

DRAFT RECOMMENDATIONS FOR DRYAD'S SUSTAINABILITY PLANNING

5 April 2010

Prepared by:

Charles Beagrie Limited

www.beagrie.com

Contents

1. Executive summary and recommendations	5
1.1. Background.....	5
1.2. Key points and recommendations.....	5
2. Strategy, performance indicators and measures	10
2.1. The vision.....	10
2.2. The organization	10
2.3. Summary of strategic goals and key performance indicators	10
3. Comparators and understanding the costs	12
3.1. Data sharing for small science	12
3.2. External comparators - journals and publishers.....	12
3.3. External comparators - other archive repositories	15
3.4. External comparators: institutional support and embedding	16
3.5. External comparators: use of students and volunteers.....	17
3.6. External comparators: activity models for archives.....	17
3.7. Cost models for Dryad	20
4. Advantages, benefits and revenue.....	25
4.1. Competitive advantages and unique selling points.....	25
4.2. Core competencies and assets.....	25
4.3. Summary benefits of Dryad	26
4.4. Prioritizing revenue streams	27
5. The proposal for sustainability	30
5.1. Considerations going forward	30
5.2. Organizational model to be adopted	30
5.3. Governance and partner commitment	31
5.4. Proposed sustainability model.....	31

5.5.	Further internationalization of Dryad.....	33
5.6.	Narrowing the gap between cost and revenue	34
5.7.	Plans for transition and implementation	34
6.	Scenarios for Dryad	36
6.1.	Introduction	36
6.2	Assumptions and caveats	37
6.3.	Scenario One – per paper charges dominant	37
6.4.	Scenario Two – annual subscriptions dominant.....	39
6.5.	Scenario Three – per paper charges and subscriptions balanced.....	40
6.6.	Differential pricing	41
6.7.	Bands or tiers for subscriptions.....	41
6.8.	Distinguishing types of member.....	42
6.9.	Founding member incentives.....	42
6.10.	Other incentives	42
7.	Risks	44
	Appendix 1: Additional Tables.....	45
	Endnotes	47

PREFACE

Dryad is an initiative led by NESCent (The National Evolutionary Synthesis Center) and the University of North Carolina at Chapel Hill Metadata Research Center, who are collaborating with a consortium of journals to develop and sustain a digital repository for publication-related data. The Dryad initiative seeks to establish a shared repository that can support a joint data sharing policy among consortium members in which any and all data underlying a publication must be archived. The initial efforts toward a shared archiving policy and repository grew out of a core of journals in evolutionary biology and ecology.

Dryad's startup funds come primarily from a four year NSF grant awarded in 2008, as well as other some other smaller sources (NESCent, the NSF-funded DataONE initiative, and a grant from the Institute for Museum and Library Services).

The NSF grant identifies as a key goal to establish stakeholder ownership and governance of Dryad, where journals serve as key stakeholders. The governing board has oversight over plans for adoption, sustainability and utility of the repository. One of the major tasks of the governing board is to agree to a sustainability plan that will ensure that Dryad is not dependent upon ephemeral grant funding, and to help to implement the sustainability plan.

1. EXECUTIVE SUMMARY AND RECOMMENDATIONS

1.1. BACKGROUND

Charles Beagrie Limited has been commissioned to work with the Dryad Project team on developing a sustainability plan. This work began on 20 October 2009 and will complete by the end of April 2010. We have incorporated within the plan results from an independent consultancy on cost models for Dryad by Lorraine Eakin and work by the Dryad Project team.

The aim of this document is to set a framework to develop the plan for future sustainability. It is intended to be a dynamic document that can be maintained, reviewed at least annually, and maybe more frequently over the first 2 years, evolving over the life of the project and beyond. It provides guidance on sustainability with the aim of informing business planning.

1.2. KEY POINTS AND RECOMMENDATIONS

Section 2 Strategy, performance indicators and measures

Dryad is two years into a four year development. It has established a clear process model from the outset to provide a repository for data associated with published articles, working hand-in-hand with journals, societies, and publishers. Dryad's goals are expressed in section 2 of this document. They are based on the project aims and objectives with appropriate modifications. The Board's acceptance of these goals is a precondition for developing a business model (the basis for a business plan.)

Recommendation 1: The Board adopts the draft strategy performance indicators and measures proposed in section 2.

Section 3: Comparators and understanding the costs

We found through desk research and interviews with journals, publishers and data centers that little is known by journals about the specific costs of handling supplementary materials and data. Costs, principally staff time, were observed to vary according to the tasks undertaken and value added. It is currently not possible to compare costs with those for Dryad as they are largely unknown, or for tasks such as adding metadata etc which are not currently undertaken. However it could be observed that proposed charges for Dryad (although it excludes some tasks and adds others compared to the journals) appear very reasonable compared to existing author charges (where these exist) for publishing supplementary data files amongst the journals we interviewed. These ranged from \$100 to \$300+.

We also found that while there is no exact archive or repository comparator for Dryad, other archive repositories do offer enough similarities to be of use in comparing some overall costs. Initial analysis, with feedback from the Dryad management team, indicates that a staff of 2-4 FTEs would be a viable initial base level of staffing to deliver Dryad's basic

operations. This is comparable to the minimum staffing of other archive comparators at launch we have considered.

The comparators we have looked are embedded within larger institutions and can thus leverage pre-existing infrastructure and effort, expertise and direction from associated staff, often co-located but funded separately working on related activity, including support services, project based research and development. This helps to maintain a dynamic and sustainable organisation that can respond to change and deal with fluctuations in staffing. Dryad currently is similarly embedded within a larger institution. Dryad needs to consider what will be the most cost effective way of providing its administration and infrastructure support going forward and whether support from a host institution can be negotiated at a cost or provided as an “in-kind” contribution.

It is very early in the Dryad development to accurately populate an activity model to derive full costs for its future activities. Initial projections of potential costs for Dryad are included in section 3.7 and section 6 to scope the likely level of costs at 5,000 papers per annum and at 10,000 papers per annum (ballpark figures of around \$200,000 or \$320,000 are indicated). We have also reviewed use of students as part of Dryad’s staffing. This provides some cost savings which are outlined as an option. These figures can be further refined and developed over the coming months.

Costs for curation will vary according to the level of additional work e.g. metadata enhancement, packaging and documentation for re-use in teaching undertaken by Dryad. Dryad is working on the development of a set of “curation levels” and associated costs. This is similar to practice in some publishers e.g. Journal of the American Medical Association or data archives such as the UK Data Archive.

Recommendation 2: There will be a minimum level of core staffing needed for Dryad to ensure sustainability and ongoing service. Relevant issues include cover for absences or staff turnover, viable staff recruitment, and a range of skill sets.

Recommendation 3: Promotion of the Dryad service will be critical in early phases and beyond. A staffing that includes a full-time executive director is likely to be required for successful implementation of the plan.

Recommendation 4: The proposed staffing level and required job descriptions and costs should be reviewed by the Dryad Project team in conjunction with the current host institution (Duke University).

Recommendation 5: Until Dryad reaches a point in its maturity and development where it can effectively be transitioned into a separate not-for-profit organisation it should work with the host institution to ensure continued support.

Recommendation 6: Dryad could pilot and review use of students as part of its staffing for curation.

Recommendation 7: Dryad should adopt an activity based costing model and build up knowledge of its activities, their costs and cost variables (volumes, file formats, etc).

Recommendation 8: Dryad should further develop the application of curation levels as a variable that can define and influence selected costs.

Section 4: Advantages, benefits and revenue

A summary table of advantages and benefits from Dryad is set out in section 4.3.

