

The International Journal of Digital Curation

Issue 1, Volume 3 | 2008

The DCC Curation Lifecycle Model

Sarah Higgins,
Standards Advisor,
Digital Curation Centre, University of Edinburgh

June 2008

Summary

Lifecycle management of digital materials is necessary to ensure their continuity. The DCC Curation Lifecycle Model has been developed as a generic, curation-specific, tool which can be used, in conjunction with relevant standards, to plan curation and preservation activities to different levels of granularity. The DCC will use the model: as a training tool for data creators, data curators and data users; to organise and plan their resources; and to help organisations identify risks to their digital assets and plan management strategies for their successful curation.

The Digital Curation Centre (DCC) is committed to promoting a lifecycle approach to the management of digital materials, to enable their successful curation and preservation from initial conceptualisation to either disposal, or selection for reuse and long-term preservation.

The necessity of a lifecycle approach, to ensure the continuity of digital material, is discussed by Pennock (2007). Digital material, by its very nature, is susceptible to technological change from the moment of creation. The curation and preservation activities undertaken, or neglected, in different stages of their management, can influence the ability to look after them successfully at subsequent stages. A lifecycle approach ensures that all the required stages are identified and planned, and necessary actions implemented, in the correct sequence. This can ensure the maintenance of authenticity, reliability, integrity and usability of digital material, which in turn ensures maximisation of the investment in their creation.

A number of disciplines and projects have researched the lifecycle approach to digital asset management, and developed domain-specific lifecycle models as frameworks for planning. A review of these (Higgins, 2007) informed the development of a lifecycle model, which would specifically address the needs of digital curation.

“The Draft DCC Curation Lifecycle Model”, along with an open call for comment, was published in the *International Journal of Digital Curation* in December last year (Higgins, 2007) and publicised by means of a poster and flyer at the 3rd International Digital Curation Conference¹. The draft and call for comment were also sent to relevant e-mail lists and a discussion started in the DCC Forum². A number of expert comments were received and, where appropriate, were integrated into the draft model.

The DCC Curation Lifecycle Model (Figure 1) offers a graphical high-level overview of the lifecycle stages required for successful curation. Generic in nature, the model is indicative rather than exhaustive. When used as an organisational planning tool, it is adaptable to different domains, and extensible to allow curation and preservation activities to be planned at different levels of granularity. It can be used to: define roles and responsibilities; build frameworks of standards and technologies; and ensure that processes and policies are adequately documented. The model identifies: curation actions which are applicable across the whole digital lifecycle; those which need to be undertaken sequentially if curation is to be successful; and those which are undertaken occasionally, as circumstances dictate.

The DCC Curation Lifecycle Model complements, rather than restates, a number of standards, which provide frameworks for the management of digital materials. While the model provides a high-level view, it can be used in conjunction with relevant reference models, frameworks and standards to help plan activities at more granular levels. Workflow design, management issues, identification of processes and use of best practice can all be enhanced through the application of standards such as OAI (International Organization for Standardization [ISO], 2003) and ISO 15489 (ISO, 2001).

¹ 3rd International Digital Curation Conference <http://www.dcc.ac.uk/events/dcc-2007/>

² DCC Forum Index <http://forum.dcc.ac.uk/>

The DCC is using the model as a training tool to help curators understand the processes involved in successful curation, and develop curation and preservation methodologies for their organisations. The DCC SCARP Project³ is using the model in their immersive case studies, to help subjects identify the possible risks that current practice places on their data, and the steps required to mitigate these risks. The model is also being used by the DCC Services team to ensure resources cover all areas of the lifecycle. To improve accessibility, the model will shortly be used as the main access point to these resources on the DCC website⁴. Standards profiled by the DCC DIFFUSE Standards Frameworks Project⁵, are being contextualised through lifecycle action classification.

The model is not definitive and will undoubtedly evolve. The next stage of the project is the development of domain-specific variations to help further contextualise training and resources, while providing more tailored advice. If you have any suggestions on how to improve the DCC Curation Lifecycle Model (see Figure 1), or ideas for domain specific-variations, they can be emailed to the author⁶.

