Collaboration to Clarify the Cost of Curation

A Draft Economic Sustainability Reference Model

Deliverable Lead:	JISC
Related Work package:	WP4
Author(s):	Brian Lavoie (OCLC), Neil Grindley (Jisc),
Dissemination level:	CO (until public draft)
Submission date:	31st July 2013
Project Acronym:	4C
Website:	http://4cproject.eu
Call:	FP7-ICT-2011-9
Project Number	600471
Instrument:	Coordination action (CA)—ERA-NET
Start date of Project:	01 Feb 2013
Duration:	24 months

Project funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Version History

Version	Date	Changed pages / reason	Modified by
0.1	03 June 2013	Project Release for initial feedback - Incomplete draft with placeholder text sections	NG
0.2	24 June 2013	Extra text included	NG
0.3	19 July 2013	Comprehensive redrafting of sections 1 and 3	BL
0.4	24 July 2013	Revising and updating sections 5and 6	NG
0.5	21 August 2013	Revising and updating section 4	BL
0.7	22 August 2013	Consolidation of sections, and changes to section 6 and revision of overall structure	NG
1.00	05 September 2013	QA by DNB & Jisc – MS9	NG/KH

Acknowledgements

This report has been developed within the project “Collaboration to Clarify the Cost of Curation” (4cproject.eu). The project is an ERA-NET co-funded by the 7th Framework Programme of the European Commission.

The authors of this report would like to acknowledge and thank Chris Rusbridge for his original contributions to the ESRM. A ‘pre- 4C’ version of the model (v0.6) can be found at:

http://unsustainableideas.files.wordpress.com/2011/09/reference-model-0-6-clean.pdf

The 4C participants are:

Participant organisation name	Short Name	Country
Jisc	JISC	UK
Det Kongelige Bibliotek, Nationalbibliotek Og Kobenhavns Universitetsbibliotek	KBDK	DK
Instituto de Engenharia de Sistemas e Computadores, Investigacao e Desenvolvimento em Lisboa	INESC-ID	PT
Statens Arkiver	DNA	DK
Deutsche Nationalbibliothek	DNB	DE
University of Glasgow	HATII-DCC	UK
University of Essex	UESSEX	UK
Keep Solutions LDA	KEEPS	PT
Digital Preservation Coalition Limited by Guarantee	DPC	UK
Verein Zur Forderung Der It-Sicherheit In Osterreich	SBA	AT
The University of Edinburgh	UEDIN-DCC	UK
Koninklijke Nederlandse Akademie van Wetenschappen -Knaw	KNAW-DANS	NL
Eesti Rahvusraamatukogu	NLE	EE

Author(s):	Neil Grindley (Jisc)
	Brian Lavoie (OCLC)

Table of Contents

Figures

Executive Summary

This report sets out the latest draft version of the Economic Sustainability Reference Model (ESRM) and provides ancillary material that will help the reader to understand aspects of the model and its purpose relative to the 4C Project. It also offers a suggested implementation of the model in the form of a Sustainability Self-Assessment Questionnaire in Appendix 1.

The purpose of a reference model is to help cut through complexity and highlight key concepts, relationships, and decision points in a complex problem space. It should also help the user of the reference model to benchmark and compare their own local models with an exemplar which has broad acceptance from a community that is relevant to the activity in question. It is clear that the economics of sustaining digital assets (by means of digital curation) is exactly the kind of complex area that might benefit from such a reference model.

Building on the work of the Blue Ribbon Task Force for Sustainable Digital Preservation and Access, the ESRM is best understood as a strategic tool for planning and discussion aimed at executive and managerial rather than operational level staff. Its purpose is to provide a foundation for progress in the development of successful sustainability strategies for digital curation. It does this by organizing the problem space; providing a common reference point of concepts and vocabulary; and introducing a layer of abstraction that hides the complexities and idiosyncrasies of individual implementations and contexts, while at the same time embodying sufficient detail to support substantive discussions of shared issues.

The ESRM organizes its discussion of economically sustainable digital curation around the concept of a sustainability strategy which is composed of several key components:

- Economic lifecycle: the economic “context” in which curation occurs
- Sustainability conditions: the conditions that must be met to achieve sustainability
- Key entities: the curation environment, including digital assets; the curation process; and stakeholders that must make decisions to ensure the sustainability conditions are met
- Economic uncertainties: forms of uncertainty that present challenges in regard to achieving the sustainability conditions

Some high level assertions can also be extracted from the model which may provide a good starting point for general discussion (and which relate to the key entities section).

- The properties of digital objects must be realistically examined and understood before they can be defined as assets;
- Sustaining digital assets (the process) is complicated but not impossible to organise and understand if you take the view that it is about ensuring there is sufficient ‘return on investment’;
- The assets and the processes are not themselves intrinsically important: the crucial issue is whether the value of the assets is sufficiently understood by the relevant stakeholders

The report contains a full elaboration of the ESRM and the objective of this phase of the 4C project is to make this version widely available to the community for comment, critique and use. As it stands the ESRM is already a useful engagement tool for the 4C Project but the model itself will only become an effective and useful *reference model* if a consensus builds around it.

1 Introduction

From a purely technical perspective, digital curation aims to secure the ongoing availability and usability of digital assets judged to be of sufficient value to warrant curatorial attention. But the technical processes through which digital curation achieves its aims do not occur in a vacuum: rather, they are usually embedded within an organization or some sort of programmatic activity that is responsible for organizing, directing, and *sustaining* these processes. These components – the organization and the processes – are not easily separable. In this sense, when we speak of sustaining digital assets over time, we cannot separate this discussion from the question of how to sustain the organizations or activities responsible for curation. And this question is chiefly one of economics.

Digital curation, like any other activity, is sustained by people making decisions to allocate sufficient resources, on an ongoing basis, to achieve well-articulated goals. This resource allocation decision must be based on a thorough understanding of the long-term costs of digital curation – i.e., the required investment – as well as the anticipated benefits from curation – i.e., the expected return on investment. Sustainability strategies that focus on costs to the exclusion of benefits, or vice versa, are not in fact sustainability strategies at all, and in all likelihood are doomed to fail. Sustainable digital curation activities must operate with a clear view of expected costs *and* expected benefits, with the corollary that if the former is perceived to exceed the latter, the curation activity should not be initiated, or if it is already underway, it should be stopped.

This is not to suggest that constructing a viable sustainability strategy is easy. Rather, it involves establishing and maintaining a complicated and often delicate interplay between a variety of stakeholders with different and perhaps even contradictory curation incentives. It requires creating and maintaining resource flows, often from multiple sources. It is a matter of choosing priorities in the face of budget constraints, including tough choices on which digital assets to select for curation, and occasionally, which curated assets should be abandoned in favour of higher-value uses for scarce resources. It calls for the development of contingency plans to address the blanket of uncertainty covering most aspects of long-term digital curation. Broadly speaking, designing a sustainability strategy might be imagined as beginning with the unachievable goal of curating everything for all time, and working backward to some realistic approximation of this ideal, given the circumstances and constraints prevailing in a particular curation context.

As this suggests, the economics of digital curation is a very complicated landscape to navigate. In circumstances such as this, a reference model is a useful tool to help cut through the complexity and highlight the key concepts, relationships, and decision points in a complex problem space. OASIS, an organization that promotes open information standards, (https://www.oasis-open.org/committees/soa-rm/faq.php) defines a reference model as “an abstract framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment.” Put another way, a reference model is a template that identifies and organizes the components that must be part of a solution in a particular problem space. It defines concepts and an associated vocabulary that promote a shared understanding of these components, and their relationships to one another. And it achieves these aims that by adopting a level of abstraction sufficient to lift the reference model above the specific circumstances or idiosyncrasies of specific manifestations of the problem space, thus extending its applicability across many contexts.

Digital curation practitioners are familiar with the Open Archival Information System (OAIS) reference model, which provides a framework and vocabulary for the technical/workflow aspects of long-term

digital preservation/curation. It has been remarkably successful in galvanizing cross-domain interactions, and communicating knowledge within the broader digital preservation community. It was instrumental in establishing a foundation of shared knowledge about the basic contours of digital archiving systems, independent of technical implementation choices or domain/project context. Having this shared foundation of knowledge was a significant step in overcoming the technical challenges of digital preservation.

The Economic Sustainability Reference Model (ESRM) described in this document is intended to accomplish for the economics of digital curation what OAIS did for the technical aspects of digital curation. It proposes a framework within which to orient discussions of economically sustainable digital curation. These discussions can take place in multiple contexts: internally (within a particular curation project), or externally (across various stakeholder groups); focusing on a particular aspect of the sustainability problem (e.g., costs), or considering the problem as a whole; designing a new sustainability strategy from the ground up, or assessing the merits of an existing strategy, or comparing the features of multiple strategies. An ESRM would be a useful tool in all of these contexts, and more. Again, curation practitioners can draw parallels here between the anticipated uses of an ESRM, and the uses to which OAIS has been successfully applied.

It is important to emphasize that an ESRM is best understood as a strategic tool for planning and discussion aimed at executive and managerial rather than operational levels. It is not, nor should it be, an off-the-shelf "template" for building a sustainability strategy for digital curation. Understanding whether and how resources can be made available and having a clear view of incentives or motivations to continue to practice digital curation is strategic work and slightly different to the work of practitioners who must have a detailed understanding of the tools, methods and standards that are applicable to the task of curating digital assets.

A Draft Economic Sustainability Reference Model (ESRM) is an early output of the 4C Project and seeks to address and support views of digital curation from a strategic and high-level perspective. It builds on existing work (http://unsustainableideas.wordpress.com/economic-sustainability-ref-model-page/) that was conceived and developed by Brian Lavoie (OCLC) and Chris Rusbridge (Chris Rusbridge Consulting) following their central involvement with the influential Blue Ribbon Task Force for Sustainable Digital Preservation and Access (BRTF). The BRTF's final report, *Sustainable Economics for a Digital Planet*[1], explained how achieving economic sustainability is complicated by the difficulty in coordinating benefits, incentives to preserve, and roles and responsibilities across the community of stakeholders attached to a particular set of digital assets. The report provided a wealth of findings and recommendations to assist decision-makers in thinking through the issues associated with achieving economic sustainability.

Lavoie and Rusbridge's work, co-sponsored by JISC and OCLC, was a first attempt at translating the concepts, findings, and recommendations of the BRTF report into a *reference model for economically sustainable digital curation*. It defines the notion of a sustainability strategy; highlights the key concepts planners must take into account when designing that strategy; and enumerates the kinds of economic risks that the sustainability strategy should defend against. The goal was to build a tool to aid planners in building, clear, persuasive arguments to help unlock the resources needed to support sustainable digital curation activities. It is a way for planners to organize their thinking about the features of the digital assets, stakeholders, and the digital curatorial process that shape the economic risks impacting a given digital curation activity, and to design remedies to overcome these risks and maximize the prospects for achieving sustainability. It focuses not just on cost, but also value to society over the long term, emphasizing that the true metric for a project's economic viability is not merely its cost, but the expected value returned net of costs.

This document revises and updates Lavoie and Rusbridge's draft model, and situates it as a key component of the 4C Project. The ESRM is one of the outputs of the *Enhancement* working group (WP4) and aligns well with its remit to think more widely and holistically about the topic of sustainability. However, development of a reference model is an incremental process that by its nature needs input and endorsement from a broad community before it can lay claim to being a valid model. Therefore, it is expected that the ESRM presented in this document will continue to evolve through an ongoing process of community consultation and feedback.

2 Background

Sorting out the economic challenges of long-term digital curation are as necessary as solving the technical issues. In the absence of a sustainability strategy designed to support clear, realistic objectives, any digital curation initiative, no matter how worthy in spirit, is doomed to fail in the practical world of limited resources, competing priorities, and fleeting attention. Despite the centrality of economic issues in the broader problem of ensuring that valuable digital assets persist and are available for use, the topic has until recently suffered from relative neglect in the expanding digital curation literature. While a number of studies have analyzed digital curation costs, or sought to regularize cost data into widely-applicable frameworks, very little work has dealt with the broader idea of sustainability – that is, a holistic view of the costs, benefits, stakeholder incentives, and roles and responsibilities associated with the long-term curation of digital assets.

This gap in the literature was addressed by the final report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access (BRTF). Published in 2010, *Sustainable Economics for a Digital Planet*[2] aimed to provide a comprehensive view of the concept of sustainability as it relates to the long-term stewardship of digital materials. The report is organized around an analysis of four important digital curation scenarios: scholarly discourse, research data, commercially-owned cultural content, and collectively produced web content. Analysis of these scenarios produced both context-specific and general findings and recommendations, in which key issues in achieving economic sustainability are identified, with actions recommended for addressing them. One of the Task Force's key findings was that the challenges of achieving economically sustainable digital curation will not be solved by simply making more funds available; instead, the prospects of achieving economic sustainability are maximized when resource allocation occurs in an environment where there is a recognition of the value of investing in preservation, where appropriate decision-makers perceive clear incentives to act, and where stakeholders throughout the digital lifecycle understand their roles and responsibilities in terms of contributing to the effort to ensure digital materials persist over the long term.

Following the release of the BRTF report, two symposia were held – one in the United States and one in the United Kingdom – to give interested parties an opportunity to react to and discuss the Task Force's findings. Domain experts representing the four curation contexts discussed in the report gave their views on the findings and recommendations in the report, and the priorities for carrying the work forward. More generally, the symposia were intended to begin the process of transferring the locus of discussion of economically sustainable digital preservation from the Task Force to the community, and inspire some ideas for future work, building on the foundation provided by the BRTF report.

In imagining ways to build on the foundations established by the BRTF report, one path forward is to translate the concepts, findings, and recommendations of the report into a framework that organizes the complex sustainability problem space, and provides some common reference points to guide those tasked with navigating this space in the course of designing effective sustainability strategies for digital curation activities. In light of this, JISC and OCLC Research jointly sponsored some preliminary explorations of the utility of establishing a *reference model for economically sustainable digital curation*. The purpose and value of such a reference model is discussed in more detail in the next section.