Repositories and established data services address their sustainability with regard to funding by establishing and seeking out multiple revenue streams. It is important to focus upon those options where there is most potential benefit of contributing to Dryad in order to maximize potential success.

Potential revenue sources are discussed in sections 4.4 and 5.4. We believe that Dryad should look to a mixed funding model and identify any tasks that can be funded from other sources than its core income. In particular it should seek project funding for innovative research and development wherever possible. The table in section 4.4 suggests the primary revenue streams are likely to be: per paper fees and partner subscriptions from journals; and grants from other funders. Journals have variable local circumstances and would need to decide individually how they would cover fees and subscriptions e.g. via author charges and/or from societies/publishers. Scenarios for different balances in these primary revenue sources are outlined in section 6.

Recommendation 9: Dryad must agree the options for revenue creation and should seek diversity as the best route to sustainable funding. It should pursue grant funding for innovative research and development wherever possible.

Section 5: The proposal for sustainability

The transition from Dryad's development phase to Dryad as a repository service requires careful planning and development of a transition strategy. The main considerations will be around organization and governance; staffing levels; maturity and reliability of automated processes to sustain the repository; and the level of active marketing and participation to build a critical mass of data available through Dryad.

The views of funders on future or continued grant support for Dryad will need to be investigated further. The views of NSF will be particularly important but European and other international funding bodies could be considered. Three areas of grant could be important: the possibility of tapered "transition funding" to facilitate the transition from project to service and to allow for the growth of the service in its early years; internationalisation of the

service (e.g. mirroring or nodes in Europe or elsewhere) to provide opportunities for widening participation and funding of the service; and research and development project funding opportunities.

There are currently 16 interim partner journals in the Dryad Consortium. The board will need to consider and agree the potential future growth or optimum size of the consortium, appropriate timescales for this, and impacts on revenues/costs as part of the transition strategy.

Recommendation 10: We recommend Dryad discusses future governance structures with its stakeholders. Possible elements might include an elected Management Board (with ex-officio representatives); an Annual Participants' Meeting; and an Advisory Board.

Recommendation 11: The views of funders on future or continued grant support for Dryad will need to be investigated further and future funding options and opportunities regularly reviewed by the board.

Recommendation 12: A separate transition planning phase and a strategy for transition funding should be developed by the end of 2010 in order to facilitate implementation of a transition to full service in 2011/2012.

Section 6: Scenarios for Dryad

Three scenarios have been developed to frame discussion of potential revenue options that would provide financial sustainability. Each scenario focuses upon different weightings of the three methods identified in the table in Section 4.4 that have been assigned the highest priority because they align most closely with stakeholders and have the most potential to raise funds. These are: journal usage based upon a per-paper deposited price; annual subscriptions paid by journal; and a one-off journal joining fee. These scenarios should be seen as a core which is supplemented by other revenue streams such as research and development grant funds.

Recommendation 13: Other revenue options such as angel donors or sponsors should be pursued in addition but as these are unpredictable sources of revenue, any receipts from such donations should be used to support discrete activity, transitioning the service, or adding to operational reserves/creating endowment funds.

Recommendation 14: A realistic estimate of the possible increase in member journals, along with a strategy and implementation plan to achieve the necessary increase needs to be undertaken.

Recommendation 15: An analysis of potential Dryad members is undertaken to assess how they might fit within tiered fees and use this as input into the development of a pricing policy.

Recommendation 16: Classify members and non-members, and review the policy on the acceptance of papers from non-members prior to the end of the current funding. Adjust the Collections Policy to align with an approach to pricing that would reflect the costs of non-members depositing and the fact that they would not be contributing an annual subscription.

Section 7: Risks

An initial risk register for the sustainability of Dryad is provided in section 7 for your review and future development.

Recommendation 17: The risk register is developed and regularly reviewed by the Board.

DRAFT

2. STRATEGY, PERFORMANCE INDICATORS AND MEASURES

2.1. THE VISION

Dryad will enable the preservation, discovery, sharing and reuse of publication-associated data for a large scientific community. It represents a unique collaboration among diverse institutions and expert communities as well as a pioneering application of digital library technology to data sharing for “small science”. It is intended that this will serve as a model for efforts to preserve and share data in other disciplines facing a similar crisis of data attrition.

2.2. THE ORGANIZATION

The Dryad Consortium Board is made up of representatives from participating journals. The senior personnel of the Dryad project serve as *ex officio* members. An Executive Committee appointed by the board is responsible for repository policy and business goals, and has the discretion to bring these to discussion and vote by the board. Issues of governance and sustainability are brought to the full board for discussion and vote.

During the NSF-funded project phase, overall co-ordination of Dryad is the responsibility of NESCent under the direction of the Associate Director for Informatics (T.Vision). The NSF funding currently runs from 2008 to 2012.

In taking Dryad beyond its current project phase the board will remain an essential element. More about the post-project organizational model is found in section 5.2.

2.3. SUMMARY OF STRATEGIC GOALS AND KEY PERFORMANCE INDICATORS

Strategic Goal 1:

To preserve the underlying data reported in partner journal papers at the time of publication.

Key Performance Indicators:

- *Adoption and implementation of the Joint Data Archiving Policy or its equivalent by the individual journals within the Dryad consortium.*
- *Achievement of a low-burden archiving process for authors through integration of the paper and data submission process, handshaking with other repositories, automated metadata generation and other technical innovations, to the satisfaction of the governing board.*
- *Achievement a high rate of compliance with the letter and spirit of data archiving at partner journals.*

Strategic Goal 2:

To promote the reuse of archived data.

Key Performance Indicators:

- *Demonstration of ease-of-discovery and retrieval of data within Dryad, and between repositories, to the satisfaction of the governing board.*
- *Assignment of permanent, resolvable, unique identifiers to Dryad holdings, and non-restrictive terms for reuse.*
- *Demonstration of the frequent and appropriate reuse of Dryad holdings, including compliance with best practices for attribution.*

Strategic Goal 3:

To empower stakeholders in the governance of Dryad.

Key Performance indicators:

- *Agreement to a long-term governance framework among a critical mass of partner journals.*
- *Continued growth of the consortium through recruitment of new partner journals.*
- *Alignment of the repository development plan with the goals of the governing board and, as determined by periodic assessments, repository users.*

Strategic Goal 4:

To achieve long-term financial sustainability

Key Performance Indicators:

- *Agreement to a business plan by the governing board.*
- *Obtaining funds sufficient for operating costs through a scalable cost-recovery model, including institutional commitments as needed, by the end of the project phase.*
- *Demonstration of the ability to obtain additional funds for repository enhancements, as needed.*

Recommendation 1: The Board adopts the draft strategy performance indicators and measures proposed in section 2.

3. COMPARATORS AND UNDERSTANDING THE COSTS

3.1. DATA SHARING FOR SMALL SCIENCE

The most established and well populated research data archives or repositories largely concentrate upon the requirements of “large science” disciplines, such as astronomy and particle physics. They have naturally arisen out of the systematic collection of the primary data by a few data centers. Although these data centers may have a mixed funding model, overwhelmingly they are supported through grants from their disciplines.

So far there have been relatively few initiatives that focus upon enabling researchers in “small science” disciplines to share and preserve research data supporting publications. Examples include Systematic Biology, which for some years has been maintaining an archive of data files from all its published papers, and the NSF BioLit pilot project to integrate the Public Library of Science (PLOS) journals with relevant entries in the Protein Data Bank. In other disciplines, the Archaeology Data Service maintains an archive, a component of which is supplementary material published by Internet Archaeology.

