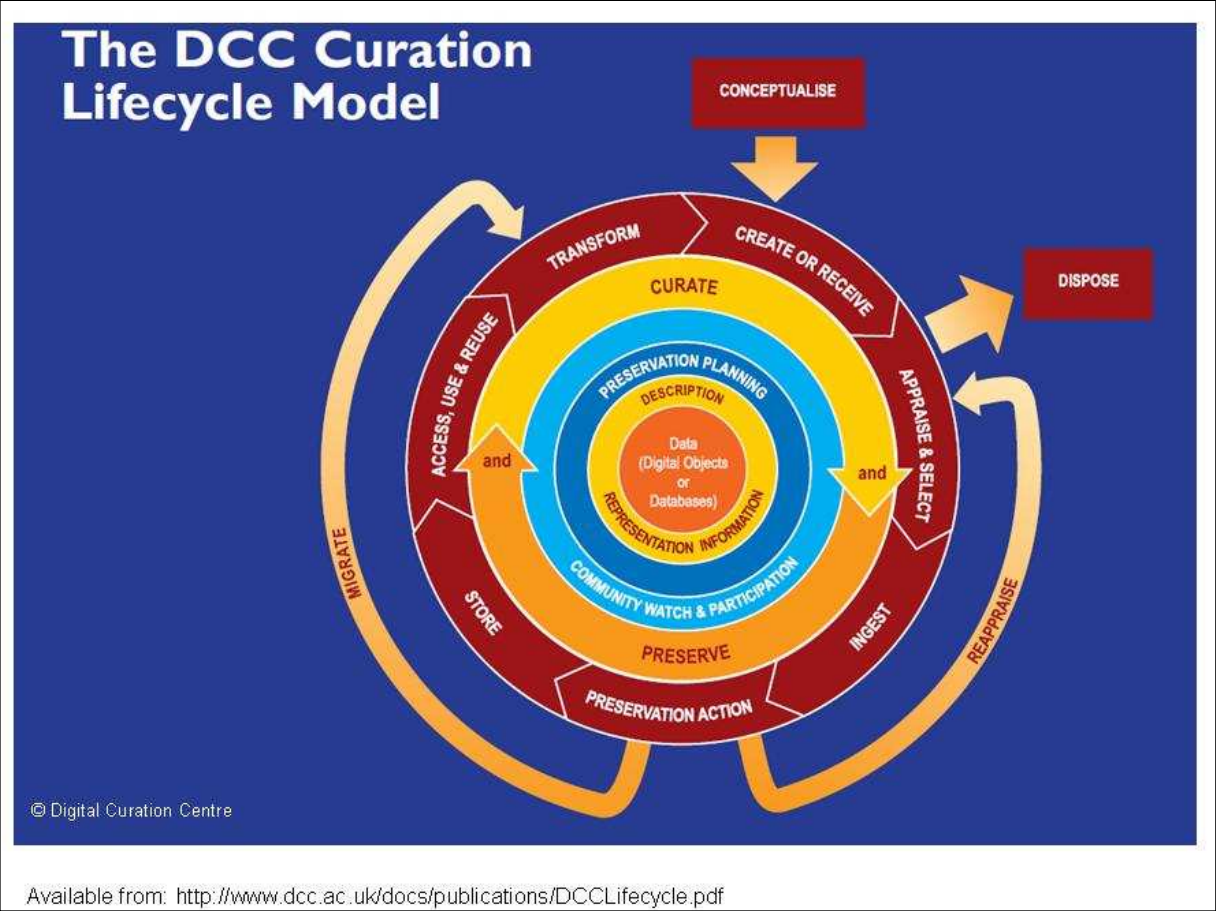
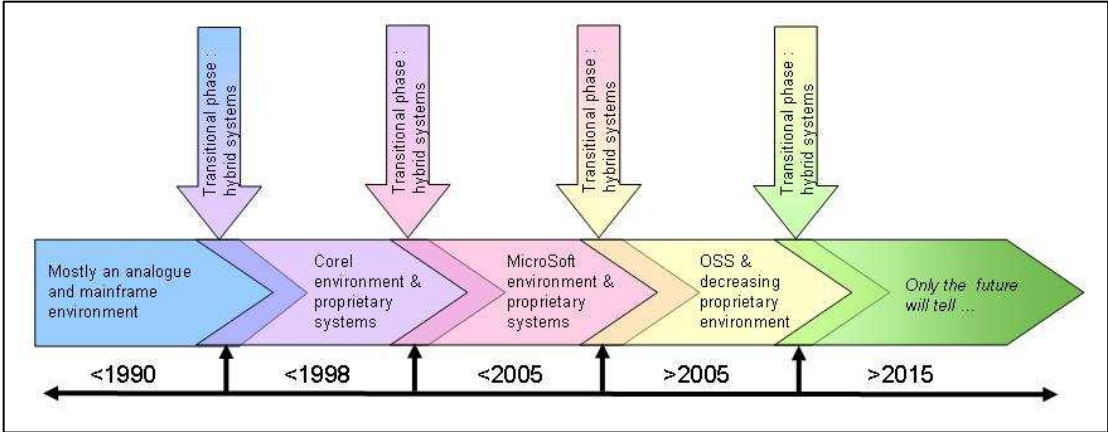