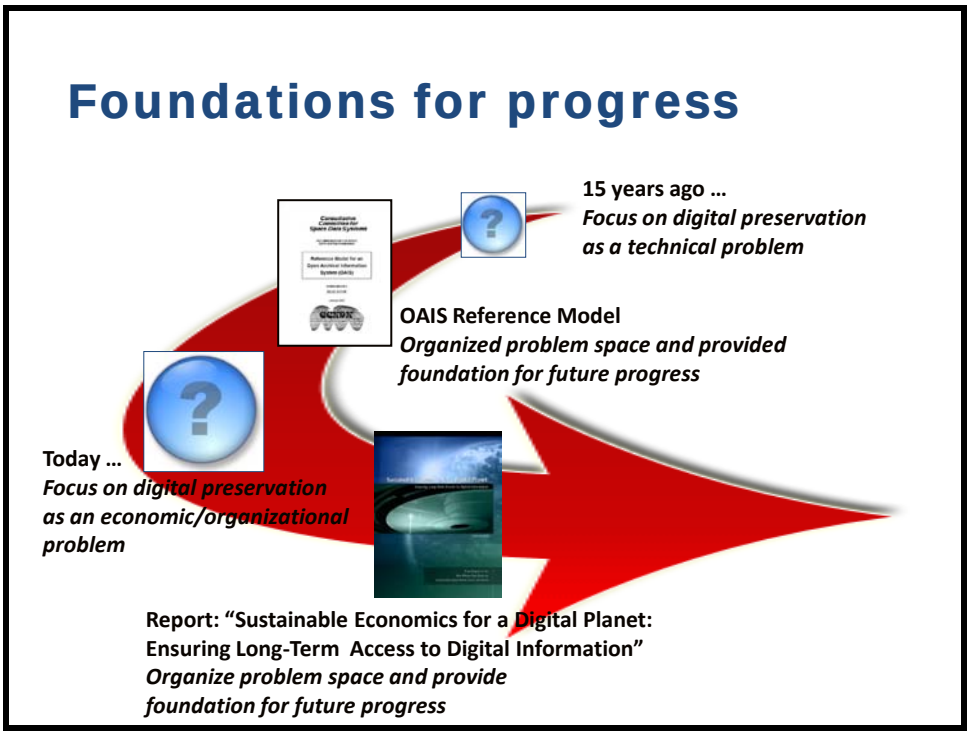
This work, undertaken in 2011 by Brian Lavoie and Chris Rusbridge, produced a draft reference model, based on the concepts and findings described in the BRTF report. The draft has been promulgated and discussed through a series of workshops and other events, with feedback generally being positive, but with a clear understanding that much more work remains to be done to establish the reference model on

a similar footing to that achieved by OAIS. The version of the reference model presented in this report updates Lavoie and Rusbridge's original draft. As explained in Section 6, this and future work on the reference model is taking place under the auspices of the 4C Project.

3 Purpose of the ESRM

The purpose of the ESRM is to provide a foundation for progress in the development of successful sustainability strategies for digital curation. The well-known OAIS reference model catalyzed progress in the development of technical solutions for digital repositories by organizing the problem space, providing a common reference point of concepts and vocabulary, and introducing a layer of abstraction that hid the complexities and idiosyncrasies of individual implementations and contexts, while at the same time embodying sufficient detail to support substantive discussions of shared issues. In the same way, the ESRM proposes to catalyze progress in creating sustainable solutions in the domain of digital curation, by performing these same functions in regard to economic issues related to the long-term curation of digital assets.

Figure 1 – Foundations for Progress



An example serves to illustrate these points. It is not unusual to encounter confusion in discussions about economically sustainable digital curation in regard to the scope of the topic. Some participants equate a sustainability strategy with a cost model. Accurate cost projections based on a good cost model are indeed a key aspect of any sustainability plan, but they are far from the whole story. As the BRTF report powerfully argues, achieving sustainability goes well beyond an understanding of how much it will cost to preserve a gigabyte of data for a given period of time. A reference model that defines the concept of sustainability, its components, and the relationships between these components, is a useful tool for promoting a shared and complete understanding of the problem space on the part of all participants in discussions that bring together stakeholders with disparate backgrounds and perspectives.

As this example illustrates, a reference model like the ESRM – or OAIS, for that matter – can serve as an aid in discussions about a particular problem space. It defines the key concepts associated with the problem space and how they relate to one another, and it does so at a sufficient level of abstraction that the discussion can transcend the specifics of a particular context without losing its substance. These

discussions can take a variety of forms: they may be intended to educate a community about the problem space the reference model describes; they may be an attempt to cultivate understanding and consensus across stakeholders with a shared interest in a range of issues associated with the problem space; they may be aimed at identifying opportunities to promote standards or shared practices in regard to certain aspects of the problem space; they may be an effort to compare proposed plans or strategies against a widely understood benchmark. In each of these cases – education, consensus-building, standardization, and benchmarking, as well as other applications – an ESRM can serve as a foundation for progress in overcoming the challenges of economic sustainability in digital curation, much as OAIS has done and continues to do in regard to the technical aspects of digital curation.

As noted in the previous section, the ESRM proposed in this document is largely based on the findings and recommendations in the BRTF final report. The specific intersections of the BRTF report on the one hand, and the ESRM on the other, are noted below.

4 The Draft Framework

Digital curation activities adopt a technical strategy to orchestrate the technical processes and workflows needed to maintain the long-term accessibility of digital materials: e.g., redundant storage, regular media refreshment, format migration scheduling, and so on. In the same way, a digital curation activity must develop a *sustainability strategy* to orchestrate the economic factors necessary to ensure that the activity has adequate resources to meet its long-term curation goals.

The OAIS reference model employs the notion of an Open Archival Information System as an organizing concept around which to present a general description of the key technical components of a digital preservation repository. The ESRM organizes its discussion of economically sustainable digital curation around the concept of a sustainability strategy. In the same way that repository planners aim to build an Open Archival Information System to address the technical foundations of digital curation, they must also aim to design a sustainability strategy to establish the parallel economic foundations. A sustainability strategy is composed of several key components:

- Economic lifecycle: the economic "context" in which curation occurs
- Sustainability conditions: the conditions that must be met to achieve sustainability
- Key entities: the curation environment, including digital assets; the curation process; and stakeholders that must make decisions to ensure the sustainability conditions are met
- Economic uncertainties: forms of uncertainty that present challenges in regard to achieving the sustainability conditions

Each of these components is explained in detail in the following sections.

The ESRM maps out the key elements of the problem space planners face when designing a sustainability strategy for their digital curation activities. The reference model focuses on the general concept of a sustainability strategy, breaks it down into its key components, and draws planners' attention to the properties of those components most relevant for economic sustainability. While the reference model does not "solve" any digital curation activity's sustainability issues, it provides a framework within which the nature of these issues can be understood, and appropriate solutions designed.

4.1 Summary of the Reference Model

Economically sustainable digital curation is a complex topic even when discussed only in generalities. It becomes even more complex when applied to the particular circumstances of a specific digital curation activity. The ESRM helps planners navigate this difficult terrain by defining a general concept that represents the "target" planners need to aim for – *a sustainability strategy*. A sustainability strategy is the means by which a digital curation activity orchestrates the economic factors necessary to ensure that the activity has adequate resources to meet its long-term curation goals. **The sustainability strategy is described in Section 4.2.**

A sustainability strategy can be broken down into four primary components: 1) an economic lifecycle; 2) sustainability conditions; 3) key entities; and 4) economic uncertainties. The *economic lifecycle* can be understood as the background against which a digital curation activity, and hence its sustainability strategy, operates. It conveys the notion that economic decision-making does not take place in a static environment, but rather a dynamic one that continually cycles through a progression of events, each of which may potentially impact economic sustainability. The first step in designing an effective sustainable

strategy is to understand the economic lifecycle over which that strategy is expected to operate. **The economic lifecycle is described in Section 4.3.**

The BRTF final report defines five key *sustainability conditions* that must be met in order to maximize the likelihood that a digital curation will remain economically sustainable over the long term. These conditions include: value; selection; incentives; resources; and organization/governance. A sustainability strategy that fails to address one or more of these conditions will likely be an ineffective means to support a digital preservation activity capable of meeting its long-term goals. **The five sustainability conditions are defined in Section 4.4.**

Meeting the sustainability conditions requires a close understanding and coordination of the *key entities* relevant to economic decision-making for digital curation: digital assets; the curation process; and the stakeholder ecosystem. Digital assets are what we invest in when we allocate resources to digital curation; the curation process is the mechanism by which our resource allocation is transformed into a "return on investment" (i.e. the ongoing availability of valuable digital assets); and the stakeholder ecosystem is the network of individuals or organizations that reap the benefits of sustainably curated digital assets, and/or contribute toward the curation process which maintains them. Understanding the properties of these entities, and how they manifest and relate to one another in a particular digital curation context is an essential component of a robust sustainability strategy. **The key entities are described in Section 4.5.**

Finally, *economic uncertainties* are potential obstacles to achieving sustainability. More specifically, economic uncertainties manifest themselves as challenges to achieving one or more of the five conditions required for economic sustainability. While these uncertainties cannot necessarily be completely avoided, a good sustainability strategy must be cognizant of the potential for their occurrence, and include appropriate plans and contingencies to address them. **A list of the general ways economic uncertainties impact digital curation, and a simple framework of possible "reactions" to economic uncertainty, are presented in Section 4.6.**

4.2 The Sustainability Strategy

A sustainability strategy orchestrates the economic factors necessary to ensure that a digital curation activity has adequate resources to meet its long-term goals. It is a plan that organizes the key entities associated with long-term sustainability – digital assets, the curation process, and stakeholders – in such a way that the curation activity becomes a *sustainable economic activity*. A sustainability strategy is enacted over the entire economic lifecycle of the digital assets, guiding the curation activity through the significant events that occur during this lifecycle. It is designed to address each of the five sustainability conditions enumerated in the BRTF final report. Finally, a sustainability strategy addresses significant economic uncertainties that may occur during the economic lifecycle by incorporating plans and contingencies to mitigate the effects of these uncertainties.

The essential components of a sustainability strategy may be understood as:

- **Economic lifecycle:** the *dynamic pattern*, or sequence of events, against which the sustainability strategy operates;
- **Sustainability conditions:** the *conditions* the sustainability strategy must address in order to achieve sustainability;
- **Key entities:** the key elements of the economic *environment* – digital assets, the curation process, and stakeholders – whose properties and relationships shape the circumstances in which the sustainability strategy operates;

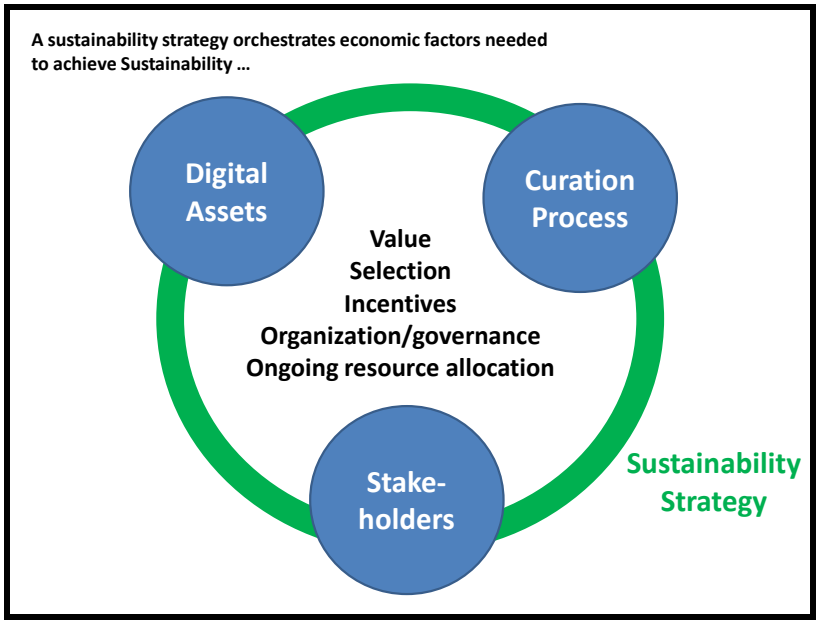
- **Economic uncertainties:** *frictions and obstacles* that may potentially act to impede the ability of a repository to achieve economic sustainability; the sustainability strategy must anticipate these uncertainties and if necessary, mitigate them.

In short, a sustainability strategy is designed to account for time, the economic environment, and uncertainties while meeting certain conditions for economic sustainability. At the highest level of abstraction, it is what repository managers must aim for to secure the economic sustainability of their digital curation activities, in the same way that the concept of an OAIS is, at the highest level of abstraction, what planners must aim for to assure the technical sustainability of their archived digital assets.

A sustainability strategy is not unlike the perhaps more familiar concept of a *business plan*. According to Wikipedia, a business plan is a "formal statement of a set of business goals, the reasons why they are believed attainable, and the plan for reaching those goals"[3]. In the same way, a sustainability strategy articulates a realistic set of curation goals, and the means for achieving them. Like a business plan, a sustainability strategy functions like a roadmap in helping decision-makers move curated digital assets smoothly across the economic lifecycle. A sustainability strategy may be more generalized, potentially expressing both value and costs in different (non-monetary) terms. However, for the particular case where a sustainability strategy is built on expected revenue, the sustainability strategy could collapse into a business plan.

Achieving sustainability requires planners to take into account the properties of digital assets, the curation process, and stakeholders, and align them in such a way that each of the sustainability conditions is addressed. The result is a sustainability strategy. It should be noted that a successful sustainability strategy is not one that guarantees long-term economic sustainability. Regrettably, such a strategy does not exist. Instead, a successful sustainability strategy is one that *maximizes the prospects* of achieving sustainability by cultivating a thorough understanding of the conditions in the economic environment relevant to the five sustainability conditions.

Figure 2 – Sustainability Strategy Components



There can be no "off-the-shelf" sustainable strategy as no two digital curation contexts are exactly alike. One reason is that stakeholders (and by extension, curation decision-makers) are humans, who will make

judgments based on perceptions of value and cost. The properties of the three entities often represent a conglomeration of conditions that can operate either to promote or to discourage sustainability. And since the nature of digital assets, the curation process, and stakeholders will differ from context to context, planners will need to think through the conditions relevant to their particular context and design a sustainability strategy that is the best fit (or at least a reasonable fit) for those circumstances.

So in reality the design of a sustainability strategy will first need to account for the properties of the key entities (digital assets, the curation process, and stakeholders) as they manifest themselves in the context in question. Next, the implications of these properties for meeting the five sustainability conditions must be considered, and in particular the key uncertainties that arise in meeting those conditions must be identified and remedies sought to overcome or at least mitigate them. The result is a sustainability strategy that fits the circumstances of the particular digital curation activity in which it is expected to operate.

It is important to emphasize again that any sustainability strategy is necessarily an approximation, in the sense that it cannot guarantee long-term sustainability for the digital curation activity. Not all economic uncertainties can be identified or foreseen; for those that are identified, not all of them can be resolved or mitigated. Moreover, solving some problems may only be possible through trade-offs. For example, restricting access to paying users in order to strengthen incentives to contribute to a digital curation activity may also have the effect of diminishing the societal value proposition for curation, since the benefits from curation would be distributed over a smaller user base.