In addition to discipline specific efforts, there has been a growing momentum around developing multi-discipline institutional repositories. Out of the information and library science community have come popular tools and open source repository platforms such as DSpace. These, in turn, are well suited to the needs of a small-science digital data repository. Dryad is in a position to both use and advance these technologies. Furthermore, being at the forefront in addressing a model for sustainability, Dryad can serve as a model for efforts to preserve and share data in other disciplines.

3.2. EXTERNAL COMPARATORS - JOURNALS AND PUBLISHERS

This section provides a summary of a survey of 12 journals’ and organisations’ experiences in working with supplementary materials and specifically supplementary data. The information has been collected to inform and provide input to Dryad’s sustainability plan.

The survey was undertaken by telephone interview, although email responses were accepted where a call was not possible. We asked about interviewees’ job roles and the impact of supplementary materials and data on these; costs and policies; funding sources; and recommendations for Dryad’s sustainability. Three Dryad participants interviewed were also asked for their views on the collaboration. A range of roles and organisational types working with supplementary materials and data were selected for interview including journal editors and collection or archive managers, scholarly societies, publishers and archives.

Findings:

Policies for supplementary materials and data

- quality control of supplementary data, in terms of peer or editorial review is variable;

- there are currently few policies by specific journals covering metadata for supplementary materials or for long term preservation or access. Exceptions include: Internet Archaeology where supplementary data was handled by a third party; and policies developed collectively by partners in the Dryad consortium.

Handling supplementary materials and data

- in four cases minimal impact on work duties was reported; in five others there was a significant but often unquantified impact (two of these might be considered data publications with a focus on publishing data papers or datasets); and in three cases the information was not available or unknown;
- the variable level of impact on work duties can often be explained in terms of the level of effort or importance applied to supplementary materials and data: the greatest levels of effort are associated with copy editing, format migration, addition of metadata, etc, whilst the least effort is required for simply hosting the material; and/or high-levels of automation in the workflow;
- supplementary data appears mostly to have grown organically at the various journals investigated, with both the work and the costs being absorbed into the daily running of journals;
- in three cases, journals also specifically mentioned that they are experiencing rapid growth in the number of articles submitted with supplementary material and data: for one journal the percentage has increased from 2% a decade ago to 87% in 2009; for another journal the growth has been from 32 articles in 2000, to 251 in 2009 – an increase of 784%; finally for the publications of another society, journal articles including supplementary material rose from 6% in 2005 to 38% in 2009;
- for the three Dryad partners interviewed, contact with Dryad is currently low, probably because the initiative is relatively new and Dryad is phasing its roll-out to participants.

Costs for supplementary materials and data

- these were in most cases unknown or only partially known. Costs mentioned but usually not quantified include: digital storage costs, often met by the publisher; salary costs of journal staff working with the archives; and long term preservation costs.
- detailed cost information was really only available from one of the interviewees who had participated in an activity based costing study.

Sources of revenue for supplementary materials and data

Current sources:

- only author fees and journal subscription fees were mentioned as current revenue sources for the supplementary materials in journals;

- three journals interviewed have author charges for supplementary materials. These are as follows:
 - Journal of Clinical Investigation - authors are charged \$300 for supplemental data to appear online with accepted articles;
 - Ecological Archives - submission of 'appendices and supplements' is free up to 10MB. Above this, there is a fee of \$250 for the first 1 GB and \$50 for each subsequent GB. The fee for publication of a data paper is \$250 for publication of the abstract in the relevant journal plus publication of up to 10 MB in Ecological Archives. An additional \$250 is charged for data sets between 10MB and 1GB, and for larger datasets there is an additional \$50 per GB fee;
 - The Federation of American Societies for Experimental Biology (FASEB) charges \$100 for each Supplemental file.
- the data archiving and sharing organisations interviewed relied primarily on (uncertain) research grants and temporary or re-current core funding, but one had access to a small endowment and another has a charging policy for some depositors.

Potential other sources suggested by interviewees were:

- government funding: particularly considering the requirement of funding bodies to make data available;
- learned societies: possible 'community support' via subscriptions to a service;
- publishers: some publishers host such material within their fee to societies or subscribers, so might contribute to a service that undertook this role: it is a bundled cost by publishers to societies and subscribers (even if some perceive it as "free").

Interviewees' Recommendations

- Principal among these were to endeavour to build a community of stakeholders by entering into a serious dialogue with publishers, authors and journal editors. Wiley and Dataverse specifically stressed that they would welcome further dialogue with Dryad.

Conclusions

With regard specifically to cost and revenue considerations, little is known by journals about the specific costs of handling supplementary materials and data. Costs, principally staff time, were observed to vary according to the tasks undertaken and value added.

Economies of scale were also noted as important. It is currently not possible to compare costs with those for Dryad as they are largely unknown, or for tasks such as adding metadata etc which are not currently undertaken.

However it could be observed that proposed costs for Dryad (although it excludes some tasks and adds others compared to the journals) appear very reasonable compared to

existing author charges where these exist for publishing supplementary data files. It suggests further work on establishing activity based costings for supplementary data and journals and for Dryad would be likely to confirm the business case for journals supporting Dryad as a shared data repository.

3.3. EXTERNAL COMPARATORS - OTHER ARCHIVE REPOSITORIES

As far as can be ascertained, Dryad is currently unique in its approach as an archive for research data associated with publications, and with its structure based on leadership by journals. Thus, direct comparators with Dryad are not possible. However, its principal repository activities align in most cases with archive comparators such as the Archaeology Data Service (ADS) and Kings College London Centre for e-Research. Like Dryad, ADS and KCL Centre for e-Research serve the needs of small disciplines (for further information on these comparator services see Appendix 1 Table 1). Differences include the ingest processes across all three repositories, and the technologies for maintaining data integrity (e.g. LOCKSS). The ADS and KCL repositories work directly with the researcher at ingest advising on formats for deposit, metadata and access interfaces and requirements. The ingest process with Dryad is more automated, and integrated into a journal's submission system. Metadata creation will be automated and curation limited to validity checking and metadata enhancement with data quality the responsibility of the submitter. In this way, Dryad will be able to save on staffing resource as compared with ADS, KCL and others as the ingest will largely automated via the journals. The work that Dryad has done during 2009 and is continuing to do in 2010 with a number of pilot journals will provide the opportunity for these workflows to be assessed and quantified within a live environment. The experimental solution for Dryad preservation based upon a LOCKSS system will address bit storage and preservation but activity around file format storage and migration will be the equivalent to comparator repositories.

Observations found in a number of studies including "UKRDSⁱ," Keeping Research data Safe, and "Digital Archive Costs: Facts and Fallacies"ⁱⁱ indicate that staffing costs are likely to be the single biggest cost of running a repository. To consider what should be the minimum viable staffing upon which a repository can be run together with required staff profiles to ensure sufficient cover, the experience of repositories at their launch is useful to note and examples from the Archaeology Data Service (ADS) University of York and the Centre for e-Research King's College London are included in Appendix 1 Table 2. The examples are similar to each other and although the staff profiles may differ for Dryad, to address its unique model, considerations about covering for absence and recruitment need to be taken into consideration. Note the absence of technical/ programmer staff below as the platform hosting and system systems support is contracted to university computing services for both ADS and Kings. Development activity is funded separately through other projects.

A similar institutional data repository effort to that of Kings at Cambridge University (DSpace@Cambridge), launched with 3 FTEs and recently increased to 4 FTEs.