Figure 1: Digital curation lifecycle model (Digital Curation Centre)

The digital curation activity goes further to include also the transformation of artefacts, the creation of new artefacts as well as establish the appraisal and evaluation criteria for the selection of valuable digital assets. Following the curation lifecycle model as a strategy ensures that digital assets are collected and verified; that digital records can be searched and retrieved; that certificates regarding the trustworthiness and integrity of the digital format are available; and that the semantic and ontological continuation and comparability of the collection remain available.

Gaining a theoretical understanding of digital preservation made it clear to us that it no longer was a question of whether or not our organization can afford data preservation but rather a question of whether we could afford not to implement a strategy that would ensure the preservation of our digital assets. Figure 2 illustrates the changing environment within the CSIR since the onset of the predominantly digital era. The 'peace of mind'-phase, regarding the safety and integrity of information, that generally existed prior to the 1990s has largely disappeared. Since the 1990s there

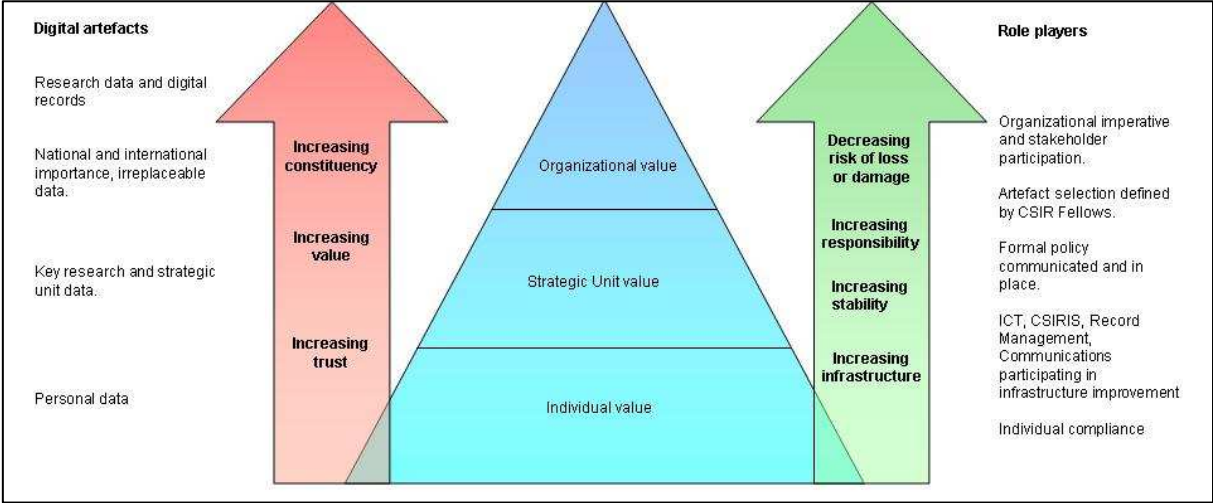
has been a rapid change in technology. In a relatively short period of time, the organization moved from a Corel environment to a Microsoft environment followed by a move towards the Open Source (OS) environment. Each of the previous eras lasted six to eight years and included a hybrid migratory period. The OS era is still in its infancy and it is not yet clear how long this era will last, but if the pattern is followed, the next change can be expected around 2015! Having no clear cut-off lines or time frames between transformation stages also hold challenges. Changes took place so gradually that individuals either didn't appreciate the impact of the switch to a new product line or they assumed that the old products will remain available indefinitely, thereby negating the need to be concerned about preservation. After all, the argument was that their data was backed-up and therefore 'safe and accessible'!