Planning a sustainability strategy is an ongoing activity. No sustainability strategy is likely to remain effective indefinitely without alteration. Planners must be alert to changing conditions in the economic environment, and be prepared to re-evaluate, adjust, or even completely re-design the sustainability strategy as needed. More generally, a sustainability strategy needs to be cognizant of and address the key economic uncertainties present in the context of a particular digital curation activity.

The notion of a sustainability strategy as a plan or roadmap for achieving economic sustainability is straightforward to grasp, but as the above description suggests, it is a complex space to navigate in practice. The following sections explain in more detail the various components of a robust sustainability strategy.

Takeaways:

- A sustainability strategy orchestrates economic factors to ensure a curation activity has sufficient resources to meet its long-term goals over the entire economic lifecycle.
- Achieving sustainability means meeting the BRTF's five sustainability conditions.
- To design a successful sustainability strategy, planners must understand the properties of the key entities; identify significant economic uncertainties associated with the properties; and identify appropriate remedies to address the uncertainties.
- No sustainability strategy is perfect; it can only maximize the prospects of achieving sustainability, not guarantee it.
- A sustainability strategy must evolve as conditions evolve.

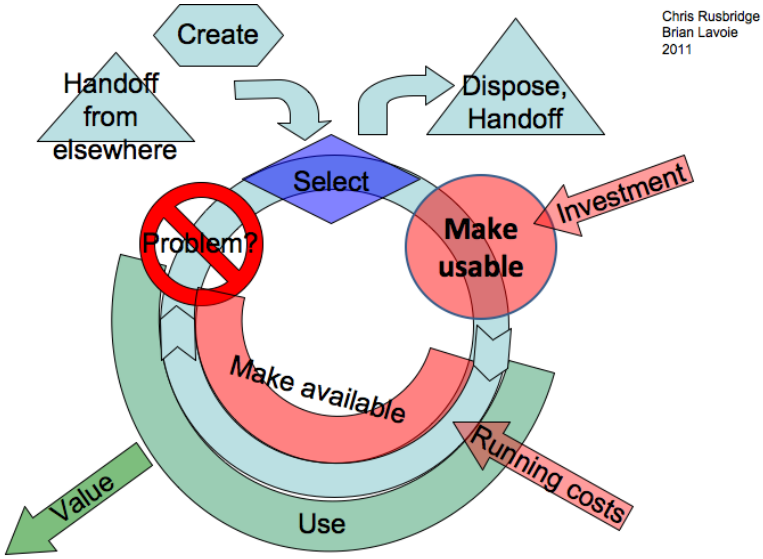
4.2.1 The Economic Lifecycle

A key point from the BRTF final report **[BRTF, p. 74-75]** is that economic decision-making for digital curation is a dynamic, sequential process unfolding over time, rather than a once-and-for-all event. Curation decisions, especially as they bear on sustainability, must be regularly revisited, re-evaluated, and if necessary, revised. The specifics of this process will necessarily vary from context to context, but it is

possible to construct an economic lifecycle that, at a sufficiently high level of abstraction, reflects the general cycle of events that characterize most if not all curation contexts.

Given the importance of the dynamic element of economic decision-making in regard to long-term digital curation, an economic lifecycle model is included as a key component of the reference model. An economic lifecycle model may be understood as one that highlights the major economic decision-points which typically occur over the "life" of digital assets, from creation to possible disposal. The lifecycle is both linear and cyclical in its construction, in the sense that it simultaneously reflects a natural progression through the normal lifetime of a digital asset, yet at the same time incorporates the idea that some economic decisions tend to "cycle" or repeat for a period of time during that lifetime.

Figure 3 – An Economic Lifecycle Model for Data Archives and Services



An economic lifecycle for data archives and services

The economic lifecycle unfolds as follows. Digital assets are created and some are selected and acquired for curation by the service or repository under consideration. The selected digital assets have to be prepared for use; this is the "ingest" phase in OAIS, the editing phase in publishing, etc. It includes adding relevant metadata required for use. This stage has been identified as one of the highest cost stages in archiving (Beagrie et al 2008[4]). Once usable, the digital objects have to be made available; running costs might continue indefinitely.

Available digital assets can be used (disseminated in OAIS terms). The assets create value through use (or availability for potential use). The economic case must be that the aggregated value over many resources and significant time exceeds the aggregated costs. One of the key ideas from the BRTF final report is "the case for preservation is the case for use" **[BRTF, p. 75-76]**. Use brings value, and value justifies preservation. It works the other way, too; if use is separated too far from the preserved content, then the value is reduced and the argument for preservation is diminished. This is not to say that there might not be other motivations for curation investment: e.g., a "bequest to future generations", or even a sense of "sentimental value" derived from knowledge that a digital asset continues to exist (see for instance Coyne 2007[5]). Moreover, there may be a need to preserve certain digital assets for litigation or evidentiary purposes, or to satisfy regulatory requirements or fulfil mandates. But demonstrable value from use is likely to make the strongest case for ongoing investment in digital curation.

As long as there is a reasonable perception of value, this situation can continue indefinitely. But eventually, it is likely that some sort of decision point or trigger event will arise. This could be a technical problem to do with the digital assets (e.g. obsolescence); it could be a technical problem to do with the service (e.g., infrastructure needs a significant upgrade); it could be a business problem to do with the service (bankruptcy, change of ownership or focus, etc). And it could also well be that the projected value of the curated digital assets failed to materialize, thereby calling into question the wisdom of continuing to allocate resources toward their curation. There is a question as to whether this resource has reached the end of its economic life. Decisions then must be made, as further significant investment may be required. Some (or all) digital assets will be retained, perhaps transformed to make them usable in the new environment. Some (or all) digital assets will be disposed of, de-accessioned, etc. The best outcome for these latter assets is that they are handed off to another service or archive; without that critical step, they will very soon be lost. (Similarly, assets might have arrived at the current service having been handed off from an earlier service.)

It is important to emphasize that the disruption of the regular cycle of ongoing availability and running costs can be anticipated – e.g., a regularly scheduled review of repository operations and resource allocations – or unanticipated – e.g., a sudden and unplanned reduction of funding. It is also the case that after addressing the disruption, anticipated or otherwise, the curation cycle may continue as before. The key point is that previous economic decisions are revisited and *possibly* altered. In some cases, alterations in previous decisions can be dramatic – for example, a decision to cease curating a particular collection of digital assets.

This general pattern of economic decision-making is characteristic of most long-term digital curation activities. Broadly speaking, this pattern includes two endpoints – the point where digital assets are ingested into the repository, and the point where they are removed from archival retention – with a cyclical period in between where the digital assets are made available on an ongoing basis, until some interruption to the cycle – a decision point or trigger event – occurs.

Takeaways:

- Economic decision-making for digital curation is a dynamic, sequential process unfolding over time, rather than a once-and-for-all event.
- Curation decisions, especially as they bear on sustainability, must be regularly revisited, re-evaluated, and if necessary, revised.
- An economic lifecycle model highlights the major economic decision-points which typically occur over the “life” of digital assets, from creation to possible disposal.
- The general pattern of economic decision-making for digital curation includes two endpoints – when digital assets are ingested, and when they are removed from archival retention – with a cyclical period in between where the digital assets are made available on an ongoing basis, until an interruption to the cycle occurs.
- Disruptions to the regular cycle of ongoing availability can be anticipated or unanticipated.

4.2.2 Sustainability Conditions

In discussing strategies for achieving economically sustainable digital curation, it is important to have a clear sense of what achieving economic sustainability actually means. The BRTF final report enumerates five conditions **[BRTF, p. 12]** that must be addressed in order to maximize the prospects for achieving economically sustainable digital preservation:

4.2.2.1 Value

This condition requires the identification of a clear stakeholder interest in the long-term curation of a particular set of digital assets. A digital curation activity is unlikely to attract funding – and thereby sustain itself – if no one sees value in the digital materials it is curating! But more than this, those that do have an interest in long-term curation need to be able to express that interest or value in compelling, specific ways. What kinds of valuable activities would be possible if the digital materials in question remained accessible over time? Or conversely, what kinds of valuable activities would *not* be possible if the materials disappeared? What might certain stakeholders be willing to pay to ensure the digital assets do not disappear? Since a sustainability case is particularly hard to build where the costs are real money but the returns are non-financial, sustainability can be greatly affected by perceptions of value.

4.2.2.2 Selection

One of the fundamental principles of economics is that resources are scarce; we usually cannot achieve everything that we might like with the resources available to us. For digital curation, this means that the slogan "curate everything for all time" is a non-starter. The resources available for curation will always be limited, and therefore we must prioritize and make choices: to the extent we can, we must proactively *select* digital assets for curation that are likely to promise the greatest value through use over time. And we should also keep in mind that the process of selection works in two directions. In particular, we must be prepared to "de-select" curated digital assets when the value from continuing to curate them no longer justifies the cost.

4.2.2.3 Incentives

It is one thing for a stakeholder to articulate an interest in having ongoing access to a set of digital materials (see "Value" above). It is quite another for a stakeholder to step forward and accept responsibility for curation. Sustainable digital curation requires stakeholders who not only recognize the value of curation, but who are also willing to sponsor or carry out the curation process. In short, there needs to be robust motivation or incentives to curate. Cultivating these incentives often means identifying and leveraging an institutional self-interest in curation: for example, curation as a business opportunity; curation as part of an institutional mission; curation as a means of fulfilling a mandate, and so on. Often, curation involves stakeholders with differing incentives to curate; these incentives need to be orchestrated over the full digital life cycle. For example, a media company may perceive a revenue incentive to curate a digital movie over a limited period while the asset has economic value. When this period expires, mechanisms need to be in place to transfer the asset to another institution with a different curation incentive, such as a library or archive.

4.2.2.4 Resources

No discussion of economic sustainability is complete without talking about resources. Curation activities, like any other activity, require sufficient resources to achieve long-term goals. Meeting this condition often boils down to developing mechanisms to transfer funding and other resources from those who benefit from and are willing to pay for digital curation, to those who are willing to provide curation services. There is a variety of market and non-market mechanisms for doing so, such as pricing models, compulsory fees or taxes, volunteer efforts, and philanthropic donations. Whatever mechanism(s) are chosen, it must support an ongoing flow of resources such that long-term curation goals can be achieved. But it is not enough to simply make resources available for curation. These resources should be used as efficiently as possible. Efficiency in this sense does not mean cutting corners, but rather getting the most value out of the resources allocated to curation. For example, we should strive to leverage economies of

scale by spreading costs over higher volumes of curation activity. We can also attempt to leverage economies of scope by spreading costs over different yet related services: e.g., locating curation and end-user access services on the same repository platform.

4.2.2.5 Organisation

Economic sustainability also requires that planners choose an appropriate organizational form for digital curation activities. A variety of organizational forms are possible. For example:

- An organization with no private interest in curation curates on behalf of others (e.g. a third-party curation service provider)
- An organization with a private interest in curation curates on behalf of itself and others (e.g. a research library)
- An organization with a public mandate to curate on behalf of society (e.g. a national archive).

To the degree there is discretion to choose, an organizational form for curation should be appropriate given the conditions prevailing in a particular context. In addition to an appropriate organizational form, a good governance mechanism is needed to ensure that curation goals are clearly articulated, a strategy is formulated for achieving these goals, curation responsibilities are appropriately allocated, and metrics and benchmarks are in place to evaluate outcomes.

Taken together, the sustainability conditions set forth in the BRTF report comprise a useful definition of what sustainability actually means in a digital curation context, and at the same time, enumerate the issues that an effective sustainability strategy needs to address. Failure to account for one or more of these conditions will seriously diminish the ability of a digital curation activity to sustain itself over time.

Takeaways:

- A clear understanding of what sustainability means in a digital curation context is essential in designing a sustainability strategy.
- The BRTF final report identifies five sustainability conditions.
- Value: identification of a clear stakeholder interest in the long-term curation of a particular set of digital assets.
- Selection: Resources are limited; we must prioritize and make choices. Where possible, digital assets should be selected for curation that promise the greatest value through use over time. Selection decisions should be re-visited and re-evaluated regularly.
- Incentives: Sustainable digital curation requires stakeholders who are willing to sponsor or carry out the curation process, based on robust motivations or incentives to curate.
- Resources: Curation activities require sufficient resources to achieve long-term goals. Mechanisms need to be established to transfer funding and other resources from those who benefit from and are willing to pay for digital curation, to those who are willing to curate. Once obtained, curation resources should be used as efficiently as possible.
- Organization/governance: An organizational form for curation should be chosen that is appropriate to the conditions under which it is expected to operate. A governance mechanism is also needed to articulate curation goals, formulate a strategy to achieve them, assign curation responsibilities, and evaluate outcomes.

4.2.3 Key Entities

The economic environment surrounding any digital curation activity is complex, involving a host of factors interacting to shape the circumstances in which a particular activity must operate. Understanding this environment, and its implications for achieving sustainability, can be challenging. Fortunately, it is possible to distil the economic environment into a few key elements that are present in all digital curation contexts. A thorough understanding of these elements and their properties is a necessary step toward building a successful sustainability strategy.

4.2.3.1 Digital Assets

Digital assets are the *raison d'être* of a digital curation activity. They can take a wide range of forms: research data sets, e-prints, executable software, web sites, and so on. To qualify as a digital asset, two criteria must be met: *first, the object must be digital, and second, it must be judged to have a value that will persist over some period of time*. The second criterion is crucial, and must be considered carefully by digital curators: is the digital object in question truly a digital asset? The question is not trivial: *not all digital objects are digital assets!*

Each class of digital asset exhibits a variety of properties that to a greater or lesser extent impact the nature of the curation activity itself. This can be readily seen from the perspective of the technical aspects of curation: the techniques and workflows needed to curate a collection of research data sets are likely to be different from those needed to curate a collection of executable software. In the same way, the properties of digital assets impact the strategies needed to support the curation activity from an economic standpoint. As in the technical sphere, different kinds of digital assets will have special properties that impact sustainability in unique ways. However, the BRTF final report **[BRTF, p. 25-28]** notes that digital assets also share certain core properties that must be taken into consideration when organizing a curation activity.

- ***Digital assets are durable yet depreciable:*** If kept in proper condition, digital assets can continue to release value for scholarship, private enterprise, education, and entertainment over extended periods of time. However, the clause "if kept in proper condition" is crucial. In order to maintain digital assets' value over time, those responsible for their stewardship will need to expend resources on an ongoing basis to support their curation. Digital assets are depreciable in the sense that if not properly maintained, they will tend to succumb to technological obsolescence, bit rot, or other "digital diseases", and in this way, their ability to release value to users will be reduced and eventually eliminated.