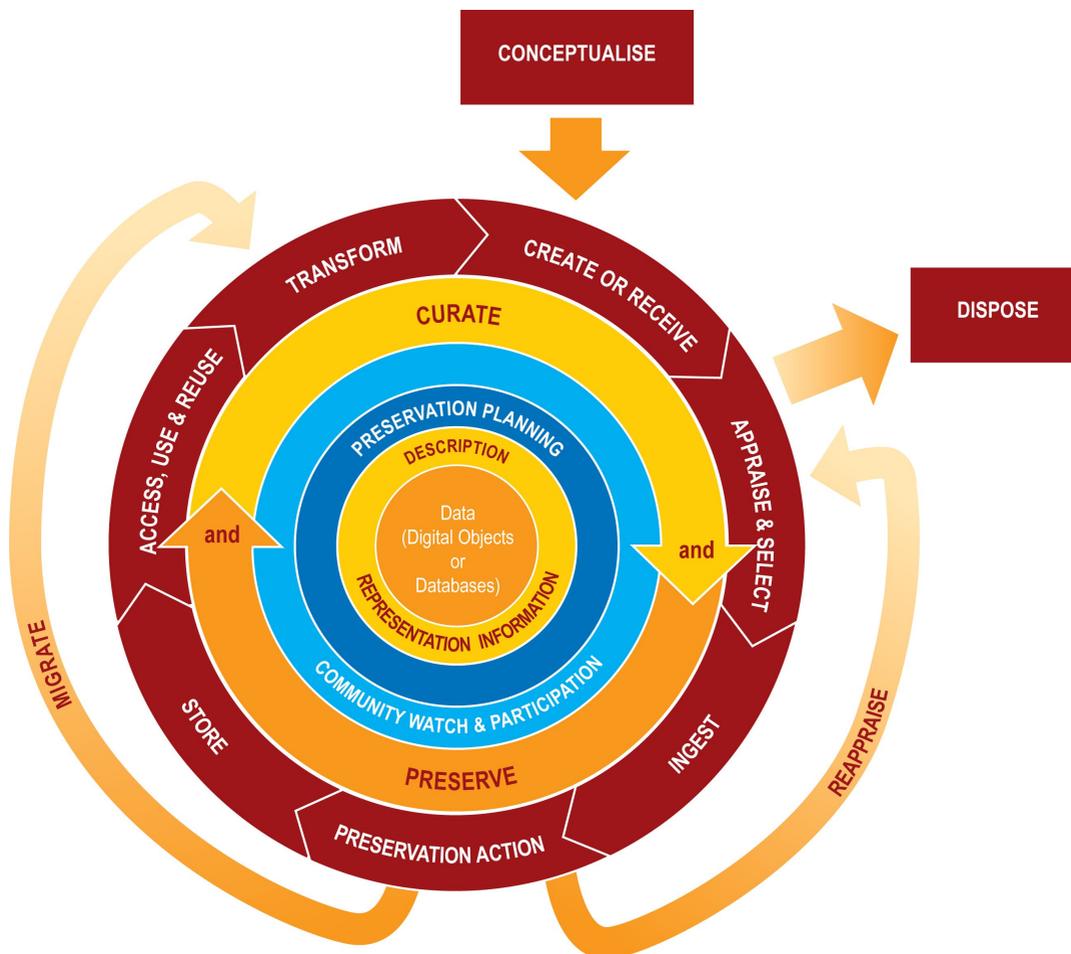


Figure 1. The DCC Curation Lifecycle Model.

³ Digital Curation Centre SCARP Project <http://www.dcc.ac.uk/scarp/>

⁴ Digital Curation Centre <http://www.dcc.ac.uk>

⁵ DCC DIFFUSE Standards Registry <http://www.dcc.ac.uk/diffuse>

⁶ Sarah Higgins e-mail Sarah.Higgins@ed.ac.uk

The Curation Lifecycle

The DCC Curation Lifecycle Model provides a graphical high-level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt. The model can be used to plan activities within an organisation or consortium to ensure that all necessary stages are undertaken, each in the correct sequence. The model enables granular functionality to be mapped against it; to define roles and responsibilities; and build a framework of standards and technologies to implement. It can help with the process of identifying additional steps which may be required, or actions which are not required by certain situations or disciplines, and ensuring that processes and policies are adequately documented.

Data (Digital Objects, or Databases)

	Data, any information in binary digital form, is at the centre of the Curation Lifecycle. This includes:
Digital Objects	<ul style="list-style-type: none"> - Simple Digital Objects are discrete digital items; such as textual files, images or sound files, along with their related identifiers and metadata. - Complex Digital Objects are discrete digital objects, made by combining a number of other digital objects, such as websites.
Databases	Structured collections of records or data stored in a computer system.

Full Lifecycle Actions

Description and Representation Information	Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long term. Collect and assign representation information required to understand and render both the digital material and the associated metadata.
Preservation Planning	Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.
Community Watch and Participation	Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.
Curate and Preserve	Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.