**Figure 2:** A dynamic and changing environment

The CSIR's records management policy, and the subsequent emphasis on digital content, and the need for digital preservation surfaced simultaneously. The records management strategy was constructed around the requirements of the National Archives and Records Service of South Africa Act (South Africa 1996). To ensure adherence to the Act our records need to be accessible. To ensure that our digital records remain accessible a digital preservation strategy has become a pressing need.

We are however, aware that preservation is not a cheap and dirty activity. One therefore needs to be very sure that only those records that are of real value are preserved. We found that the model depicted in Figure 3 provides good guidance when determining value.



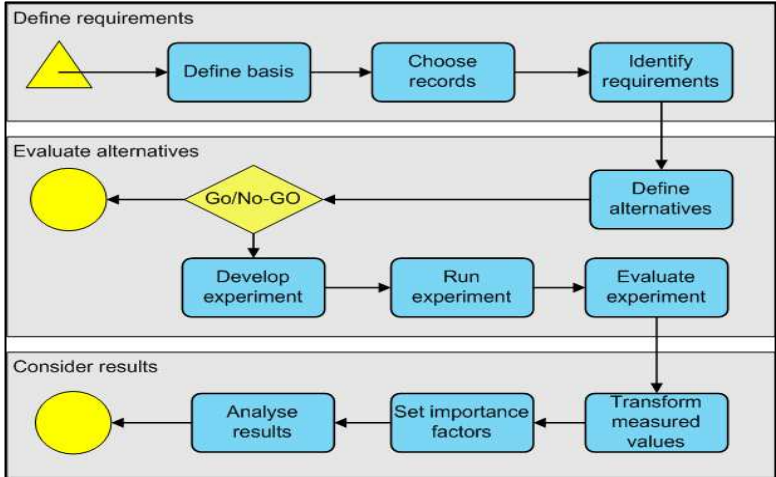
**Figure 3:** The value pyramid (based on the data pyramid of Berman (2008))

Berman’s model (2008) presents a data pyramid to be used for the selection and preservation of artefacts in the context of Institutional Repositories. Figure 3 is based on the data pyramid but was adapted to reflect the value of research artefacts and the role players in terms of the CSIR’s preservation strategy.

Trust is of the utmost importance for preservation to be effective. Trust must exist from an individual level up to an organizational level for compliance to take place. It is also important to understand that the concept of ‘value’ will differ between individuals, community and society.

Similarly preservation cannot exist in a vacuum and within the CSIR the efficient collaboration of a multi-disciplinary team consisting of the CSIR’s ICT, Records Management, Information Services and Communications functions, together with our research staff will be required before our goal, to do effective preservation, could be reached.

When it comes to the selection of the research artefacts particularly the collaboration of the CSIR’s Research Fellows will play an essential role in identifying those artefacts that urgently require preservation. The workflow (Figure 4) presented by Strodl (2007) provides an overview of the different stages of identifying an effective preservation planning workflow. We intend making use of the workflow to ensure that our process will run smoothly. Phase one requires that one defines the purpose of the preservation activity, the selection/identification (also see Figure 6) of the records that should be preserved and formalising the organizational requirements. In phase two, attention is given to the selection of the preservation approach (Table 1 refers) and the testing of the selected approach(es). In the third and final phase the selected approaches are analysed according to previously determined criteria.



**Figure 4:** Digital preservation planning workflow (Strodl et al. 2007)

With basic understanding of the CSIR’s context as well as with input gained from a variety of sources and models it is then appropriate to say more about designing the CSIR’s preservation strategy. As a first step though it is necessary to understand what the essential elements of a good preservation strategy are.

## **4. ESSENTIAL ELEMENTS OF A PRESERVATION STRATEGY**

The statement has already been made that preservation goes beyond mere backups. Therefore the elements of preservation must be examined in more detail. Clarity is required to understand the implications of providing ongoing access to digital artefacts and the associated usability of those artefacts. On a basic level artefacts, whether digital or analogue, are subject to external influences that will affect the integrity of the artefacts, mainly due to changes in the digital environment. For digital artefacts, these threats include physical deterioration, e.g. exposure to external elements such as dust, magnets, excessive heat, direct sunlight, scratching and so forth. If these were the only threats, a centralized backup strategy would have been sufficient. Another, perhaps not so well known threat is digital obsolescence. As can be deduced from Figure 2, digital obsolescence is a serious and potentially 'lethal' threat for the CSIR's research artefacts. Since the emergence of the digital era, both software and hardware technology have been going through rapid development phases. The danger becomes apparent when either or both software and hardware are no longer available to access the artefacts (ICTU: Digital Preservation Testbed 2002). However, the question is raised whether an item should be preserved in such a manner that it is an exact representation of the original or whether preservation should focus on preserving the integrity while the contents remains authentic and usable even if it is not identical to the original item (ICTU: Digital Preservation Testbed 2002).