  The implication of this property for economic sustainability is clear: in assessing the costs of a proposed curation activity, planners must look beyond the current costs of gathering digital assets into a collection and consider something akin to "total cost of ownership". What will be the costs of maintaining the collection in a usable condition for an extended period of time? And what mechanisms can be put in place to support a sustained flow of resources to the curation activity? From the outset, curation planners must discard notions of one-time chunks of funding, and think instead in terms of ongoing flows of funding.

- ***Digital assets can be curated by one, but used by many:*** A digital asset residing on a public server is, at least in principle, accessible to any Web-enabled user. In this sense, the user base for a curated digital asset is potentially far larger than that of a physical item like a print book. Moreover, a digital asset can be used simultaneously by many users, also in contrast to a

physical item, where use by one generally precludes use by another (again, consider a print book). Expressed in economics terms, the use of digital materials is *non-rival in consumption*.

The important implication of this property for economic sustainability is that although the potential user community – those that benefit from curation – is potentially quite large, the incentive to contribute resources to the curation activity by any one user is weak. Once a digital asset is curated by one organization, it is at least in theory available for use by all; there is no need for each organization to curate a local copy of the digital asset, as long as a single curated copy is available on the Web. In contrast, there is a far stronger need to curate local copies of print materials, since a particular print copy can only serve a limited pool of primarily local users. Since one organization can curate a digital asset on behalf of all, it tends to be in the interest of any one organization to have someone else incur the trouble and expense of curation, while still enjoying free access to the digital assets.

Economists call this the *free-rider problem*, which can make it a challenge to collect sufficient resources among beneficiaries to sustain the curation activity. To overcome this problem, it may be necessary to exclude some beneficiaries from access to the digital assets if they do not contribute toward curation. If the planners for a curation activity are not prepared to exclude non-contributors, or for some reason it is impossible to do so, this may create significant problems for long-term sustainability; the incentives for contribution to the activity will likely be weak.

Takeaways:

- Digital assets are digital objects that are judged to have a value that will persist over some period of time.
- Digital assets share several core properties that impact sustainability planning.
- Because digital assets are durable but depreciable, planners must consider the "total cost of ownership" associated with maintaining them in a usable condition for an extended period of time. Rather than one-time chunks of funding, sustainable curation activities require flows of funding.
- A digital assets can be curated by one stakeholder, but used by many users simultaneously. Because of this, the incentive for any user to contribute resources to curation is weak. This *free-rider problem* can make it difficult to collect sufficient resources among beneficiaries to sustain curation. One solution is to exclude beneficiaries from access to the digital assets if they do not contribute toward curation.

4.2.3.2 Curation Process

Long-term accessibility to digital assets is achieved through the *curation process*, which is the set of activities involved in maintaining digital assets in a usable form for an extended period of time. It is important to distinguish between the *process* of digital curation, and the *outputs* of digital curation. In particular, the value of digital curation is experienced through the value-creating capacity of the curated digital assets themselves. Put another way, we generally do not value curation for the sake of curation; we value it because of the uses to which we can put curated digital assets. As the BRTF final report notes **[BRTF, p. 24-25]**, this establishes an important property of the curation process that is crucial when considering economic sustainability:

- The value of the curation process (and the associated digital information service in which it is embedded) is derived from the value-from-use of the curated digital assets. A curation activity

that cannot make a compelling case for the value of the curated digital assets that it manages will find it difficult to attract the funding and other resources needed to sustain itself over time. This ties in with the distinction noted above between digital objects and digital assets. A digital asset is a digital object with a perceived future value. The value of the curation process, therefore, derives from its ability to deliver the value of the digital asset. In economic terms the demand for the curation process is a derived demand. In other words, the demand for a curation service is derived from the demand for curated digital assets.

Many of the activities associated with the curation process are technical in nature, consisting of curation techniques and workflows invoked on the digital assets to ensure they persist in a suitable condition for use. For the purposes of economic analysis, however, it is not the technical aspects of the curation process that are key; instead, it is the *decision-making process* that overarches the day-to-day management of curated digital assets. This decision-making process also exhibits an important property relevant to economic sustainability **[BRTF, p. 28-30]**:

- Curation decision-making is path-dependent. Digital materials pass through a sequence of stages and decisions (the economic lifecycle; see above), with endpoints of creation and disposal. Decisions made at one stage of the lifecycle often shape the choices available to decision-makers at later stages. This point is illustrated in its starkest terms by the initial choice of whether or not to curate; if the decision is “no”, it is often impossible to revisit and change this decision at a future date, because by then the digital asset may be unavailable or have deteriorated to such an extent that effective curation becomes prohibitively expensive or even infeasible.

Takeaways:

- The curation process is the set of activities involved in maintaining digital assets in a usable form for an extended period of time.
- The value of the curation process derives from its ability to deliver the value of the digital asset. In economic terms the demand for the curation process is a derived demand – in this case, derived from the demand for curated digital assets.
- Curation decision-making is path-dependent. Decisions made at one stage of the economic lifecycle shape the choices available to decision-makers at later stages.

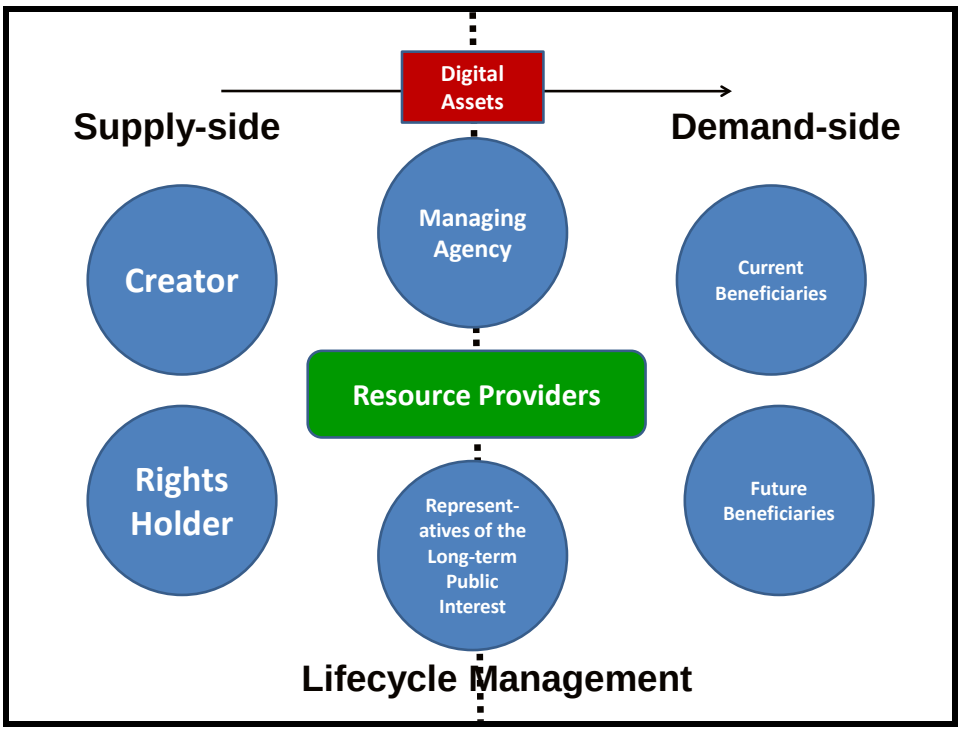
4.2.3.3 Stakeholders and the Stakeholder Ecosystem

The network of stakeholders surrounding a particular curation activity can be complex and difficult to characterize. Moreover, these stakeholders can represent an equally wide range of interests in regard to the long-term future of the assets in question. The organization of these stakeholders, and in particular, the distribution of curation roles across them, heavily impacts the prospects for achieving sustainability, and by extension, the shape of the sustainability strategy best suited for those circumstances.

In order to understand the basic contours of the stakeholder interests associated with a set of digital assets, it is useful to employ an organizing device called a stakeholder ecosystem. A stakeholder ecosystem designates the key stakeholder roles in lifecycle digital management, and articulates important relationships between them as they relate to sustainability. The stakeholder ecosystem takes the form of a high-level abstraction that helps organize stakeholders into broad categories of interest. Figure 4 provides a view of the types of stakeholders associated with a digital curation activity. It is important to recognize that the stakeholders illustrated in the figure are *roles*, and not necessarily distinct entities. A single organization (or individual) can fulfil multiple roles simultaneously. As we will see, this can have

important implications in terms of creating conditions that either encourage or discourage long-term sustainability.

Figure 4 - Types of Stakeholders



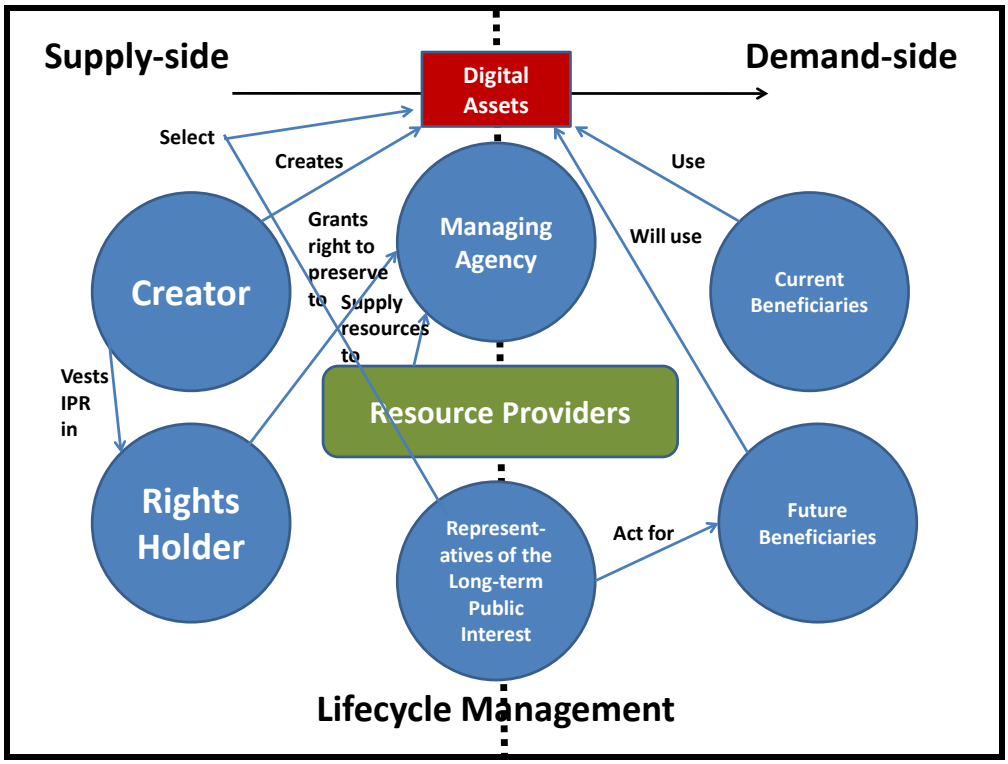
The stakeholder taxonomy depicted in Figure 4 is divided into three areas:

- **Supply-side:** stakeholder roles pertaining to those who create and/or own the digital asset.
  - Creators: entities responsible for creating digital assets (e.g., scientist who creates a data set; artist who creates a digital work of art; company that develops a new computer game)
  - Rights Holders: entities who currently own key intellectual property rights related to digital assets, such as the right to provide access, the right to preserve, etc. This will often be the same entity that created the digital asset (i.e. the same stakeholder has both the Creator and Rights Holder roles), but not always. For example, a publishing company may hold the rights to an e-journal article authored by someone else; a social media site may own the collective contributions of its members, etc.
- **Demand-side:** stakeholder roles pertaining to those who benefit from the availability of the digital asset.
  - Current beneficiaries: those who currently have well-defined uses for the digital asset, and derive value from its ongoing availability.
  - Future beneficiaries: those who could be expected to benefit from the digital asset in the future, or those whose uses of the digital asset are as yet unknown/undefined. The BRTF report notes that the interests of future beneficiaries are often left out of decision-making on digital assets.
- **Lifecycle management:** stakeholder roles pertaining to those responsible for curating the digital asset, and for ensuring its ongoing inclusion in the cultural/scholarly record.

- Managing Agencies: entities responsible for managing, curating, preserving or providing access to digital assets.
- Representatives of the Long-Term Public Interest: entities who advocate for the digital asset's inclusion as part of society's ongoing cultural or scholarly record. This reference model focuses on digital assets in which there is a long-term public interest. These are materials that would be included in society's cultural and scholarly record. Typically cultural heritage institutions (libraries, archives, museums) play a prominent role in advocating or actively intervening to ensure certain materials are in fact curated on behalf of the long-term public interest. Various philanthropic organizations, funding agencies and informal volunteer networks would fall into this category as well.
- Resource providers: entities who provide resources (i.e., funds, in-kind transfers, etc) to support the curation process. Provision of resources can occur in a variety of ways: payment of a fee-for-services (e.g. where the Managing Agency is a third-party digital archiving service and the Resource Provider is a customer of this service); donation (e.g. where the Resource Provider is a corporate sponsor of a digital collection or of a curation organization); grant award; advertising (e.g. where the Resource Provider pays the Managing Agency for the placement of advertisements in association with the latter's access services); and so on.

As Figure 4 illustrates, many different categories of stakeholders are associated with a digital curation activity. But this in and of itself does not create an ecosystem. The stakeholder ecosystem is created by the pattern of relationships between the stakeholder roles that holds for a particular activity. This pattern of relationships can be quite complex, as Figure 5 illustrates.

Figure 5 - Stakeholder Relationships: Example



Relationships in the ecosystem define how the various stakeholder categories interact with one another. For example, as the illustration in Figure 5 shows, in many digital curation contexts Representatives of the

Public Interest such as libraries and archives act on behalf of Future Beneficiaries by making sure that the latter's interests are represented in today's curation decision-making. Similarly, the Rights Holder and Managing Agency often have a relationship in that the former must grant an exclusive or non-exclusive right to curate to the latter. Resource Providers also have a relationship with the Managing Agency, in that they supply resources sufficient for the Managing Agency to carry out its curation activities.

The pattern of relationships depicted in Figure 5 is but one of many possibilities – many other relationships between the various stakeholders could be imagined. However, it suggests the potential complexity that might prevail in the pattern of relationships associated with a particular digital curation activity. It also points up the importance of orchestrating the interaction of *all* stakeholders in a sustainable digital curation activity. A sustainability strategy that neglects to account for the interests and participation of any of the stakeholder categories, and hence their roles and responsibilities in the curation activity, is likely to founder at some point in the economic lifecycle **[BRTF, p. 22-23]**.