In addition, although again not a direct comparator, there is generic guidance about possible staffing levels to establish a DSpace Institutional Repositoryⁱⁱⁱ from MIT. It does assume an embedding within an institutions’ existing support structures, and accounts for operations and programming effort through a category of “systems and equipment costs”. The staffing considered to have primary responsibility for the service is listed as follows with the caveat that “both the number of staff and the level of commitment (full or part-time) will be defined by the service model”.

Senior manager	0.05 FTE (5%)
Librarian	0.2 FTE (20%)
User Support Manager	1 FTE (100%)
Total	1.25 FTE (excluding operations and programming)

Table 3: DSpace Institutional Repository minimum level estimates for primary staff.

Recommendation 2: There will be a minimum level of core staffing needed for Dryad to ensure sustainability and ongoing service. Relevant issues include cover for absences or staff turnover, viable staff recruitment, and a range of skill sets.

Recommendation 3: Promotion of the Dryad service will be critical in early phases and beyond. A staffing that includes a full-time executive director is likely to be required for successful implementation of the plan.

Recommendation 4: The proposed staffing level and required job descriptions and costs should be reviewed by the Dryad Project team in conjunction with the current host institution (Duke University).

3.4. EXTERNAL COMPARATORS: INSTITUTIONAL SUPPORT AND EMBEDDING

The comparators we have looked at leverage effort, expertise and direction from associated staff, often co-located but funded separately working on related activity, including support services, project based research and development. This helps to maintain a dynamic and sustainable organisation that can respond to change and deal with fluctuations in staffing. The fact that they are embedded within larger institutions means that they can leverage advantage from the stability and longevity of these establishments and pre-existing infrastructure. Dryad currently is similarly embedded within a larger institution. Dryad needs to consider what will be the most cost effective way of providing its administration and

infrastructure support going forward and whether support from a host institution can be negotiated at a cost or provided as an “in-kind” contribution.

Recommendation 5: Until Dryad reaches a point in its maturity and development where it can effectively be transitioned into a separate not-for-profit organisation it should work with the host institution to ensure continued support.

In addition to core or priority staff, there are a range of additional staff funded via research grants. In established repositories, we have found typically around 50% of income and staff is core and 50% supported via research project funding. We believe Dryad should look to mixed funding model and identify any tasks that can be funded from other sources than its core income. In particular it should seek project funding for innovative research and development wherever possible.

3.5. EXTERNAL COMPARATORS: USE OF STUDENTS AND VOLUNTEERS

None of the external archive comparators use students or volunteers for archive tasks. However use of volunteers in some institutions such as museums or use of students as a labour source in universities is very common. In recent years there has also been an increase in involving volunteers via the web in scientific projects. This is often referred to as “citizen science” when individual volunteers or networks of volunteers, many of whom may have no specific scientific training, perform or manage research-related tasks such as observation, measurement or computation. The use of citizen-science networks often allows scientists to accomplish research objectives more feasibly than would otherwise be possible.

It has been suggested to us by members of the Dryad Board that some of the most basic curation tasks in Dryad might be undertaken by students as paid work. This will provide a potentially cheaper labour force although there will be some higher costs for core staff associated with training and supervision of students if this option is taken.

We believe Dryad would not currently benefit from outsourcing any activity to foreign labor markets although if the work grows significantly in future years this option could be kept under review.

Recommendation 6: Dryad could pilot and review use of students as part of its staffing for curation.

3.6. EXTERNAL COMPARATORS: ACTIVITY MODELS FOR ARCHIVES

Activity (or “lifecycle”) based costing is a very well-established technique across a wide range of sectors and has recently been applied to a number of research data centers and their activities for research data. The activity model below (see Table 4) from “Keeping Research Data Safe”^{iv} was developed in consultation with a number of data repositories. It provides a guide as to the most common activities that need to be considered in costing a

data repository. The activity headings in the table are at a relatively high level and a detailed description and further breakdown can be found in the refined model emerging out of “Keeping Research Data Safe 2”^v.

<i>Pre-Archive Phase</i>	Outreach
	Initiation
	Creation
<i>Archive Phase</i>	Acquisition
	Disposal
	Ingest
	Archive Storage
	Preservation Planning
	First Mover Innovation
	Data Management
	Access
<i>Support Services</i>	Administration
	Common Services
<i>Estates (includes leasing of premises, and space management / maintenance)</i>	

Table 4: Main phases and activities of the KRDS2 activity model

A case study example where this activity model has been applied is that of the UK Data Archive. This organisation is a large social science data archive and not a direct comparator to Dryad as it is a large multi-functional service and research centre employing over 50 staff. However, this example does illustrate how this model can be used.

The overall activity costs in the UK Data Archive are shown in Table 5.

Activity	2009	2009
	% cost	% time
Archive: Acquisition	5.8	4.8
Archive: Ingest	21.5	22.2
Archive: Archive Storage / Preservation Planning	3.1	2.8
Archive: Research and development	6.9	6.9
Archive: Data Management / Information Development	15	14.7
Archive: Access	16.9	16.3
Support Services: Administration	21	23.8
Support Services: Common Services	5.1	4.9
Other	4.8	3.7
Total	100	

Table 5: Allocation of costs by activity for the UK Data Archive.

Ingest costs cover the activities of receiving, reading, quality checking, cataloguing, of incoming data (including metadata, documentation, etc.) to the point of insertion into the archive. Ingest can be manual or electronic with manual steps involved in quality checking. In common with many other national data archives there is a high-level of added value from normalisation of file formats, addition of metadata and documentation to make the submitted material more accessible and suitable for re-use. Access focuses on the services and functions which make the archival information holdings and related services visible to consumers, including a high-level of user support (help-desk, user workshops and training etc). The “other” activities, making up approximately 5 per cent of the spending, covered: disclosure checking, reformatting services, project management of external projects as well as some non-work related activities. Note the UK Data Archive, hosted by the University of Essex, carries out almost all its own financial and human resources activities with limited assistance of the host institution so administration is 21% of effort. Removing these activities from the costs of the UKDA would reduce the cost of support services by around

12 per cent, but might increase the institutional overhead costs for UKDA charged by the university to external funders.

When applying this model to Dryad the differences as compared to other repositories need to be considered. For example, journal integration puts higher upfront cost on recruitment and coordination of journals, but then ingest per data package is expected to be relatively inexpensive, and much of the “cataloguing” work is taken care of by the journals, with limited further curation of data files and some metadata needed. Dryad does not have copyright and disclosure issues, but does need to monitor embargo and DOI assignment. A higher visibility to end-users and an evolving expectation of service requires some programming and technical maintenance effort. The pilot work with journals presently underway provides an opportunity to map Dryad activity and costs across the data life cycle.

3.7. COST MODELS FOR DRYAD

Dryad is a specialist repository with a clear data collection policy. Some work has already started within Dryad on breaking down and analysing curation tasks and estimating time to undertake work at different levels of curation. This is something that can be expanded upon and reviewed and developed as the repository matures. As Dryad gains more practical application and experience of its workflows it will be able to use the kind of activity modelling described above to account for how costs are apportioned, thus assisting with future planning and resourcing decisions, and being able to target areas for efficiency improvements, where there will be the most impact. It is important that an accurate analysis of costs is undertaken and reviewed periodically to feed into sustainability planning and business decisions over time. It should be noted that a range of drivers influence costs (formats, volume, quality of metadata and documentation, etc). Care will be needed that the analysis of costs is based on a broad range of different journals and supplementary materials, and experience built up and reviewed over an appropriate timeframe.