An effective digital preservation strategy should not only identify the risks but should also present workable and affordable solutions. The threats facing digital artefacts are relatively straightforward and will be described in more detail below. It is however, difficult to determine the area with the highest potential impact.

### **4.1. Managing changes in the digital environment**

In the digital environment changes are generally physical of nature, unlike the relative stability of paper, while the contents stays the same. Some of these changes are relatively simple to manage, whilst others have the potential to damage the content of the artefact.

Just like paper, the physical media can become damaged and will consequently require the transfer of data. Paper is relatively easily protected by applying climate control in a secure area. However, the protection of digital media is more complex than just physically protecting them against dust, excessive heat, moisture and magnets. It is not just damage to the physical format but the dangers resulting from changes in the format or media that should be managed. It is easy to transfer data between occurrences of the same or similar storage media without changing the contents of the artefact, e.g. transferring the data from a DVD to another DVD. The only benefit is that the storage medium remains 'fresh' and usable. In the context of the preservation strategy, transfer of data could be limited to backups and ensuring that the backups remain accessible as long as the software and technology remain the same.

However, accessibility of both the physical media and the content of the artefact are threatened as a result of technological development. Preservation of data will be required when new technologies or systems are implemented. Preservation is influenced by changes in system environments, operating systems or even programming languages. An example of preservation is the migration of data from WordPerfect to MSWord and recently to OpenOffice Writer. Another example is the migration of research data sets from a proprietary to an OSS environment. A very real threat is the corruption of data and the questions that can be raised regarding the integrity of the data.

## **4.2. Mitigating damages resulting from changes in the physical environment**

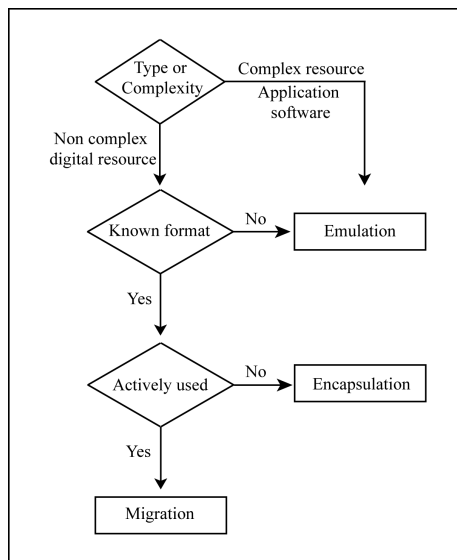
The Digital Preservation Testbed White paper (ICTU: Digital Preservation Testbed 2002) analyses seven preservation approaches. They point out that each of approaches has different technical requirements and cost implications. The first two approaches are valid but not deemed a usable solution and will not be discussed in this paper. They are technology preservation and printing-to-paper. Technology preservation is not deemed sustainable and printing-to-paper defeats the digitization efforts currently in place. The other approaches listed deserve more investigation regarding the suitability to the organization. These approaches are emulation, encapsulation, migration, XML, virtual machine software approach and conversion/digitization.

The virtual machine software approach can be regarded as a form of emulation in the sense that a 'Universal Virtual Computer (UVC)' is used as an interpreter. For future access, the UVC interpreter (language) should be installed on a computer prior to access (ICTU: Digital Preservation Testbed 2002). Unfortunately this doesn't seem to be a viable option for the CSIR at this time and will not be discussed in any more detail.