It is important to emphasize again that all of the stakeholders represented in the ecosystem are *roles*, and not necessarily distinct entities. So a single person or organization could simultaneously fill multiple roles. Many of the important issues regarding incentives to preserve and allocation of long-term curation responsibilities hinge on whether various roles are combined within a single organization, or separated across distinct entities. The more that stakeholder roles are dispersed across distinct entities, the more negotiation and consensus must be achieved to reach sustainability goals, and the greater the probability that individual interests will conflict. This leads to an important property of stakeholders that has enormous implications for the prospects of achieving long-term economic sustainability:

- The distribution of curation roles across the network of stakeholders is *critical for understanding both the factors encouraging sustainability in a particular digital curation context, as well as the uncertainties that might prevent sustainability from being achieved.*

The ecosystem diagram illustrates the major categories of stakeholder interest in curated digital assets. The distribution of stakeholder roles across distinct individuals or organizations is a key factor in influencing the nature of the economic uncertainties associated with a given curation activity, and the kinds of economic remedies that are appropriate for addressing them. For example:

- The incentives to curate are strengthened when the Rights Holder for a particular digital asset is also a Current or Future Beneficiary. They are weakened when this is not the case.
- Allocation of responsibility for undertaking curation is clearer when the Rights Holder is the same as the Managing Agency. When this is not the case, it is less clear who should undertake the curation process, especially when ownership (the Rights Holder role) is split among many entities.
- The interests of Future Beneficiaries are often not represented in the curation decision-making process. Do Representatives of the Long-Term Public Interest adequately represent these interests and are they being taken into account in today's curation decision-making?
- When Resource Providers are distinct from Current and Future Beneficiaries, it is vital that a compelling value proposition exists to incentivize the Resource provider to supply resources for long-term curation.

In summary, the stakeholder ecosystem is derived from the layering of three components on top of one another: first, the *set of stakeholder categories* depicted in Figure 4; second, the *pattern of relationships* between stakeholders depicted in Figure 5; and third, the *distribution of stakeholder roles* discussed above. The contours of the stakeholder ecosystem associated with a given digital curation activity are fundamental in determining the potential "traps" that might impede the activity from achieving long-term

economic sustainability. By extension, then, the design of an effective sustainability strategy to mitigate these traps must rest on a thorough understanding of the underlying stakeholder ecosystem.

The impact of the stakeholder ecosystem on the economic uncertainties faced by a digital curation activity can be illustrated with an example. ArXiv is a repository for e-prints in physics, mathematics, and other sciences. Access to the contents of arXiv is free and open to all. Over the years, arXiv has become a key mode of scholarly communication between scientists, who rely on it both as a means of exposing their work to colleagues around the world and eliciting comment and discussion. Cornell University currently serves as the Managing Agency for arXiv, and prior to 2010, was the sole Resource Provider as well. Although faculty and students at Cornell certainly benefit from use of arXiv, the vast majority of those who obtain value from arXiv – that is, the Current Beneficiaries – are unaffiliated with Cornell University.

A key feature of arXiv's stakeholder ecosystem is that the Resource Provider is distinct from Current Beneficiaries. Those who benefit from using arXiv are not obligated to contribute toward its costs, and indeed have a weak incentive to do so, since they are not excluded from the benefits even if they contribute nothing. This creates a free-rider problem, since those who benefit from arXiv are free to do so without contributing resources to its ongoing operations. This situation was sustainable only as long as Cornell University was willing to fund arXiv's approximately $400,000 annual budget on behalf of the general scientific community. This in turn introduces an economic uncertainty into the curation activity: the Current Beneficiaries have little incentive to provide resources to support a curation activity undertaken in their interest, which calls into question the ability of the curation activity to marshal sufficient resources on an ongoing basis to continue its operations. One way to mitigate this uncertainty is to induce or persuade Beneficiaries to contribute to arXiv – that is, to combine the roles of Resource Provider and Beneficiary in the same entities.

In 2010, Cornell made its first step toward doing just that, by issuing a call to other institutions whose affiliated faculty and students benefited from arXiv to make voluntary donations to support arXiv's operations, with the requested donations scaled to each institution's level of use (as measured by number of downloads). In describing this new funding strategy, Cornell noted[6] that "[s]cholars worldwide depend upon the stable operation and continued development of arXiv. Sustainability is best assured by aligning revenue sources with the constituents that realize value from arXiv, and by reducing dependence upon Cornell University Library's budget. We have decided to pursue a collaborative business model that will engage the institutions that benefit from arXiv". In short, Cornell has realized the potential pitfalls of a stakeholder ecosystem where beneficiaries are completely distinct from Resource Providers. The donation program is a first step toward re-aligning the ecosystem such that the roles of Resource Provider and Beneficiary become merged within the same entities. It is significant to note, however, that Cornell considers the donation program to be only an interim solution until a long-term business plan can be worked out. An obvious weakness of the interim plan is that it is voluntary, with no "repercussions" if institutions choose not to participate. A long-term strategy might consider restricting arXiv access only to those institutions that contribute toward its upkeep. However, given its long history as a freely available resource, such a strategy is bound to encounter resistance.

This example demonstrates that knowledge of the stakeholder ecosystem underlying a particular digital curation activity helps identify potential economic uncertainties impacting long-term sustainability, as well as possible remedies to overcome them.

Takeaways:

- The wide range of stakeholder interests associated with many digital curation activities, and the relationships between them, heavily impacts the prospects for achieving sustainability, as well as the shape of the sustainability strategy best suited for the circumstances.
- A stakeholder eco-system designates the key stakeholder roles in long-term digital curation, and articulates important relationships between them as they relate to sustainability.
- Stakeholders fall into three general categories: Supply-side (Creators; Rights Holders); Demand-side (Current Beneficiaries; Future Beneficiaries); Lifecycle management (Managing Agencies; Representatives of the Long-Term Public Interest; Resource Providers). A sustainability strategy that neglects to account for the interests and participation of any of the stakeholder categories is likely to founder at some point in the economic lifecycle.
- The stakeholder ecosystem can be understood as three layers of abstraction that, taken together, illustrate the way stakeholders are organized within a particular digital curation context, and how that organization might harbor economic uncertainties impacting sustainability. The first layer is the types of stakeholder roles associated with most digital curation activities; the second layer describes the relationships existing between stakeholder roles in a given digital curation context; the third layer describes the distribution of stakeholder roles across distinct individuals or organizations.
- It is important to recognize the distinction between stakeholders as distinct entities and stakeholder roles. A single organization (or individual) can fulfil multiple stakeholder roles simultaneously.
- The distribution of curation roles across the network of stakeholders is critical for identifying economic uncertainties inherent in a particular configuration of the stakeholder ecosystem.

4.2.4 Economic Uncertainties

Understanding the nature of the economic lifecycle underpinning digital curation, the conditions which must be met to achieve sustainability, and the key entities composing a digital curation activity, is an essential step toward developing a robust sustainability strategy. However, even a sustainability strategy based on this strong foundation is not necessarily destined for success. Myriad economic uncertainties exist throughout the economic lifecycle that can potentially derail even the most well-designed sustainability strategy **[BRTF, p. 35-47]**. While it is impossible to completely eliminate the potential effects of these uncertainties, anticipating as many of them as possible, and incorporating appropriate plans and contingencies to mitigate their impact, is an essential part of any sustainability strategy.

It is wrong to think of uncertainty as always being bad: consequences can be positive or negative. Economists refer to positive (yet uncertain) outcomes as upside risk, and negative (yet uncertain) outcomes as downside risk. We take on an upside risk when we invest in something (whether with money or personal time or effort), which is exactly what someone has to do for sustainable archives or information services!

Economic uncertainties potentially impact a repository's ability to achieve sustainability. These will vary from case to case, and it is therefore important to assess and address uncertainties for the specific circumstances associated with a particular digital curation context. This in turn requires a thorough understanding of the economic lifecycle, the sustainability conditions, and the key entities – digital assets, the curation process, and stakeholders – constituting the environment in which the digital curation

activity will operate. This is the starting point for identifying the most significant uncertainties associated with a curation activity.

Uncertainties could affect many different aspects of a digital curation activity. Assuming some sort of service for the management or curation of digital assets over time, uncertainties can affect:

- The creation of the service
- The continuation of the service
- The termination of the service
- The succession or transformation from one incarnation of the service to another, whether in succession or in parallel, e.g. through the handoff process (including legal and other agreements), through migration of the assets, identifiers etc, through revised technology for the underlying service, or new owners or policies
- The quality of the service
- All or some digital assets in the service.

This list is not exhaustive!

The economic lifecycle implies uncertainty at nearly any point, but explicitly at the stage labelled "Decision point". Economic uncertainty will most likely come into play when significant changes are expected, for example to the cost structure (where cost may include effort, time etc as well as money), the group of interested stakeholders, or to the underlying resources. As such they are more likely to apply to the whole service rather than part of it.

All of the five sustainability conditions described above are subject to economic uncertainty. For example, the perceived **value** of a set of archived digital assets can change over time, significantly altering the cost/benefit equation underpinning their curation; **selection** (and de-selection) policies can become mis-aligned with stakeholder expectations; the **incentives** to curate can diminish as the stakeholder ecosystem evolves and changes; insufficient start-up **resources** may not be available to get the curation activity off the ground; a change in mission or even closure of the program responsible for a curation activity leads to **organizational** issues concerning curation roles and responsibilities.

A complete list of potential digital curation uncertainties would almost seem endless, but a well-designed sustainability strategy will include provisions to address at least the uncertainties that are most significant given the circumstances of the digital curation activity it is intended to support. It is sensible to do a thorough assessment when deciding which uncertainties to explicitly address in the sustainability plan, and how to do so.

Broadly speaking, there are four ways to respond to a perceived uncertainty:

- *Avoidance:* stop the activity with the uncertainty.
- *Mitigation:* lower probability and/or impact of uncertain event.
- *Transfer:* pass uncertainty to another party (e.g., insurance).
- *Acceptance:* recognize the uncertainty but choose to "ignore" it.

4.2.4.1 Avoidance

Some organizations are extremely risk-averse, and are unwilling to expose themselves to any significant risks. In this sense, they are likely to stop, or not initiate in the first place, activities that are deemed "risky" in one way or another. For example, orphan works are works whose rights-owners cannot be traced even after reasonable effort. Legally, making these available or even carrying out some

preservation actions could be viewed as copyright violations, so this can be an economic uncertainty. Avoidance in this situation would be implementing a strict policy of curating only resources whose rights-ownership status clearly allows this. Uncertainty could be reduced even further by curating only public domain works, and/or making the repository "dark" (i.e., eliminating access to archived materials). As this example makes clear, a significant trade-off usually resides in avoidance strategies, in the form of a reduction in capacity and/or benefits, in return for avoiding uncertainty.

4.2.4.2 Mitigation

Once the more serious economic uncertainties are identified, and a decision has been taken that the uncertainties cannot be avoided, then where possible steps should be taken to reduce or mitigate those uncertainties. Two approaches are available, preferably to be used together but if necessary in isolation. These are to reduce the likelihood of the uncertain event occurring, and to reduce the effect or impact if it does occur. The precise means of doing so will depend, of course, on the specific circumstances involved, but some examples help illustrate the point:

- Uncertainty: loss of trust by user community; Mitigation: improve curation procedures and/or undergo quality assurance or certification process (such as the Data Seal of Approval[7]).
- Uncertainty: doubt that archived digital assets will have future value; Mitigation: carry out a strategy where initial curation is relatively light-weight and low-cost, yet is sufficient to preserve the option to engage in more intensive curation later on if future demand materializes.
- Uncertainty: right to curate will be disputed or withdrawn; Mitigation: seek appropriate legal advice prior to curation actions; establish proper licenses and disclaimers; respond promptly to concerns raised about curation/access rights.
- Uncertainty: flow of funds to curation activity may be interrupted; Mitigation: develop multiple funding streams, which will reduce the impact on sustainability if one stream is interrupted.

Note that in all of these examples, uncertainty is not fully eliminated, either in terms of probability or impact. However, the mitigation strategy helps reduce the likelihood and impact of the uncertainty, which in turn helps adjust the perceived cost/benefit equation of the uncertain activity into a more favourable ratio. This can make all the difference in deciding to go ahead with a particular curation activity, as opposed to adopting an avoidance strategy instead.

4.2.4.3 Transfer

A transfer strategy can be effected in several ways. One is through something akin to an insurance policy, where in exchange for premium payments, the insured risk is transferred to the issuer of the insurance policy. The insurance policy issuer, of course, aggregates many risks and relies on the probability that most will not in fact occur. There are insurance policies to cover many types of risks; unfortunately, we are not aware of any specifically tailored to those associated with digital curation!

Another important form of transfer is to simply transfer a risky or uncertain activity to another organization. For example, the arrangements that publishers have with Portico[8] or CLOCKSS[9] are essentially transfers, because they are transferring the responsibility (and therefore the attendant uncertainties) of digital curation to these organizations. Why should other organizations be willing to accept such transfers – i.e., take on risks to which other organizations are unwilling to expose themselves? One reason may be differences in mission: while one organization may not see long-term curation as part of its core mission (or business model), and therefore is reluctant to take on the uncertainties attendant with curation activities, another organization may perceive curation to be an essential component of its mission, and the associated uncertainties simply a "cost of doing business".

4.2.4.4 Acceptance

Acceptance is simply recognizing that a particular uncertainty exists, but taking no particular action to avoid, mitigate, or transfer it. For example, a repository could decide to curate and make available orphan works without explicit permission, without making an effort to locate the rights holder or design and manage a take-down policy. The uncertainty, of course, is that the owner of a work may eventually materialize and seek legal redress. As this example suggests, the advantage of an acceptance strategy is that it incurs no extra costs involved in mitigating or transferring uncertainty, or in reduced capacity or benefits in the case of avoidance. The disadvantage is that if the uncertainty manifests itself, the repository is fully exposed to its impact.

It would seem that there are few situations where it wouldn't be sensible to take some form of mitigation or transfer action in the face of a perceived uncertainty – at least those which are non-trivial in their impact, or whose likelihood of occurring is more than just "within the realm of possibility". However, the option of choosing an acceptance strategy reminds us that avoidance, mitigation, and transfer have costs, which need to be weighed against their perceived benefits. The digital curation community would benefit from the emergence of clear best practices emerged in regard to which economic uncertainties need to be addressed in a sustainability strategy, and which approach – avoidance, mitigation, or transfer – is the preferred method for addressing each of them.