Recommendation 7: Dryad should adopt an activity based costing model and build up knowledge of its activities, their costs and cost variables (volumes, file formats, etc).

Estimating a baseline

Following on from the KRDS2 activity model above (Table 4), by making a number of assumptions and adjustments, it is possible to begin to estimate an annual base-line cost for Dryad (see Table 6 below). One must keep in mind that costs will differ for the “start-up” or “transition” to a service repository from the current project. Also, as is true for all repositories, costs will change over time – an early stage project faces different costs than a mature service running at capacity. Creating a cost model is an iterative process. In particular it should be expected that a significant component of staff time at start-up should

be devoted to marketing and advocacy to ensure appropriate growth in support from journals, scientific societies, and authors.

Beyond the estimation of a raw baseline, certain adjustments need to be considered including economic adjustments covering inflation/deflation, and service adjustments covering other major variables affecting research data preservation costs over time e.g. file formats, data volumes or required metadata, documentation and IPR.

Assumptions and caveats for the preliminary baseline

1. The preliminary baseline is based on initial figures supplied by the Dryad project team and will need to be refined and developed in coming months.
2. The **minimum** level of appropriately skilled staff is sufficient to (a) deal with the **lower** estimated number of datasets per annum ingest rate of 5000, based on an analysis by the Dryad project team of partner journal articles assuming 30 participants; (b) grow journal participation and support for the repository at levels required for sustaining the repository.
3. More curation effort will be required if the ingest rate raises above the minimum or if curation is required above the lowest level.
4. No peer review for supplementary data, (JAMA curation level “zero” re editing of data – see JAMA model in next section) and lowest level of curation by Dryad.
5. A 5 year equipment cycle with procurement fund to be accumulated per annum to provide for capital replacement in 2015. (Based on equipment budget for hardware for period 2010-2015.) Hardware maintenance costs not included.
6. No inflation factor built in and cost estimates not verified against local conditions.
7. Personnel benefits at a rate of 20%. Salaries only based on estimates until future job descriptions and grade levels agreed with host institution.
8. All personnel are co-located for viable service cover and flexibility. Fractions of FTEs are amalgamated as far as possible into viable full-time or half-time FTE job descriptions for ease of management and recruitment and retention.

	FTE %	Annual cost estimate (\$k)
Staff costs (FTEs)		
Part-time Director (10% of PI)	10%	10.8
Technical Manager, including technical coordination / support (1 FTE but supported half time by external funding leaving 0.5 FTE funding for Dryad base-line)	50%	36.0

	FTE %	Annual cost estimate (\$k)
Communications and outreach officer	50%	24.0
Programmer	25%	20.0
Curator [Curating a max of 5000 papers at a minimum level, <i>OR equivalent using student curators</i>] See note¹ and table below for “hours” calculation and student curator approach with a breakdown based on different levels and throughput of papers.	72% OR 1250 hours	43.2 OR 18.75 + management o/h
Total Staff	2.07	\$134,000 (OR) \$109,550
Non-staff items		
Equipment costs averaged p.a. over 5 years (requirements <i>to be reviewed</i> once workflows fully established and new hardware in place after 2010) Excludes hardware maintenance costs.		18.6
Travel (staff)		5.0
Travel (board meetings)		10.0
Promotional materials		6.0
Outsourcing – facilities management (FM)		<i>To be determined</i>
Total non-staff items		\$39,600 (+FM)
Support services (in-kind contribution to be investigated)		
Secretarial and accounting	50%	24.0
Equipment costs (networks, office etc)		6.08
Office supplies and expenses		3.0
Estate costs (office accommodation)		<i>Local calculation to be applied</i>
Support services total		\$33,000 (+ accommodation)

Table 6: Estimated minimum viable base-line cost for Dryad at launch.

¹ **Note** that the number of hours per annum for the Curation Officer in table 6 is based on 46 working weeks x 5 days per week x 7.5 hours per day = 1725 hours per annum. Also the management overhead associated with employing student curators needs to be accounted for in addition to the rates in table 7 below. Effort will be dependant upon the numbers employed and the student curator turnover rate and the type of QA processes put in place.

Note that the number of hours per annum for the Curation Officer in the table above is based on 46 working weeks x 5 days per week x 7.5 hours per day = 1725 hours per annum. The management overhead associated with employing student curators needs to be accounted for in addition to the rates in table 7 below. Effort will be dependent upon the numbers employed and the student curator turnover rate and the type of QA processes put in place.

Curation level	Estimated time to curate per paper ²	Cost for 5,000 papers	Cost for 10,000 papers
1	15 minutes	1250 hours*\$15= \$18,750	2500 hours* \$15= \$37,500
2	30 minutes	2500 hours*\$15= \$37,500	5000 hours* \$15= \$75,000
3	60 minutes	5,000 hours*\$15= \$75,000	10,000 hours* \$15= \$150,000

Table 7: Alternative cost profiling for curation of data by Dryad based on employing student curators at \$15 per hour.

Change over Time Factors

In order to estimate annual costs over time it is useful to take into account key findings from case study observations.^{vi}

Long-Term Digital Preservation Costs

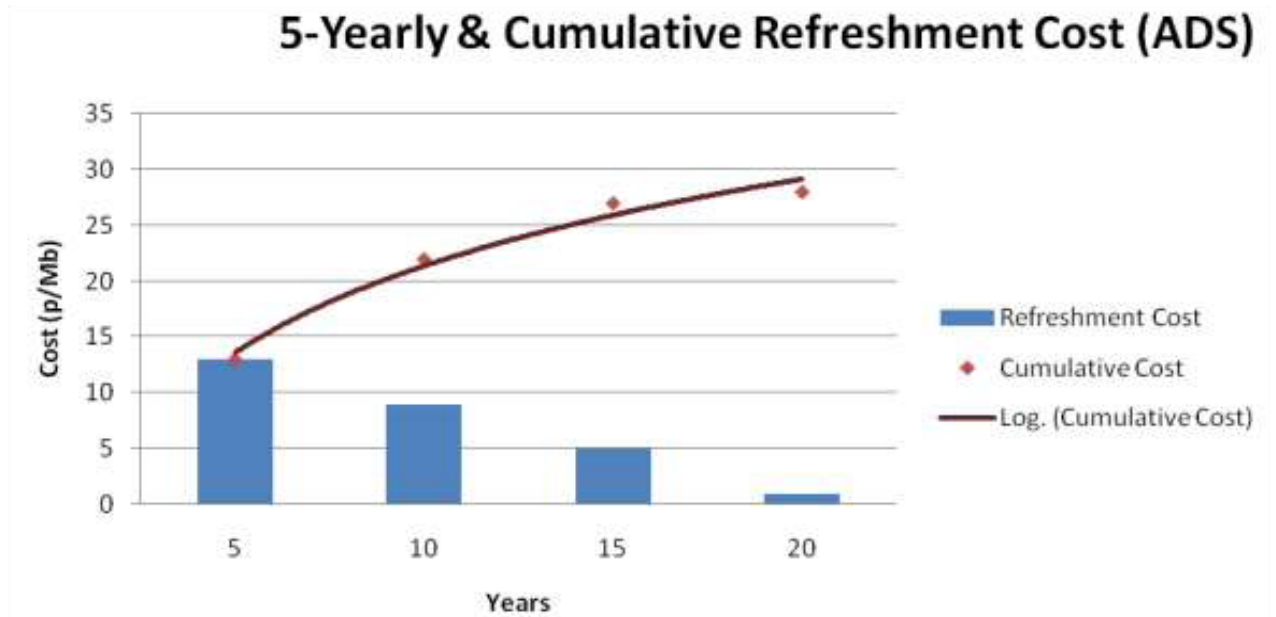
Over time, as the content within the repository grows, it is important to take into account the long-term preservation responsibilities of the repository and the cost implications. Accessioning and ingest costs within repositories investigated were higher than ongoing long-term preservation and archiving costs. For example, the following approximate division of costs across high-level archive functions of the activity model was suggested for the UK Data Archive:

Acquisition and Ingest - c. 42%; Archival Storage and Preservation - c. 23%; and Access - c. 35%.