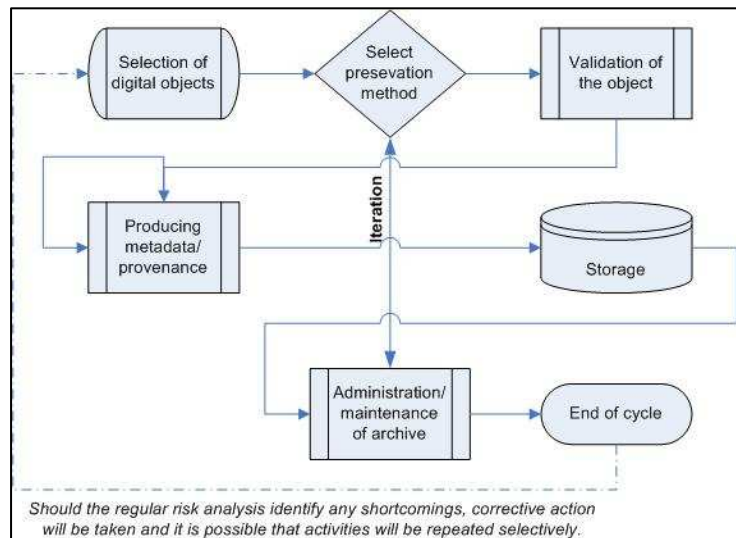
Although it is not a priority area conversion/digitization (of paper-based material into a digital/electronic media) is an option that will not be neglected. Our proviso is however that the digitization technology should be suitable to the OS environment. This option too will not be discussed in any further detail.

The remainder of the approaches need to be evaluated and selected in relation to the types of product and the variety of outputs associated with the CSIR. Preservation is definitely not a case of 'one size fits all'. In an audit recently conducted at the CSIR (CSIR 2008b), it was established that research data is created, manipulated, captured or analysed using approximately 500 different software products, excluding the different versions of each. This variety of products and outputs complicates the selection of a feasible approach. Fortunately not all of the outputs will fall into a category that will require preservation!

Deciding where to start will however remain a challenge. Fortunately several workflow models are available to assist. These models also help with the 'which-preservation-option-to-use' decision. To us the model depicted in Figure 5 and designed by Lee et al (2002) provides a schematic diagram that will help with the selection of the most appropriate preservation approach. Lee makes it clear that the complexity of the digital artefact will determine the approach used. Granger (2000), in turn, illustrates in Figure 6 where the selection of the preservation approach fits in with the overall preservation workflow.



**Figure 4:** Schematic diagram for selection of preservation technique (Lee et al. 2002)



**Figure 5:** Digital preservation workflow (Granger 2000)

ICTU's Digital Preservation Testbed White Paper (ICTU: Digital Preservation Testbed 2002) explains the basic principles of encapsulation, emulation, migration and XML while Lee et al (2002) also provides a comparison of the advantages and disadvantages between emulation, migration and encapsulation as options. (The advantages and disadvantages of XML have not yet withstood the test of time, but it does seem to be a viable approach.) Table 1 provides summarized information about each of the preservation techniques.