Takeaways:

- Myriad economic uncertainties exist throughout the economic lifecycle. While it is impossible to eliminate these uncertainties, anticipating as many of them as possible, and incorporating appropriate plans and contingencies to mitigate their impact, is an essential part of any sustainability strategy.
- There are four general ways to respond to a perceived uncertainty: *avoidance* (stop the activity with the uncertainty); *mitigation* (lower probability of uncertainty occurring, or reduce its impact if it does occur); *transfer* (pass uncertainty to another party (e.g., insurance)); *acceptance* (recognize the uncertainty but choose to "ignore" it).
- Risk-averse organizations may be unwilling to expose themselves to any significant risks. But a significant trade-off usually resides in avoidance strategies, in the form of a reduction in capacity and/or benefits, in return for avoiding uncertainty.
- Mitigation reduces the likelihood of an uncertain event occurring, reduces the effect or impact if it does occur, or both. Mitigation does not fully eliminate the probability or impact of the uncertainty, but reduces them sufficiently to make incurring the uncertainty acceptable to repository planners.
- Transfer can be effected through something like an insurance policy, or by transferring a risky activity to another organization. Other organizations may be willing to accept such transfers because the risky activity is central to their perceived mission.
- Acceptance involves recognizing that an uncertainty exists, but taking no particular action to avoid, mitigate, or transfer it. The advantage of acceptance is that it incurs no extra costs involved in avoiding, mitigating or transferring uncertainty; the disadvantage is that if the uncertainty manifests itself, the repository is fully exposed to its impact.
- Avoidance, mitigation, and transfer have costs, which need to be weighed against their perceived benefits. Best practices are needed as to which economic uncertainties should be addressed in a sustainability strategy, and which approach is the preferred method for addressing each of them.

5 The ESRM Role in the 4C Project

The 4C project was conceived to look across a complex area of work, assess the different perspectives and approaches, and then to synthesise them and represent them more meaningfully to new and existing stakeholders. This aligns well both with the purpose of an EC coordination action and with the needs of the broad community who have witnessed many attempts over twenty or so years to either describe, model or quantify how much digital curation costs; or how organisations might best be configured or strategically persuaded to invest in their digital assets. Lists of the various relevant initiatives (quoting sources as far back as 1980) have been compiled[10] and there has been commentary urging the community to stop duplicating effort and exploit existing work more effectively.[11]

So the 4C project is working in the context of a community that is beginning to voice dissatisfaction (or perhaps impatience) with attempts to tackle the cost of curation. But what is also apparent from these bibliographies and commentaries is that although the topic is clearly understood to be multi-faceted and complex, it would appear (based on the non-hierarchical and uncategorised way that the lists are compiled and the references made) that it is convenient and logical for those wishing to understand the connections and boundaries of the topic to categorise it all as broadly to do with taking an economic perspective on curation.

The following is a non-exhaustive and conceptually overlapping list of the types of issues that are addressed by the literature[12]:

- Contingent valuation
- Ecosystem valuation
- Intangible assets
- Investments
- Information markets
- Sustainability
- Incentives
- Cost modelling
- Lifecycle modelling
- Loss costing
- Business modelling
- Cost-driven design
- Cost/benefit analysis
- Net financial benefit
- Net economic benefit
- Net social benefit
- Shadow pricing

Given the expectation emerging from the community that any new work on the costs of curation should be cognisant of the much more extensive work referred to above, the Economic Sustainability Reference Model is a highly useful method of concisely capturing a whole range of concepts and approaches that have a bearing on the economics of curation and provides the 4C project with a framework and a jumping off point for the work outlined in Work Package 4 (Enhancement).

5.1 The ESRM in relation to WP4 (Enhancement)

The ESRM work is being done in parallel with another task within WP4 which seeks to identify issues which are likely to have an indirect determining effect on the cost of curation. The general objectives of WP4 as stated in the project Description of Work are as follows:

> "The objective of this group is to ensure that comprehensive consideration is given to all indirect factors that might still be economic determinants of digital curation, initially through:
>
> - a collation and prioritised assessment of those determinants
> - an expansion on partners recent experience in audit and certification with a focus on trust and quality; both key economic determinants of current priority within the community
>
> These will inform initial work on an Economic Sustainability Reference Model which will be iteratively updated in response to further Assessment [WP2] and Engagement [WP3] work and to a subsequent evaluation of risk, benefit, impact and value as determinants."
>
> (*Taken from the 4C Project Description of Work, (hereafter referred to as 4C DoW)*)

From the outset then, it was anticipated that a number of factors would need to be identified and assessed as having a bearing on the economics of curation, particularly issues of trust, quality, risk, benefit, impact and value. The work that has subsequently been done to generate the list of indirect economic determinants will be available on the 4C project website at http://4cproject.eu/outputs-and-deliverables/ as deliverable 4.1. The draft list available at the time of writing includes the following items:

- authenticity
- benefit
- efficiency
- impact
- innovation
- interoperability
- quality
- reputation
- risk
- sensitivity
- skills
- sustainability
- transparency
- trustworthiness
- value

This initial list (and its accompanying definitions) will form the basis of an ongoing community consultation to quantify and prioritise their relative importance amongst the diversity of ways that organisations and groups manage their information and plan the curation of their digital assets. It is clear, for instance, that a large corporate entity engaged in the manufacture of aviation components will have a very different perspective than, say, a university research department. But what should be possible is to arrive at a list that makes overall sense across a wide variety of domains and becomes an initial checklist of concepts against which curation planning can take place.

As well as planning, however, the point of the list is to indicate that if any of these concepts are important, then their degree of importance will influence the amount of care (and ultimately the investment) that will be required to ensure that they are adequately addressed by the entity that is considering its ongoing curation requirements; which is where the indirect economic determinants overlap and feed into the economic sustainability reference model. The ESRM is (amongst other things) an attempt to build many concepts into a workable and understandable framework that organisations and groups can use to try and make sense of what is otherwise a blizzard of terms and concepts, all of which may be said to have economic significance.

As stated above, two other tasks within WP4 (T4.3 and T4.4) will more closely examine the economic significance between curation and some particular conceptual areas that were identified at the outset of the project as being potentially significant. It is anticipated that the list of indirect economic determinants and the ESRM will enable the 4C project to engage with various stakeholder communities and elicit views providing input and shape for the two reports addressing issues of: trust, risk, impact, benefits and value.

The final task of WP4 (T4.5) will examine business models for digital curation with a particular focus on the provision of services and building on the knowledge and analysis of cost and economic issues emerging from earlier phases of the project. The ESRM will, along with various other components frame this task and provide it with stakeholder views and positions from the outset. It is evident that there is logically not much conceptual distance between a sustainability model and a business model, so it follows that a model that aspires to provide generic and widely applicable guidance about economic sustainability should have much to offer a series of business models aimed at defining service options for digital curation.

5.2 The ESRM in relation to WP2 (Engagement)

The task of Work Package 2 (Engagement) is to "identify, get involved and build partnerships with individuals, groups and institutions that are active or interested in the issue of curation costs, and foster a better understanding of the issue amongst the community." *(4C DoW)*

So this activity is designed to setup interaction between the project and interested stakeholders both for the purpose of getting their input (to help refine and develop the project outputs) and as the audience for those outputs when they have been developed. One of the ways in which that interaction can happen is by using the early outputs of the project to foster discussion and elicit opinion and this is one of the principal early purposes of the ESRM in relation to the 4C work plan.

It is anticipated that the ultimate purpose of the ESRM will be to act as a community-owned reference model that will assist organisations to define and plan sustainable strategies for digital curation (somewhat akin to the role that the OAIS Model[13] has played within the preservation community for the last decade). By this logic, if it receives endorsement and validation from the community and can be maintained sustainably, the framework of the model (set out in sections 3 and 4 of this document) will hopefully continue on from the 4C project as a discreet resource. At this early stage however, the context in which the model is developed within the 4C project, and the way that it is introduced and explained to stakeholders will be an important part of getting it positively embedded into the community.

One of the early tasks of WP2 (Engagement) was to define a list of stakeholder groups that were of relevance to the project and would be targets for engagement and dissemination. The following table sets out the various groups that were alighted upon.

Table 1 – 4C Stakeholder Matrix

Type	Difficulty	Impact	Priority
Research funders	Low	**High**	9
Big data science	Low	**High**	9
Cost model experts	Medium	**High**	6
Digital preservation vendors	Low	Medium	6
Universities	Low	Medium	6
Government agencies	Medium	Medium	4
Publishers & content producers	Medium	Medium	4
Industry	**High**	**High**	3
Memory institutions and content holders	Low	Low	3
Small and medium enterprises	High	High	3
Miscellaneous contacts / Other interested parties	-	-	-

It should be apparent that the diversity of stakeholders that 4C aims to reach requires it to have a number of different strategies for engaging with people at all levels and in very different working contexts. In order to generalise across all of these categories and provide 4C with a paradigmatic way of characterising stakeholders, the groupings suggested by the APARSEN project[14] have been adopted with the addition of a category denoting funders and investors (see figure 6).

Figure 6 – High level descriptions of stakeholder roles/perspectives

It is possible that these high-level descriptions of stakeholder roles will not be applicable in some working contexts (e.g. small commercial enterprises) and may describe more than one discreet role that in reality devolve to one individual. However, they are meant to indicate role perspectives rather than jobs *per se*

and should be interpreted loosely. The rationale for the layout will become apparent in the following section.

5.3 The ESRM in Relation to WP3 (Assessment)

The purpose of Work Package 3 (Assessment) is as follows:

> "To establish the most effective current methods for private and public sector organisations to estimate and compare the cost of digital curation, and to identify the most beneficial paths for future development of solutions and services. This will enable stakeholders to more effectively and comprehensively assess the investment of resources that may be required to sustain their digital preservation activities; and allow comparisons of existing and future tools and models with the knowledge that a broad range of criteria: e.g. price, savings, quality, value, risks, benefits, sustainability, etc., are implicit to the comparison." *(4C DoW)*

Whilst acknowledging the implicit importance of what it refers to as a 'broad range of criteria', the main stated focus of WP3 is: to look closely at the issue of costs; the ways they have been measured before now; the shortcomings and gaps in the methods that have been proposed; and to suggest more effective approaches for the future. So ostensibly, the work is very much engaged with the sort of activity (cost modelling) that the Blue Ribbon Task Force (out of which the ESRM originated) had to frequently and emphatically differentiate itself from. The BRTF was a conscious attempt to look more holistically at the issue of sustainability through an economic lens and the ESRM is a continuation of that purpose. It maintains the principle that focusing on costs without also looking at a range of other factors, in particular benefits and incentives, is of limited value. So what is the relationship between the ESRM and a work package that is principally focused on cost models?

The main difference between the work now and the work that happened during the period 2008-2010 (during which time several prominent cost models were being developed or refined; and the BRTF was convened) *is the existence and publication of the BRTF reports themselves* - along with the subsequent dissemination of those results via symposia and related workshops (see section 2: background). Irrespective of whether the conclusions and recommendations of the BRTF work have been a strong influence on organisations or strategic bodies (and they may well have been; it is not the remit of the 4C project to make such an assessment), it is clear from the frequency with which the BRTF is mentioned by the digital preservation community that it has been broadly influential.

So the 4C Project is very much a 'post-BRTF' community collaboration and as such the workpackage focused on the assessment of cost models - and where to go next with them – fully acknowledges the fundamental importance of considering benefits alongside costs and is committed to ensuring that such thinking is embedded into the 4C approach. Early discussions of the 4C project team have spent some time focused on this issue. The following series of figures is a graphical summary of some ideas that have helped the 4C team to work out where components sit relative to each other; how they relate to the stakeholder grouping introduced in figure 6 above; the aspects stakeholders are interested in; how those interests map onto models; and how the models relate to each other.

Figure 7 - What are different stakeholders focused on and interested in?

Figure 7 reuses the stakeholder roles mentioned above and introduces a dashed line to gently demarcate territory to the left and right.

It places Operational Staff (i.e. practitioners with a hands-on digital curation role) on the left-hand side of the diagram indicating more of a focus on the actual management of digital assets and a knowledge of curation as a detailed series of actions and processes that result in well managed, robust and accessible data.

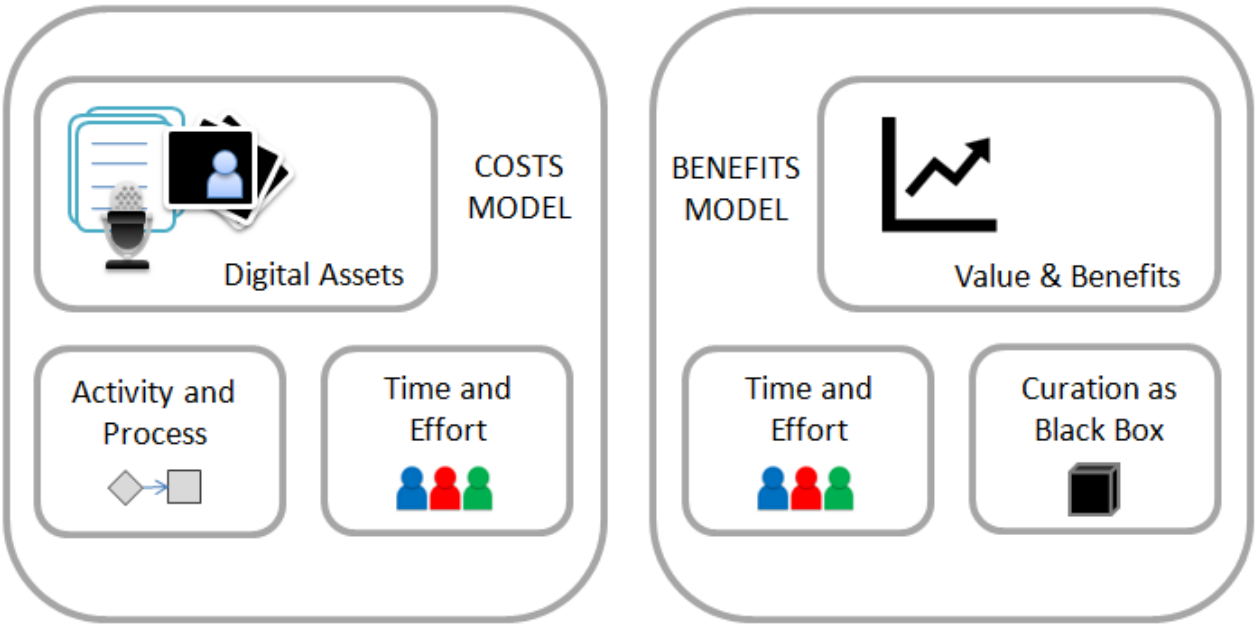
The Operational Manager (i.e. the person with responsibility for overseeing those responsible for digital curation or for delivery of a curation service) is likely to share an interest in the practical aspects of curating data effectively but may also be interested in the right hand side of the diagram which focuses on the value of assets and the benefits of curation; and the amount of time and effort it takes to realise those benefits.