The Archaeology Data Service (ADS) provides a useful projection of its long-term preservation costs for research data based on its costs to date and ongoing trends. It shows relatively high costs in the early years after accessioning but costs declining to a minimal

² **Note:** Estimated curation times are preliminary figures supplied by the Dryad team.

level over 20 years. Nonetheless, the 20-yr outlay is approximately twice that for the initial 5-yrs.



The ADS projection is a complex mix of underlying trends such as long-term declining data storage costs, costs for ongoing actions such as preservation interventions (file format migrations), and assumptions of archive growth which provide economies of scale.

Different Preservation Aims and Data Collection Levels

Differential spending will be dependent upon the extent of collection level data being handled and editorial effort applied. The Journal of the American Medical Association (JAMA) has developed a model for editing supplements which has relevance for Dryad. JAMA categories for the editorial effort applied are as follows:

Level 1 (or zero!)

- No editorial review with a disclaimer to that effect
- Not used, inconsistent with our brand promise

Level 2

- A “little” editing and no composition or proofreading
- Content bypasses the production department
- Used for JAMA and some Archives Journals

Level 3

- The complete editorial and production workflow
- Same as print

- Used for some Archives Journals

Consideration should be given to developing Dryad curation levels that reflect the variability of curation required for Dryad's data packages and differential curation effort to be applied based on the download frequency or educational value of the data package. The "handshaking" functionality being developed with TreeBASE and GenBank will assist depositors and enrich Dryad metadata. An assessment of the resources needed to develop and maintain this service should contribute to the longer term understanding of repository costs. Possible synergies in staffing among the repositories may be exploited in order to reduce the need for some capacity within Dryad itself.

Recommendation 8: Dryad should further develop the application of curation levels as a variable that can define and influence selected costs.

4. ADVANTAGES, BENEFITS AND REVENUE

4.1. COMPETITIVE ADVANTAGES AND UNIQUE SELLING POINTS

Dryad has emerged from a recognized and well articulated need within the scientific community. Dryad's primary goals have been formulated and agreed by its community of interest through workshops and meetings. It has emerged as a collaborative and cooperative venture with stakeholders, through its governing board, having the ultimate responsibility for the nature of the enterprise. This collective engagement provides a very strong foundation and advantage to Dryad in realizing its vision to become a self-sustaining repository in the service of its stakeholders.

Currently based within the research community, Dryad is able to leverage credibility and support for its work with current and potential researchers through education, outreach and publication, and link with other informatics projects that contribute to the digital landscape of evolutionary biology and ecology.

In raising its profile, the name "Dryad" should be protected as an asset including possibly trademark registration.

4.2. CORE COMPETENCIES AND ASSETS

There are three main areas to draw attention to that, in combination, provide Dryad with a strong foundation upon which to build:

Stakeholders

The growing consortium of journals, associated with a variety of publishers and societies, their willingness to work together on the Dryad governing board, and their ability to progress with strategy and policy issues (such as the Joint Data Archiving Policy), is a significant asset.

Staff and their expertise

A major asset is the staff involved with the delivery of Dryad. Expertise has been gathered from multiple organizations and disciplines, including both domain scientists and information technologists.

Establishing and hosting the repository

The early work of NESCent in establishing the repository and from the outset being able to provide a production environment for Dryad through the Digital Library Initiatives (DLI) group at North Carolina State University is a positive benefit given its experience and partnership in the Library of Congress's National Digital Information Infrastructure and Preservation Programme (NDIIPP). It establishes a confidence that Dryad is serious about reliability and availability.

4.3. SUMMARY BENEFITS OF DRYAD

Summary Benefits of Dryad	
Publishers	<ol style="list-style-type: none">1. Providing a permanent data hosting environment, reducing editorial effort.2. Citation and discovery of data is likely to increase readership and citations of related research papers.
Societies	<ol style="list-style-type: none">1. A conduit to support and promote its disciplines research outputs and encourage new research.2. Independent verification of research findings is possible.
Journals	<ol style="list-style-type: none">1. Enriching the product through access to underlying data and "handshaking" with specialized repositories.2. Improving author provided metadata through both automatic and manual curation.3. Source of data curation expertise.4. Stakeholder governance.5. By pooling resources, avoid duplication of effort and leverage economies of scale.6. Citation and discovery of data is likely to increase readership and citations of related research papers.
Research Funders	<ol style="list-style-type: none">1. Maximizing investment in research by promoting and enabling Open Data.2. Educating future authors in the practice and benefits of

Summary Benefits of Dryad	
	<p>Open Data by enabling data reuse.</p> <p>3. Reducing redundancy of funding for primary data collection through the efficient reuse of existing data .</p> <p>4. Enabling new research directions through data availability.</p>
Depositors	<p>1. Removing barriers to data deposit by integration with journal submission and one-stop data-deposition.</p> <p>2. Enabling underlying data to be effectively cited, leading to increased citation rate.</p> <p>3. Freeing depositors from the burden of data management for sharing and preservation.</p> <p>4. Increasing discovery and impact of published works, opportunities for new collaborations.</p>
End-users	<p>1. Providing easy access to data both through data citations and through direct data discovery.</p> <p>2. Enabling the use of published data for research and education.</p> <p>3. Providing the opportunity to develop and use new research tools that exploit the integration of existing data.</p>
Universities and Research Institutes	<p>1. Promoting the work of the institution and its researchers.</p> <p>2. Reducing unplanned data loss.</p>
Related repositories	<p>1. Reciprocal harvesting of metadata with other repositories, facilitating data discovery and retrieval.</p> <p>2. Increasing data holdings through low-burden data submission and handshaking.</p>

Table 8: Summary of Dryad's benefits and advantages to main stakeholder and interest groups^{vii}

4.4. PRIORITIZING REVENUE STREAMS

In the experience of the study team, national or subject repositories are funded in the main through a mixed economy of core and project funding where a maximum of 50% core funding is the norm. Although this can lead to tensions in balancing priorities, diverse revenue streams offer a realistic path for sustaining continued funding and provide some flexibility to decisions around future development.

The easiest and often most successful approaches for projects looking at sustainability issues and possible revenues are to seek out the “paths of least resistance”. That is to examine which “groups” most benefit from the outcomes of the initiative or project and assess their ability and willingness to provide continuing support. Multiple revenue streams

can be hard to manage and will bring an additional overhead to the organisation that should not be underestimated, so a necessary balance has to be found between the risk of being dependant upon just one or two revenue streams or that of spreading risk across many but then having to deal with managing them.

A number of possible revenue options have been considered and discussed. These are listed in table 9 below. In addition to these, another option of in-kind contributions is useful to consider in the context of a mixed funding model. These can be both explicit and implicit, from academic institutions, sponsors, and other organizations. These in-kind contributions—including, for example, office space, administrative support, technical hosting and support, etc.—effectively offset costs that would otherwise be incurred.^{viii}

The following table prioritizes revenue options against perceived benefits and provides a crude assessment of where application is likely to result in most reward. Although somewhat subjective, it does point to options that will have low or high impact. Also, consideration is given to what point in the planning cycle it will be feasible and appropriate to initiate any of these options. At the point where function and service is being delivered rather than awaiting the end of the current project funding period should be considered.