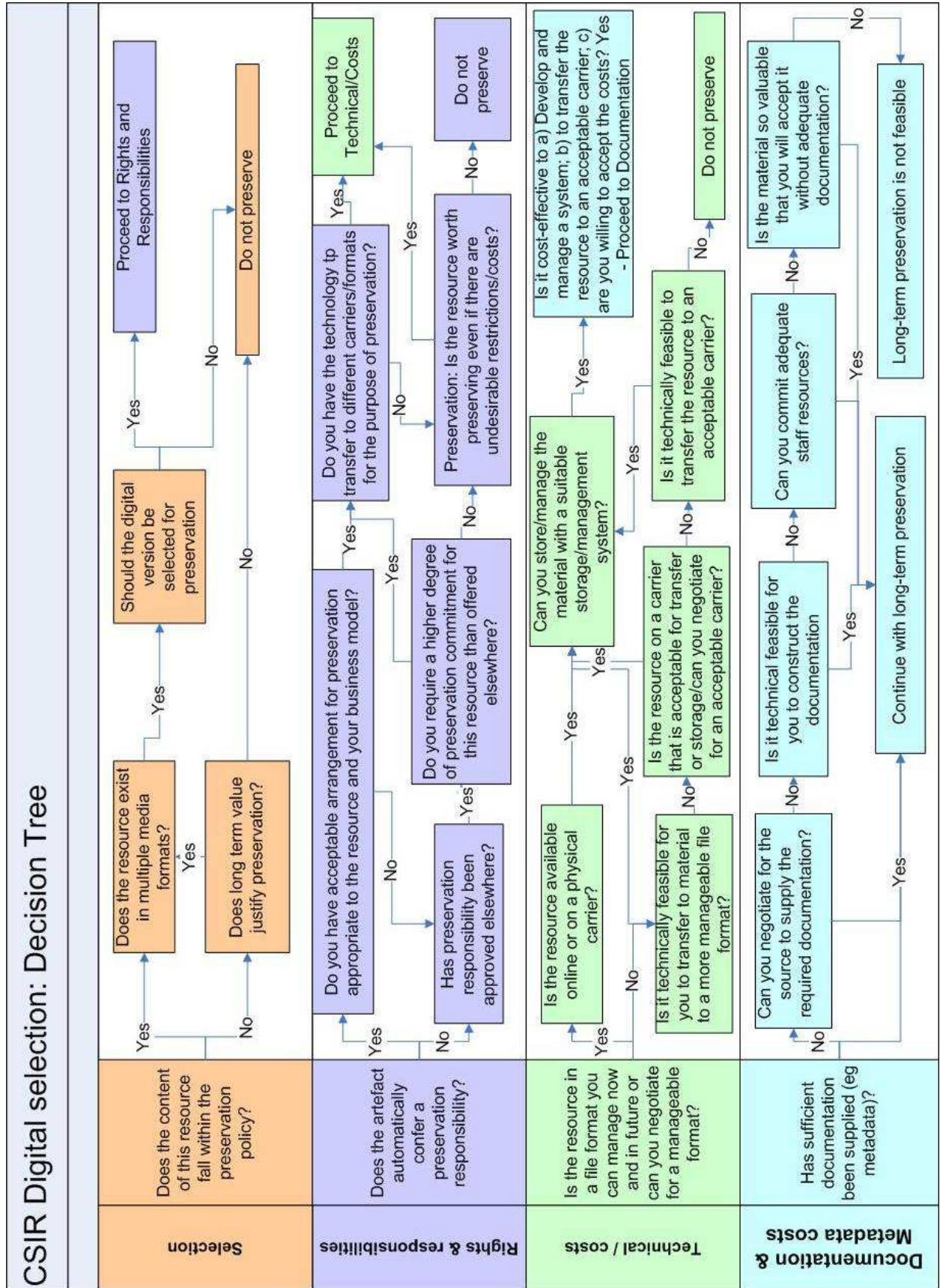
The selection of a suitable approach for preservation will also be affected by the retention period assigned to the artefacts within the records management file plan. The file plan clearly identifies those items where long-term preservation is required by law, those with a short or temporary value and those that are of value but where the curation responsibility has been handed over to the National Archives of South Africa (South Africa 1996).

In terms of acceptable file formats: the National Archives states that "At this stage PDF seems to be the better choice because an archival version known as PDF/A was recently published as an International Standard. XML also currently seems to be a format of choice since it is hardware and software independent and a text based, self describing, and human readable mark up language." Although the final format will have to meet the standards of the National Archives, items that have value for the CSIR might require a different approach. Therefore a complete understanding of each of the remaining formats is required.

All these preservation related decisions could become confusing but fortunately the preservation decision tree presented in Figure 7 provides more information that can be used for the selection process.

Table 1: Description, advantages and disadvantages of preservation techniques (Adapted from Lee et al. 2002; ICTU: Digital Preservation Testbed 2002; National Archives of Australia 2004)

<b>Technique</b>	<b>Advantage</b>	<b>Disadvantage</b>	<b>Domain</b>
Emulation/Recreation: provides access to the artefact in its original environment	Maintains the look and feel	Complexity of creating emulator specifications Large amount of information that must be preserved  Archaic software required to access information	Application software Complex digital resources such as those that contain executable files Resources for which there is a lack of sufficient knowledge Resources whose look and feel are important Very costly and required high level of diversified expertise.  IP and copyright issues might become stumbling blocks
Migration: transfer of files from one physical format (hardware/software) to another, e.g. MSWord to OpenOffice writer	Does not need to retain original applications  Supports active access and management	Substantial investment in resources for long-term preservation High probability of information degradation Lack of preservation metadata  Need for continued diligence on the part of archivists	Resources that are actively accessed and managed, such as scientific data or databases  Resources whose formats are sufficiently well known
Encapsulation: document is retained in the original format but contains instructions on how the original should be interpreted. Contains metadata.	Maintains preservation information	Knowledge about the format must be preserved Must be used with migration or emulation Systems required for capturing the digital information  Risks of not including important metadata	Resources that are unlikely to be accessed and managed actively Resources whose formats are sufficiently well-known
XML (Extensible mark-up language): Links with migration and encapsulation. The structure and the contents of the document is specified (can also be regarded as a form of migration although originals can be created by using XML)	Platform and program independent Open standard Extensible and controllable similar to natural language Readable by both humans and machines Seen as normalisation, stabilization and standardization  Reduces the need for repeated/regular migration	Still in pioneering phase Limited to text documents Expansion interfaces requires in-house development  Requires an interface to produce the original format of the content	Currently limited to text based documents