The Senior Manager (i.e. a decision-maker or budget holder who may have responsibility for resourcing digital curation but no topic expertise) may also straddle the central line in as far as they take an interest in the effective management of digital assets but they are less likely to be interested in the detailed processes involved with curation. They will be interested in the value and the benefits of curation and they will want to be able to measure that against the time and effort (cost) involved. They are more likely to regard curation as a black box (i.e. a system which delivers an outcome without the detailed workings of that system needing to be apparent to the beneficiary).

The Funder or Investor (i.e. the entity making resources available for digital curation to occur) is even less likely than the Senior Manager to concern itself with the detail of curation processes and is highly likely to be very interested in the resultant benefits of activity and how that translates into impact and positive outcomes.

The way that these break down into areas for modelling follows the left and right hand sides of the diagram taking 'time and effort' as a common component of both – see figure 8.

Figure 8 – Costs and Benefits Models with Variant Stakeholder Interest Assumptions



This is undoubtedly a simplistic rendering of the implicit assumptions of modelling approaches but is intended to convey the need for costs models to deliver inputs to benefits models that are relatively unencumbered with the kind of detail that will be hard to digest and difficult to make judgements against. Conversely, it is important for benefits models to be as clear as possible about the value it expects and the level of benefits that it is reasonable to anticipate (and possibly to realise) in order to make the cost/benefit analysis across the models meaningful. Given the complexity of both sides of this modelling for digital assets, there is a role for an overarching economic analysis that should help to frame all of this work and rationalise the decisions that are required.

Figure 9 – Economic model as overarching guide and principle for other modelling activity

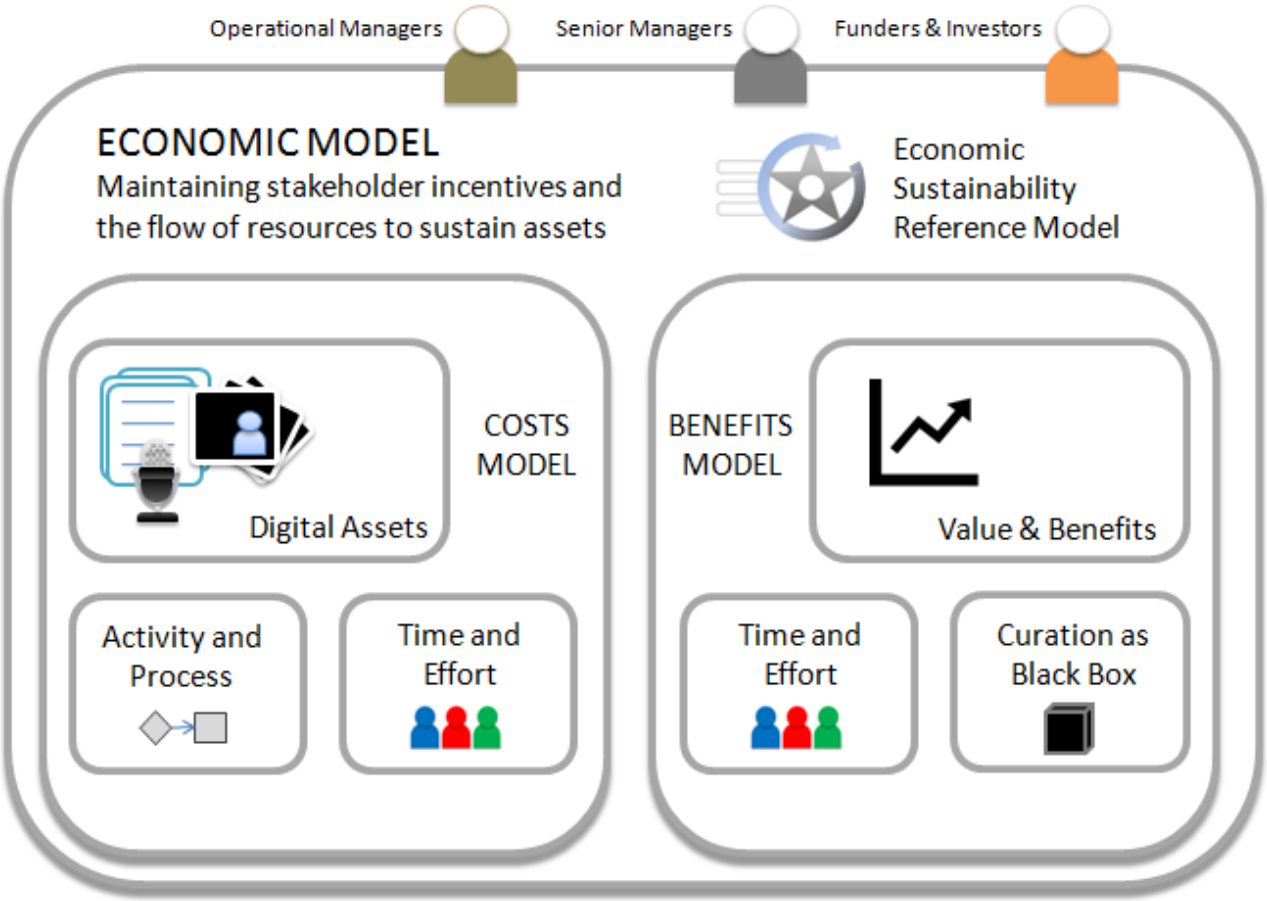


Figure 9 above nests the costs and benefits modelling activities within the overarching framework of an economic model; and indicates that this form of analysis is primarily strategic in nature and therefore of most interest to the managerial and funder/investor level stakeholders. It also sets out the place of the ESRM amongst all of the modelling and assessment activity that originates out of WP3.

5.4 The ESRM in Relation to WP5 (Roadmap)

The object of WP5 (Roadmap) is "to arrive at coherent and evidence-based recommendations for future action and strategy in relation to the economic aspects of digital curation. The focus will be on measures that will assist diverse types of organisations to better understand and take control of the cost of managing digital assets over varied timescales, including the provision of cost-effective solutions and services to others." *(4C DoW)*

The production of a Roadmap report is a commonplace output for an EC-funded project of this type and by design it should gather together intelligence from all parts of the project. The ESRM is a tool for thinking about the component parts of an economic model and (like the BRTF report before it) it comes with some built-in opinions. It advocates regarding access to and use of the digital asset as the motivating factor behind digital curation (or in BRTF language – digital preservation is a 'derived demand'). It advocates sustainability as the guiding underlying economic principle that should inform decisions (rather than say 'affordability'). It steers in the direction of public entities taking ultimate responsibility at the end of chains of stewardship when materials are handed on between organisational entities that are no longer capable or interested in sustaining the assets (rather than anticipating responsibility for long-term preservation and access should be a market-driven commercial enterprise)... and so forth. One of the purposes of incorporating the ESRM work into the 4C project is to use this platform of opinion and perspective to find out whether and how consensus can be reached across different stakeholder

communities about economic strategies for digital curation. The discussions and reactions to the implicit structures, agendas and advice and guidance contained within the ESRM will be a key input into the Roadmap report.

6 Implementation and Tools

The ESRM in its current iteration sets out a number of concepts, many of which originate from the work done by the Blue Ribbon Task Force on Sustainable Digital Preservation and Access. The model as elaborated and fleshed out in section 4 is designed to be read as a standalone resource (perhaps with some accompanying introductory text) and in future will almost certainly be presented discreetly from the rest of the contents of this report and the work of the 4C project. However, the intention of the 4C project is to attempt to see how the ESRM work can be applied and implemented so the following section (and appendix 1) is an examination of *one possible implementation* of the ESRM according to the requirements and purposes of the 4C Project. As will be made clear in section 7 (Next Steps) there is an expectation that community input will continue to shape and hone the model but there is, as yet, no certainty that either the graphical representations shown above (in section 5) or below (or the tables in appendix 1) will become intrinsic or attached components of the model as it is taken forward for validation (or critique) by the broader community.

6.1 The 4C Candidate ESRM Graphic

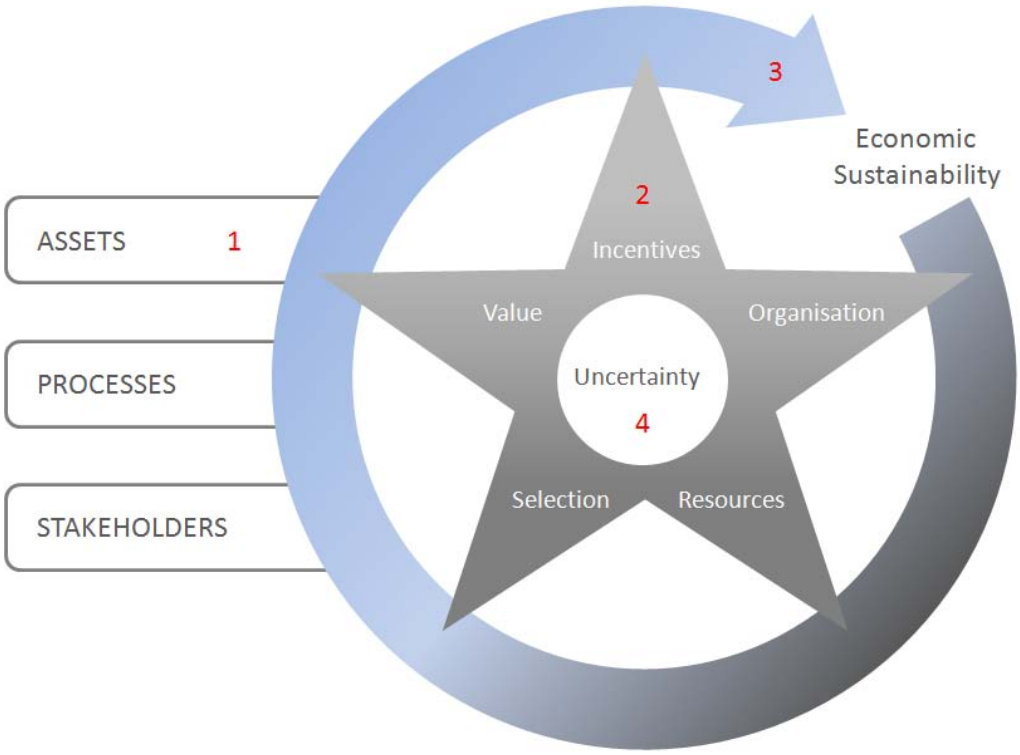
A concise candidate representation of the ESRM is shown in figure 10 below and it has two primary purposes. Firstly, it is emblematic of the full description of the model (which is mainly a text-based resource) and is designed to be a concise visual reference (a logo) for the model and useable at a range of different scales, with or without the text annotations (see figure 11).

Figure 10 – Scaled usage of the proposed ESRM graphic representation

Secondly, the larger annotated version sets out the four main components of the ESRM and is intended to act as an aide-memoire for components of the reference model. The four elements are as follows and their position is highlighted by red numerals in Figure 11 :

1. The 3 Key Entities (see section 4)
2. The 5 Sustainability Conditions (see section 4)
3. The Economic Lifecycle (see section 4)
4. The Concept of Uncertainty (see section 4)

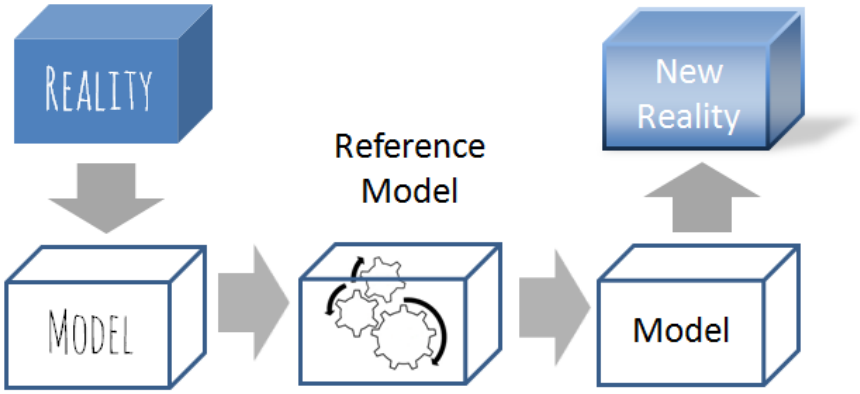
Figure 11 – The layout of the components of the ESRM



6.2 Reference Model as Change Tool

Early discussions within the 4C project questioned the purpose of a reference model and what extra meaning it conveys over and above the simple term 'model'. One possible definition of the term 'model' is that it is *a simplified version of the real world.** Taking this definition, the purpose of a reference model is to allow a local model to be compared with a community validated and generic yet comprehensively thought-through elaboration of the concept. This may provoke ideas for changing the local model to address shortcomings or imprecision, which in turn should suggest ways of changing the reality that the model represents.

Figure 12 – Reference model as change tool



* Included in the 4C Project Glossary and paraphrased from http://www.businessdictionary.com/definition/model.html. This definition of 'model' and slight linguistic variants of it are widely referred to in relation to economic modelling.

6.3 Working with the Model

A detailed elaboration of the latest draft of the model is contained in sections 3 and 4 of this document. This provides a rationale for its use; sets out the components and an order for working through them; defines many of the terms; provides links to earlier sources where appropriate (particularly where the text is related to the BRTF work); and provides summaries in the form of 'takeaways' at the end of each sub-section.

The additional content of this document, particularly section 5, aims to explain how working on the ESRM model and thinking about the nature of economic sustainability has helped the 4C project to clarify its own thinking about the relationship between costs, economics, benefits, etc. What this represents then is the first occasion when the ESRM has been put to good use to help another project to think through issues in relation to the economics of curation. That is not to say that the 4C Project has not also helped to shape the latest iteration of the ESRM but it also anticipates that community use of the model will be the way that the ESRM evolves into a resource that will ultimately be of broad utility to a wide range of people.

The point of this section is, therefore, to emphasise that the objective of the remaining period of the 4C project is to find out how people react to the model and in what ways they may wish to engage with it and work with it in the context of their own organisation. Some reasons why the ESRM might be useful have been suggested in section 3 (education, consensus-building, standardization, and benchmarking) but it is down to the community to really work out whether and how this model helps them to achieve things that are important to them. If it does, it would be very useful if they could pass that information back to the 4C Project along with the specifics of how exactly it did help them with their strategic or practical activities. Further details about how that interaction may be possible will be outlined in section 7 (Next Steps).

Whilst the aim is *not* to be prescriptive about how to work with the model, the 4C project does offer an opportunity to interpret the model and extract a series of assertions from it that can be turned into a relatively concise exercise which may be of use to managers with overall responsibility for orchestrating digital curation activity. This interpretation of the model is set out in Appendix 1 and will be available in other formats on the Community Resources page of the 4C website (http://www.4cproject.eu/community-resources).