Sequential Actions	
Conceptualise	Conceive and plan the creation of data, including capture method and storage options.
Create and Receive	<ul style="list-style-type: none"> - Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation. - Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.
Appraise and Select	Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.
Ingest	Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.
Preservation Action	Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.
Store	Store the data in a secure manner adhering to relevant standards.
Access, Use and Reuse	Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.
Transform	<p>Create new data from the original, for example</p> <ul style="list-style-type: none"> - By migration into a different format. - By creating a subset, by selection or query, to create newly derived results, perhaps for publication.

Occasional Actions	
Dispose	Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data are destroyed. The data's nature may, for legal reasons, necessitate secure destruction.
Reappraise	Return data which fails validation procedures for further appraisal and reselection.
Migrate	Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.

Acknowledgements

I would like to thank not only members of the DCC Staff but also the various experts who submitted contributions to this article. Additional thanks go to Chris Blackall for suggestions in respect of graphics.

References

- Higgins, S. (2007). Draft DCC curation lifecycle model. *International Journal of Digital Curation*. Vol. 2 (2), p 83. Retrieved June 18, 2008, from <http://www.ijdc.net/>
- International Organization for Standardization. (2001). Information and documentation. Records management. General and ISO/TR 15489-2:2001 Information and documentation. Records management. Guidelines. Retrieved June 18, 2008, from http://www.iso.org/iso/catalogue_detail?csnumber=31908
- International Organization for Standardization. (2003). Space data and information transfer systems — Open archival information system - Reference model. ISO 14721:2003. Retrieved June 18, 2008, from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=24683
- Pennock, M. (2007). Digital curation: A life-cycle approach to managing and preserving usable digital information. *Library and Archives Journal, Issue 1*. Retrieved (preprint) June 18, 2008 from http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf

Bibliography

- Boehm, B (1988), *A Spiral Model of Software Development and Enhancement*. Retrieved June 18, 2008, from <http://www.cs.usu.edu/~supratik/CS%205370/r5061.pdf>
- Chen, Y., Chen, S., & Lin, S. (2003). A metadata lifecycle model for digital libraries: methodology and application for an evidence-based approach to library research. In *Documents in Information Science, Working Papers Series*. Retrieved June 18, 2008, from http://www.ifla.org/IV/ifla69/papers/141e-Chen_Cheng_Lin.pdf
- Dukes, P. (2006). *Making the most of our data: MRC's data sharing & preservation initiative*. Retrieved June 18, 2007, from <http://www.rin.ac.uk/files/Peter%20Dukes%20v2.pdf>
- Green, A. (2006). Conceptualizing the digital life cycle. *IASSIST Communiqué*. Retrieved June 18, 2008, from <http://iassistblog.org/?p=26>
- Humphrey, C. (2006). e-Science and the Life Cycle of Research. Retrieved June 18, 2007, from datalib.library.ualberta.ca/~humphrey/lifecycle-science060308.doc (sic).

-
- InterPARES 2 Project. (2007). *Chain of preservation (COP) model*. Retrieved June 18, 2007, from http://www.interpares.org/ip2/ip2_models.cfm#
- Knight, G. (2006). A lifecycle model for an e-print in the institutional repository. Retrieved June 18, 2007, from http://www.sherpadp.org.uk/documents/wp213-lifecycle_model.pdf
- Lord, P., & Macdonald, A. (2003). *e-Science Curation Report: Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision*. Retrieved June 18, 2007, from http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal.pdf
- Lyon, L. (2003, July). eBank UK: Building the links between research data, scholarly communication and learning, *Ariadne, Issue 36*. Retrieved June 18, 2007, from <http://www.ariadne.ac.uk/issue36/lyon/>
- Paradigm Project. (2006). Workbook on Digital Private Papers Draft lifecycle for the long-term preservation of digital archives. Retrieved June 18, 2007, from <http://www.paradigm.ac.uk/workbook/introduction/paradigm-lifecycle.html>
- Royce, W. (1970). *Managing the development of large software systems*. Retrieved June 18, 2007, from <http://www.cs.umd.edu/class/spring2003/cmse838p/Process/waterfall.pdf>
- Rusbridge, C. (2005). *Information Life Cycle and Curation*. Retrieved June 18, 2007, from <http://www.dcc.ac.uk/docs/dcc-life-cycle.ppt>
- Rusbridge, C. (2007). *Curation mind map*. (Unpublished diagram.)
- Serco. (2008). Model requirements for the management of electronic records update and extension –Version 1.02. Retrieved June 18, 2008, from <http://www.moreq2.eu/downloads.htm>
- Upward, F. (1997). Structuring the record continuum part one: Post custodial principles and properties. *Archives and Manuscripts, vol 24, no 2*, pp 268-285.
- Wheatley, P., Ayris, P., Davies, R., Mcleod, R., & Shenton, H. (2007). The LIFE model v1.1..Discussion paper. LIFE Project: London.