**Figure 7:** CSIR Decision Tree as adapted from the Decision Tree published by the Digital Preservation Coalition (2006)

Figure 7 is based on the Digital Preservation Coalition's Decision Tree (Digital Preservation Coalition 2006). The Decision Tree focuses on the selection of digital materials for long-term preservation and the same criteria can successfully be applied to the selection of the items that should be preserved in the long term.

### **4.3. Identifying a suitable preservation method**

Thibodeau (2002) states that preservation methods can be positioned according to three basic principles and he explains his principles as follows:

- On the "preserve technology" end of the spectrum there are the methods that attempt to keep data in specific logical or physical formats and to use technology organically associated with those formats to access the data and reproduce the object.
- In the middle of the spectrum we find the methods that migrate data formats as technology changes, enabling use of state-of-the-art technology for discovery, access and reproduction.
- On the "preserve objects" end of the spectrum, methods focus on preserving essential characteristics of objects independently of specific hardware or software.

Thibodeau (2002) also provides four criteria to use for the selection of a preservation method, namely feasibility, sustainability, practicality and appropriateness. If the intention is to preserve the digital object in its original format, it is doubtful whether the approach would meet any of the four criteria. It will be prohibitively expensive to try and maintain old hardware and software for a 'just-in-case' scenario. Sustainability won't be possible as it can be a case of a physical lack of expertise. Based on the aforementioned, it is by implication neither feasible nor appropriate to try and preserve the item in its *original* format. Logically speaking, valuable digital content should therefore be preserved *irrespective* of its original format. However, if it is still an object that is actively used, care should be taken to preserve it in a format that will facilitate future re-use of the content.

The work by Lavoie and Dempsey (2004) places the preservation activities further in perspective. They provide four prerequisites that are applicable to any method or approach used. These are: Coverage must be predictable and ongoing in order to remain trustworthy and to retain interests and participation; Interoperability and ease of use are essential and should not require specialised tools or skills; an authoritative record of all transactions is required to ensure privacy and integrity; and, preservability must be maintained to ensure sustainability and participation. Whichever system is in place must be tested regarding the system's preservability and method that will be used.

### **4.4. Objects requiring preservation**

Thibodeau (2002) subdivides the description of a digital object into three basic levels, namely physical, logical and conceptual.

- Physical level: this refers to the tangible format of the digital object namely the media on which the digital information is stored such as CDs, DVDs, magnetic tapes, etc. In Thibodeau's own words: '*A physical object is simply an inscription of signs on some physical medium*'
- Logical level: Thibodeau uses the term 'logical' to represent the interaction between the digital object and the software application. In Thibodeau's own words: '*A logical object is an object that is recognized and processed by software*'.

▪ Conceptual level: this level is typified by the representation of the objects by metadata, structure and contents but can be presented in more than one way without changing the meaning of the contents. In Thibodeau's own words: '*A conceptual object is the object as it is recognized and understood by a person ...*'

There seems to be a move towards focusing on the conceptual object rather than on the physical and logical objects, the argument increasingly being that it is the contents and its context that are important rather than the look and feel or the format in which it is produced. However, this approach raises concerns in terms of legally acceptable objects should the object not be a 100% true reflection of the original document – including the look and feel and format. The deciding factor will therefore be based on the legal implications and the liability associated with proving the integrity of the object. It can also be argued that system generated provenance should be sufficient to prove integrity. Furthermore, it can be argued that all preservation activities can raise questions regarding the integrity of an object. Fortunately, research data focuses more on the re-use of research data, which is the conceptual object, than either the physical or logical object. The value of the object is in relation to the value of the content, the metadata and its provenance.