It is the first practical example of how the model can be used to generate other resources and tools and is intended to provide encouragement to others to think of alternative ways of engaging with it.

7 Next Steps

An early objective of the 4C Project in relation to the ESRM is to make a new version of the model available for community scrutiny so the first initial step beyond the finishing of this report is to present the model and the ancillary materials at the forthcoming IPRES conference due to be held in Lisbon on September 2nd – 6th 2013.

Should the model not find traction with the community it may still have a role to play as a 4C engagement mechanism (see section 7.2). However, it is hoped that there will be an audience and a demand for it and 4C has designed some opportunities into the project workplan to facilitate further discussion.

7.1 The ESRM as a community developed resource

From the outset, 4C has been identified as a good vehicle for enabling further development of the ESRM and perhaps more importantly, as a method of increasing community engagement with the model in order to improve it, make it more fit-for-purpose, and to orchestrate the sort of broad community buy-in that will be needed for it to become a true 'reference model'.

The 4C Project intends to provide the following ways of enabling further broad discussion on this topic (alongside other issues arising from the project).

- 3 workshops (September 2013; July 2014; October 2014)
- 4 Focus Groups (September 2013; December 2013; March 2014; July 2014)
- The 4C Conference (October 2014)

In addition to these formal engagement opportunities, the 4C Project website will provide a mechanism for engaging the community in debate about the issues. The website includes the facility for the community to start conversations or leave comments. http://wwww.4Cproject.eu

The Community Resources page (http://www.4cproject.eu/community-resources) will list all of the outputs and deliverables (including this one) and it is possible to leave comments on the individual page relevant to the resource.

The 4C twitter account is another way of fostering debate (@4C_project) and tweets that include that twitter name or include the 4C Project hash tag #4Ceu will be monitored by the project team.

Other opportunities will also be sought to elicit views on the ESRM and other aspects of the project work by giving presentations and papers at external workshops, symposia and conferences. If it transpires that the model is used by the community and this comes to the attention of the 4C project team, then case studies of that usage will be written up and made available on the 4C website, in addition to being included in the final ESRM deliverable which is scheduled for February 2015.

It is also possible that copies of the ESRM will be available from other locations. OCLC and Jisc have separately invested time and money into developing the model and it would be logical for copies of the ESRM to reside on the websites or in the repositories of those organisations. Those copies in turn may attract their own commentaries. The 4C Project is scheduled to draw up a sustainability plan in the course of the project and the ESRM is one of the resources that needs to be considered as part of this. Discussions will be had to determine the most appropriate long-term stewardship of the model and the coordination of its further development once the 4C Project concludes.

7.2 The ESRM as an engagement mechanism

If it does transpire that the ESRM is not favourably received or simply doesn't elicit much interest from the community then there may be a good case for continuing to refer to the ESRM anyway (for the duration of the project at least) as a form of 'straw man'. By drawing attention to the ESRM and suggesting that stakeholders adopt its strategic recommendations, this might then provoke those stakeholders to better articulate in what ways the model is insufficiently descriptive of the context in which they work. It is already known that different communities have different perspectives about the economics of digital curation and 4C has a wide remit to engage across those communities. It is perfectly possible that different communities may not be aligned with (or not prepared for) the sort of conceptual structure or assumptions that underpin the ESRM (see section 5.4 above).

Small and medium sized businesses and similarly sized public-sector bodies especially may not fully grasp all the terminology or embrace the implicit assumptions of the ESRM but they should at least be able to form some opinions about the following high level assertions – which arise out of the key entities discussed in **section 4.2.3**.

- The properties of digital objects must be realistically examined and understood before they can be defined as assets;
- Sustaining digital assets (the process) is complicated but not impossible to organise and understand if you take the view that it is about ensuring there is sufficient 'return on investment';
- The assets and the processes are not themselves intrinsically important: the crucial issue is whether the value of the assets is sufficiently understood by the relevant stakeholders

Unlike discussion of cost models and cost modelling techniques which may be beyond the capacity or capability of some organisations, the assertions above can be used as a basis for discussion and engagement with almost any organisation. Additional next steps will consequently involve less visible use of the ESRM on the part of the 4C Project as it reaches out to individuals and uses its extended network of contacts to try and find additional affiliate partners and active stakeholders to help it progress the work in hand.

Anyone who is interested in talking to the 4C project about the ESRM or any aspect of the work on the costs and economics of digital curation should contact the project via the website (http://www.4Cproject.eu) or via email at info@4cproject.eu.

References

The references appear after Appendix 1 (see page 63)

Appendix 1

The following self-assessment exercise represents one potential implementation of the ESRM. The questions in the tables are based on a reading of the BRTF reports and an interpretation of the ESRM and may be of use to senior managers and decision-makers when considering the costs and benefits of digital curation and how sustainable it is likely to be and over what timescale.

This interpretation of the Economic Sustainability Reference Model (ESRM) and its implementation as a series of questions has been developed by the 4C project in conjunction with Brian Lavoie (OCLC).

The following tables are also available in .docx and pdf/a formats

see http://www.4cproject.eu/community-resources

4C Sustainability Self-Assessment Questionnaire (ESRM)

Introduction

The following tables interpret the advice and guidance in the Economic Sustainability Reference Model (ESRM) and present it as a sustainability self-assessment exercise.

Objective

To challenge present assumptions about sustaining designated digital assets

Who is this aimed at?

The questions are non-technical and are aimed at respondents who have strategic-level responsibility for digital curation and the sustainability of digital assets.

Instructions

The issue level throughout the exercise should be rated as follows:

1. Not an issue
2. A minor issue
3. An issue
4. A major issue
5. Not resolvable

There are 40 questions in total.

The exercise begins by recording who undertook the exercise and when they did it so that the perspective of the participant can be factored in to any subsequent analysis of the form.

Name	
Affiliation	
Department / Group / Section	
Role	
Date	
Which digital assets* are the specific focus of this sustainability assessment exercise?	

* Digital assets is the term used throughout this exercise to refer to all digital information/materials/ objects/data that will require digital curation over time to ensure their sustainability

CHALLENGES

ASSETS

Digital assets have a variety of properties and characteristics that have implications for digital curation and sustainability.

#	Questions	Responses	Issue Level
1	Are the digital assets in manageable formats?		
2	Can digital curators access them and work on them?		
3	Are they homogeneous in type and capable of being batch processed or are they heterogeneous, chaotic and difficult to process?		
4	Is it known whether others have copies of the assets and if so, whether they are sustaining them effectively?		
5	Are the rights to the assets understood and do they allow digital curation to be undertaken?		

	# STAKEHOLDERS The network (or ecosystem) of stakeholders that have interests in digital assets is often complex and may affect digital curation and sustainability strategies, particularly where curation roles are distributed.		
#	Questions	Responses	Issue Level
6	Has the stakeholder ecosystem been surveyed or mapped?		
7	Are the curation roles in the ecosystem clear to everyone?		
8	Has at least one stakeholder declared responsibility for the assets?		
9	Is it understood what interests various stakeholders have in the assets?		
10	Is it clear whose job it is to supply access to the assets?		
11	Is it clear who is demanding access to the assets?		
12	Is there a ‘free rider’ problem with the assets, i.e. is anyone getting access to them that isn’t paying for them and is this an issue?		

	# PROCESSES Long-term accessibility to digital assets is achieved through the curation process, which is the set of activities involved in maintaining digital assets in a usable form for an extended period of time.		
#	Questions	Responses	Issue Level
13	Are the curation processes that are currently used to sustain the assets adding sufficient value?		
14	Is it understood that the activities that are undertaken at an early stage of the digital asset lifecycle determine the later value of those assets?		
15	Is there enough flexibility in the curation processes to withstand a change in external economic conditions? (i.e. can processes be changed or scaled up or down)		
16	Can the curation processes change in response to emerging technology?		

STRATEGY & TACTICS

The following 21 questions are divided up into five short sections and address strategic or tactical issues that may need to be overcome for sustainable curation of digital assets to occur.

The issue level should be rated as before

1. Not an issue
2. A minor issue
3. An issue
4. A major issue
5. Not resolvable

The 'when might this change?' column is to record at what point the current situation is likely to change. For example, there may be an agreement in place for the next 3 years to transfer funds to support digital curation activities but there may be little likelihood of the funding arrangements being renewed. So a response to this question might be as follows:

EXAMPLE

#	Questions	Responses	Current Issue Level	When Might This Change?	New Issue Level
20	Are functioning agreements in place to transfer funds and/or other resources from those who are willing to pay to those who are able and willing to curate?	*We have an agreement with the Foundation for Digital Sustainability that their final phase of funding to support our digital curation programme will end in three years time. We have no alternative funding currently lined up.*	*1*	*2016*	*4*

If the situation or the response to the question is unlikely to change, then the box should be left blank.

Value

The term ‘digital asset’ implies that the object in question has value. This value has to be understood and realised over time to avoid the digital object becoming a liability.

#	Questions	Responses	Current Issue Level	When Might This Change?	New Issue Level
17	Is there a written business case that clearly sets out the purpose of sustaining the assets?				
18	Have the assets been assigned quantitative financial value?				
19	Has the value of the assets (financial and/or non financial) been agreed upon by the stakeholders for whom they are being sustained?				

Resources

Are sufficient resources in place to achieve long-term goals?

#	Questions	Responses	Current Issue Level	When Might This Change?	New Issue Level
20	Are functioning agreements in place to transfer funds and/or other resources from those who are willing to pay to those who are able and willing to curate?				
21	Is the mechanism for payment or transfer of other resources arranged such that it is effective, robust and reliable? (e.g. pricing, fees, philanthropic donations, effort)				
22	Are curation activities organised and managed to make the best and most efficient use of the funds or resources received?				
23	Can sufficient human resource with adequate capability be employed to ensure effective curation takes place?				

24	Are the appropriate tools and infrastructure available to allow effective curation to take place?				

	Selection Are strategic choices being made about the assets being sustained?				
#	Questions	Responses	Current Issue Level	When Might This Change?	New Issue Level
25	Is it understood that the best use of resources involves making choices based on value judgements and selecting material for curation?				
26	Is the selection of materials based on an agreed framework where the selected items are those which are most likely to realise the greatest value over time?				
27	Is the demand for materials sufficiently understood to inform selection decisions (e.g. types of demand: measured, certain, non-excessive)				
28	Is it possible to de-select and transfer or dispose of assets that are deemed no longer worth sustaining?				

Organisation

If an organisation chooses to sustain its digital assets, it must know why it is making that investment and why it is motivated to do so. Alternative organisations might have more plausible reasons to take on the responsibility for sustaining the assets.

#	Questions	Responses	Current Issue Level	When Might This Change?	New Issue Level
29	Is the organisation under an obligation to sustain the assets? (e.g. funder requirements, project conditions, customer expectations)				
30	Are the governance arrangements clear for taking responsibility and approving sustainability strategies?				
31	Is there a written policy that mandates a way of sustaining the assets?				
32	Is there a written strategy that sets out how the assets will be sustained?				
33	Are roles and responsibilities for sustaining the assets clearly allocated?				
34	Are roles and responsibilities for sustaining the assets appropriately allocated?				

35	Have benchmarks and metrics been assigned to evaluate the outcome of curation activity and its impact on sustainability?				

	Incentives Sustaining digital assets requires stakeholders to invest in curation processes and investments require incentives.				
#	Questions	Responses	Current Issue Level	When Might This Change?	New Issue Level
36	Do the incentives for sustaining the assets align with the interests of the rights holders?				
37	Are there sufficient incentives to continue to invest in sustaining the assets?				
38	Have the costs of sustaining the assets been calculated?				
39	Have the benefits of sustaining the assets been clearly articulated?				
40	Are the benefits that the assets realise more significant than the cost of sustaining them?				

REFERENCES

[1] BRTF-SDPA. (2010). *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*. (A. Smith Rumsey, Ed.). San Diego. Retrieved from http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

[2] BRTF-SDPA. (2010). *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*. (A. Smith Rumsey, Ed.). San Diego. Retrieved from http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

[3] Wikipedia authors. (2011). Business plan. *Wikipedia, the free encyclopedia*. Retrieved October 5, 2011, from http://en.wikipedia.org/wiki/Business_plan

[4] Beagrie, N., Chruszcz, J., & Lavoie, B. F. (2008). *Keeping Research Data Safe: A cost model and guidance for UK Universities*. Bristol. Retrieved from http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pd f

[5] Coyne, M. (2007). *The Case for Investment in Digital Archives*. Retrieved from http://digitalpreservation.ssl.co.uk/asset_arena/text/-7/D12-7.pdf

[6] Cornell University Library. (2010). arXiv Business Model White Paper. Retrieved October 7, 2011, from http://arxiv.org/help/support/whitepaper

[7] DANS. (2008). Data Seal of Approval. The Hague: Data Archiving and Networked Services. Retrieved from http://www.datasealofapproval.org/

[8] Fenton, E. (2006). Portico: An Electronic Archiving Service. *iPres 2006*. Retrieved from http://ecommons.cornell.edu/bitstream/1813/3684/1/Eileen_Fenton-Portico.pdf

[9] CLOCKSS. (2008). A Trusted Community Governed Archive. *CLOCKSS Overviews*. Retrieved from http://www.clockss.org/clocksswiki/files/CLOCKSS_Handout.pdf

[10] Open Planets Foundation, Digital Preservation and Data Curation Costing and Cost Modelling, http://wiki.opf-labs.org/display/CDP/Home (accessed 21/07/2013)

and Library of Congress, A Digital Asset Sustainability and Preservation Cost Bibliography, http://blogs.loc.gov/digitalpreservation/2012/06/a-digital-asset-sustainability-and-preservation-cost-bibliography/ (accessed 21/07/2013)

[11] Wheatley, P. (2012), Digital Preservation Cost Modelling: Where did it all go wrong?, http://www.openplanetsfoundation.org/blogs/2012-06-29-digital-preservation-cost-modelling-where-did-it-all-go-wrong (accessed 21/07/2013)

[12] This list is taken from a presentation given at PASIG, Dublin 2011: Grindley, N. (2011), Costs Versus Benefits, http://www.slideshare.net/neilgrindley/digital-preservation-costs-versus-benepasig-dublin-oct-2012-dp-costs-final2 (accessed 21/07/2013)

[13] Open Archival Information Systems, developed by the Consultative Committee for Space Data Systems, ISO 14721:2003, (revised in 2012), http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=24683

[14] The APARSEN Project is an EC-funded FP7 Network of Excellence, http://www.alliancepermanentaccess.org/index.php/aparsen/