	Revenue raising options	Target group & perceived benefit	Potential to raise funds and priority	Timing
A	In-kind contributions	Host institutions and participants - collaborators in Dryad	HIGH. Contributions to offset costs that have to be met and to also demonstrate commitment to the initiative. They can add up to be quite significant and may include offices; general infrastructure needs such as networks and office systems support; etc.; use of a contracts / legal office; HR office.	Discuss and negotiate what is possible early on.
B	Annual Subscriptions (in return for full partnership benefit)	Journals: benefits include stakeholder governance.	HIGH. A fee levied across all the journals could account for significant fraction of the operational budget.	Look at bringing this in as soon as Dryad is offering a functional service.
C	Fees – joining, and use fee based on per paper charge	Journals: Participating journals paying a fee to support Dryad, maybe in the form of a one-off joining fee followed by annual billing based around use – a per paper charge. High degree of control through the Dryad governing board – special relationship with repository.	HIGH. Dependant upon the number of journals joining annually and the number of papers with supplementary materials being deposited in Dryad. Per paper fees could be tiered based on usage or revenue levels.	Look at bringing this in as soon as Dryad is offering a functional service.

	Revenue raising options	Target group & perceived benefit	Potential to raise funds and priority	Timing
D	Grants	Research funders. Where a Dryad repository helps to further policies around preservation and access to research data. Also maximises benefits and returns on investment in research.	HIGH. Funding via NSF and others to be explored. Proposals for related and development activity.	At least 1 year before current funding ceases.
E	Charging model for large data packages	Research Scientists: benefit in addressing preservation and access challenge, sometimes a requirement of a research funder. Datasets requiring customised interfaces could be subject to additional charge (as done by ADS).	MEDIUM. A charging policy for data packages that exceed limits of Dryad's collection policy. Likely to be limited take-up but possible to deal with on a case-by-case basis. Likely success only if researchers include cost of data preservation and access in their proposals.	Only look at bringing this in once Dryad is fairly established. Take into consideration demand.
F	Angel donors	Organisations or benefactors with particular interest in the subject specialty combined with an interest in access to and preservation of data.	LOW. Tend to be one-off. Difficult to predict. Can require considerable work to solicit such donations.	N/A
G	Advertising or sponsorship	Advertisers and Sponsors wishing to draw attention to their organization or product through the Dryad website	LOW. Academics tolerance to advertising can be low. However, sponsorship can be effective and may come from a wider variety of organisations. Work needed to attract both.	Need to demonstrate website hits and make-up of audience/users to make attractive to advertisers and sponsors.

Table 9: Revenue options for Dryad and their potential.

Recommendation 9: Dryad must agree the options for revenue creation and should seek diversity as the best route to sustainable funding. It should pursue grant funding for innovative research and development wherever possible.

5. THE PROPOSAL FOR SUSTAINABILITY

5.1. CONSIDERATIONS GOING FORWARD

This section aims to provide a realistic direction of travel for Dryad, providing the stability it needs to establish itself as the repository of choice for evolutionary biologists and ecologists, extending to other “small science” disciplines over time. Dryad must address:

- The organization, its status, overall structure, and leadership Governance model.
- Generating “good will”, promoting an understanding of Dryads core aims and objectives with subject specialists, and a transparency with stakeholders to maintain and grow support.
- The ability to generate new revenue streams to match costs.
- Development planning and resourcing.
- Delivering a stable and reliable service environment.

5.2. ORGANIZATIONAL MODEL TO BE ADOPTED

Status

Given the nature of Dryad and its chief goals, the intention to remain “not-for-profit” is completely in line with the approach other specialist subject repositories have taken. This still leaves options to be considered. Dryad is currently a project hosted and embedded within a “not-for-profit” organization. If the host institution is willing, a recommended transitional approach would be for Dryad to remain “hosted” beyond its project and into its service phase. The host organization will be the legal entity that enables Dryad to formalize its relationship with partners and stakeholders relatively quickly, enabled by the mature host infrastructure.

The step to being a separately registered not-for-profit organization will be significant. The implications of becoming a separate not-for-profit organization should be investigated in detail and planned. It should take into account: set-up costs; employment issues (including possible transfer of staff); relationships with “delivery” organizations that Dryad might contract with to deliver parts of the service and/or development e.g. facilities management, curation activity, possibly R&D; legal, financial and tax issues.

Organizations such as Portico and JSTOR have successfully taken the approach of not-for-profit organizations but even these large services have merged with a supporting organization, Ithaka.^{ix} A different kind of organization that has been spun off to form a separate “not-for-profit” business is UK based JISC Collections established by the UK’s Higher Education Funding Agencies.^{xxi} Each of these may provide lessons and exemplars for Dryad should it wish to pursue a similar future organizational transition.

Structure

High cost structures impose a steeper hurdle for any income-generating model to clear, while low cost structures will allow Dryad greater flexibility in selecting a sustainability model. Thus, in terms of initial risk perception, Dryad may elect to deploy the lowest sustainable operating structure in order to substantially limit costs.

At start-up the Dryad repository service will have a directly-employed minimal viable team. The structure needs to take account of the management of infrastructure and support arrangements with external parties whether within the host or within other organizations. The estimated initial staffing will depend upon outsourcing of a range of activity, including facilities management for the technical infrastructure hosting; finance; contracts and HR. A priority has to be a stable and reliable service.

5.3. GOVERNANCE AND PARTNER COMMITMENT

What makes Dryad different, with the potential to sustain its viability over time, is the principle upon which it is built, as a repository for data relating to published articles. Since the governance model needs to reflect this relationship between the repository and the journals, the governing board and Executive Committee will continue to have a role in supporting Dryad. Whilst Dryad continues to be hosted within an institution, the interests of that institution should also be taken into account.

Dryad will need to give careful consideration to evolving its governance structures for its service phase, particularly if it aims to grow the number of participants substantially. The host organization will need to be represented and involved in the governance framework.

Recommendation 10: We recommend Dryad discusses future governance structures with its stakeholders. Possible elements might include an elected Management Board (with ex-officio representatives); an annual participants' meeting; and an Advisory Board.

5.4. PROPOSED SUSTAINABILITY MODEL

At the heart of any sustainability model should be a clear articulation of the “value proposition”– how the organization provides a solution to a problem or delivers an attractive product otherwise difficult, expensive or impossible to obtain to its stakeholders and users. The questions are: what value can be placed on these solutions and products; what size are the communities that can benefit; and the size of Dryad consortium and of the economies of scale delivered to its participants. The feasibility of sustainability for Dryad will depend upon the following factors:

- The costs of maintaining the Dryad organization and its supporting technology
- The number of partner journals and the rate at which new data packages are ingested.

- Success in addressing the varying interests of multiple stakeholders, including journals, scientific societies and publishers
- The extent to which Dryad increases its visibility in the research community, there is increase in the practice of data reuse, and there is increased adoption of data citations.
- The extent to which Dryad can attract funding / revenue for both operating costs and continued development of its service.

Table 9 in section 4.4 gives an indication that the model Dryad should be looking at is one of mixed revenue streams with a primary focus upon per paper fees, partner journal subscriptions, and grants. Scenarios for different balances in these revenue sources are outlined in section 6.