## **5. Way forward**

As mentioned earlier, the CSIR make use of about 500 different software products. It is unrealistic to preserve all the digital objects created within the organization. It is also impossible to develop a 'one-size-fits-all' policy for preservation, especially the re-use of data, due to the organization's dispersed areas of research. For this reason, the Decision Tree approach (see Figure 7) will be used to identify and define those artefacts that will require preservation. In order to guarantee the sustainability of the preservation strategy a cost and benefit analysis will be included in the preservation policy as background. The cost of preservation needs to be benchmarked against the benefit associated with the preservation activities within the context of the artefact's intrinsic value with the emphasis on unit and organizational levels as illustrated in the value pyramid. A suitable approach, based on the artefact's format, will be selected and tested according to the process illustrated in the digital planning workflow. With the ten laws for preservation understood, even if not in place, the planning and implementation of the preservation strategy can commence. The strategy will include activities that will validate the object as well as the metadata. The CSIR has a very tight timeframe for the submission and implementation of the preservation strategy. The first target is the development and approval of the preservation policy within the 2009/2010 financial year followed by the development of the infrastructure, especially storage, and the identification and training of the required human resources. The preservation strategy coincides with the wider records management plan and it is expected that these initiatives will run parallel to each other in order to meet the digital curation targets.

## **6. References**

BERMAN, F. 2008. Got data? A guide to data preservation in the information age. *Communications of the ACM*, 51(12), pp. 50-56. Available from: <<http://portal.acm.org/citation.cfm?id=1409376>> (Accessed on 09 Jan 2009).

CSIR. 2008a. About us. Available from: <[http://www.csir.co.za/about\\_us.html](http://www.csir.co.za/about_us.html)> (Accessed on 09 Jan 2009).

CSIR. 2008b. Project Vula. Available from: <<https://vula.csir.co.za/quests/results.php>> (Accessed on 05 Jan 2009).

DIGITAL CURATION CENTRE. No date. the DCC curation lifecycle model. Available: <http://www.dcc.ac.uk/docs/publications/DCCLifecycle.pdf> (Accessed on 09 Jan 2009).

DIGITAL PRESERVATION COALITION, 8 March 2006, 2006-last update, decision tree: interactive assessment. Available: <http://www.dpconline.org/graphics/handbook/dec-tree.html> (Accessed on 09 Jan 2009).

GRANGER, S., RUSSEL, K. and Weinberger, E. 2000. *Cost elements of digital preservation*. Available from < <http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc> > (Accessed on 2 Jan 2009)

ICTU: DIGITAL PRESERVATION TESTBED, 2002-last update, XML and digital preservation [Homepage of ICTU] [Online]. Available from: < [http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white-paper\\_xml-en.pdf](http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white-paper_xml-en.pdf) > (Accessed on 09 Jan 2009).

LAVOIE, B. and DEMPSEY, L. 2004. Thirteen Ways of Looking at... Digital Preservation. *D-Lib Magazine*, 10(7/8), pp. 1082-9873. Available from: <<http://dlib.org/dlib/july04/lavoie/07lavoie.html>> (Accessed on 09 Jan 2009).

LEE, K.H., SLATTERY, O., LU, R., TANG, X. and MCCRARY, V. 2002. The State of the Art and Practice in Digital Preservation. *Journal of Research-National Institute of Standards and Technology*, 107(1), pp. 93-106. Available from: <<http://nvl.nist.gov/pub/nistpubs/jres/107/1/j71lee.pdf>> (Accessed on 9 Jan 2009).

NATIONAL ARCHIVES of AUSTRALIA. 2004. *Digital recordkeeping: guidelines for creating, managing and preserving digital records*. Canberra, Australia. Available from: <<http://www.nla.gov.au>> (Accessed on 26 Jan 2009).

SOUTH AFRICA. 1996. *National Archive and Records Service of South Africa Act 1996 (Act 43 of 1996)*. Available from: <[http://www.national.archives.gov.za/arch\\_act.htm](http://www.national.archives.gov.za/arch_act.htm)> (Accessed on 12 Jan 2009).

STRODL, S., BECKER, C., NEUMAYER, R. and RAUBER, A. 2007. How to choose a digital preservation strategy: evaluating a preservation planning procedure, *Proceedings of the 2007 conference on Digital libraries*, 2007, ACM Press New York, NY, USA pp 29-38. Available from: <<http://www.ifs.tuwien.ac.at/cp>> (Accessed on 12 Jan 2009).

THIBODEAU, K. 2002. Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. *The State of Digital Preservation: An International Perspective*. Available from: <[http://chnm.gmu.edu/digitalhistory/links/pdf/preserving/8\\_37e.pdf](http://chnm.gmu.edu/digitalhistory/links/pdf/preserving/8_37e.pdf)> (Accessed on 09 Jan 2009).