The amount that can be raised from partner subscriptions will depend upon how many partners there are. Currently, there are 16 interim partners. Societies may publish one or more of these journals and in some cases contract with commercial publishers to do so. A limited number of partner journals are owned wholly by commercial publishers. This builds a relatively complex set of relationships when looking at the application and categorization of fees or subscriptions.

The approach must be transparent and simple, maximizing revenue at the same time as seeming to be fair so that goodwill is maintained.

The clearest approach is for the journal to have the direct relationship with Dryad. A membership and fee model based upon journal partnership is relatively straight forward and could result in the following:

- A one-off joining/membership fee for each journal. Consideration may be given to a reduced fee for every third / fourth/ fifth etc. journal from the same society or publisher. For example, a one-off fee of \$5,000 (without any reduced fees) for 20 journals would yield \$100,000. While this is not recurrent revenue, it could be used to contribute to a contingency reserve (angel donors might also contribute to building up endowment funds and the contingency reserves).
- An annual retrospective billing based upon pro-rata use. This could also be tiered based on usage or revenue levels.

Other approaches to longer term regular income can come through subscription. This may result in the following:

- A journal would be subject to an annual subscription, say negotiated for a multi-year term at a fixed price. This would give a predictability of income for Dryad and indicate a clear commitment from journals. The perceived benefit of such subscription dictates the level. For example, subscribing organizations may be eligible to become a member of the Board. Therefore, $N(\text{number of subscribers}) * AY(\text{annual rate}) * (\text{committed years}) = R(\text{revenue})$. Consideration should be given as to whether this

should be tiered dependant upon the resources or usage of the subscribing organization.

The number of partners will put limits on achievable revenue. Increasing the number of partners clearly will help, but it will be necessary to actively bid for transition grant funding that can support elements of Dryad activity whilst the consortium is grown.

As Dryad matures it will still need to be responsive to user and stakeholder requirements and to address changes and advances to technology that could impact upon its systems and interfaces. Two complementary approaches to addressing this continuing activity are through research and development projects, separately funded; and through association with other groups and organizations active in related fields. Dryad already has the foundations, links and experience to enable this to happen and we have recommended (Recommendation 9) that Dryad should pursue grant funding for innovative research and development wherever possible.

5.5. FURTHER INTERNATIONALIZATION OF DRYAD

Although based in the USA, Dryad is already operating with journals that publish science from authors worldwide and with scholarly societies and publishers operating both in the USA, Europe or elsewhere. A component of the sustainability plan should be to develop and consolidate its existing international links and opportunities for funding as part of its longer-term business plan. We suggest that if LOCKSS is implemented as part of the preservation infrastructure for Dryad some LOCKSS nodes might be based outside the USA as a starting point of this strategy. Dryad might also consider mirroring, funding and development partnerships along similar lines to those that have evolved between PubMed Central in the USA and UK PMC or PMC Canada through PMC International (PMCI)^{xii}.

The long term goal of PMCI is to create a network of digital archives that can share all of their respective locally deposited content with others in the network. There are three primary reasons for doing this:

- the probability of an archive surviving over the long term is greater if there are working copies of the archive in regular use at multiple sites around the world;
- a producer or funder of research literature often will be more inclined to make the primary deposit of its material to a locally or regionally affiliated archive, rather than to one operated elsewhere in the world;
- Each site can integrate the journal articles in the archive with related material, such as national or regional practice guidelines, that has particular significance to its users.

Collaboration with each center is based on a formal agreement between the NLM and the center's sponsors concerning the management of the archive. This agreement protects the rights of those who deposit content in PMC and ensures the integrity of their content by requiring each center to use NCBI's pPMC software. It also ensures that users will always

have free access to the content. This may be an attractive exemplar for longer-term development of Dryad.

5.6. NARROWING THE GAP BETWEEN COST AND REVENUE

During the transition from project funding, it will be necessary to address how to narrow the almost inevitable gap between cost and revenue in light of realistic projections for how revenues and members will grow over time. Some costs may be off-set by in-kind contributions from the host or partner organizations. This might include stakeholders taking on activities that entail little overhead to their organization but are costly to Dryad, such as having some of Dryad's publicity and outreach needs provided by societies and publishers. Section 6 below examines some possible revenue raising scenarios and provides discussion around how a funding gap may be addressed.

An early dialogue with NSF about continuing partial funding and transitional support as other revenues are grown will help in the prioritization of other proposals and approaches that may be made to funding organizations.

Recommendation 10: The views of funders on transitional and ongoing grant support for Dryad will need to be investigated further and future funding options and opportunities regularly reviewed by the board.

5.7. PLANS FOR TRANSITION AND IMPLEMENTATION

The transition from Dryad's development phase to Dryad as a repository service requires careful planning and development of a transition strategy. The main considerations will be around organization of business relationships and governance; staffing levels; maturity and reliability of automated processes to sustain the repository; and the level of active marketing and participation to build a critical mass of data available through Dryad.

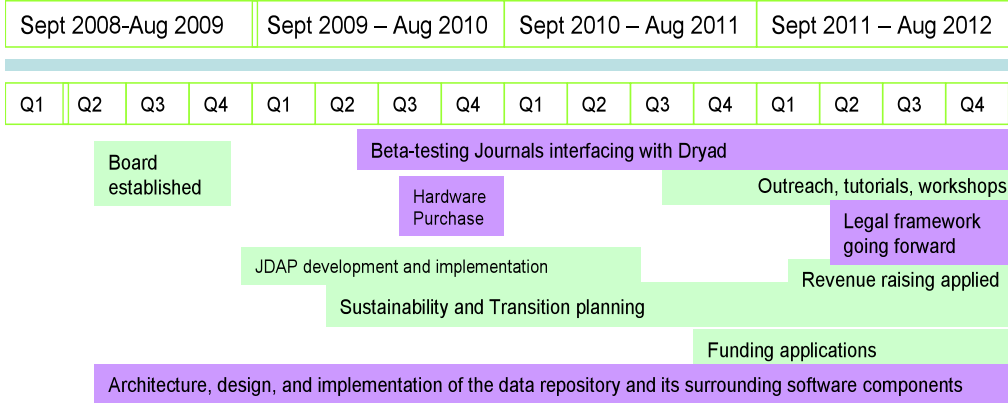
The views of funders on future or continued grant support for Dryad will need to be investigated further as recommended in section 5.6 above. .

Recommendation 12: A separate transition planning phase and a strategy for transition funding should be developed by the end of 2010 in order to facilitate implementation of a transition to full service in 2011/2012.

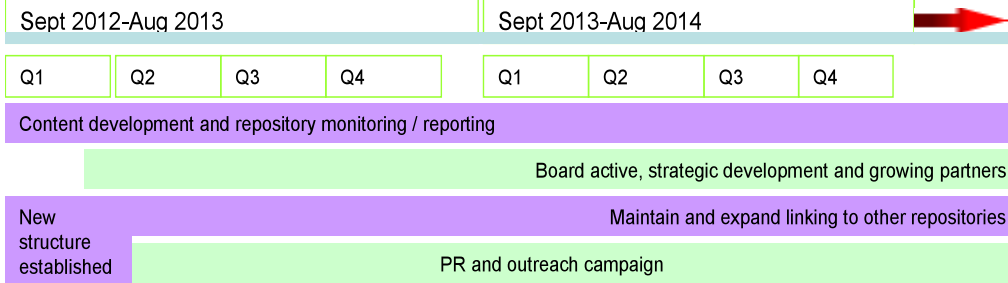
The following chart maps parallel activity and milestones associated with current development of the repository and the preparation to move forward the organization and to deliver a robust service.

Key Dryad milestones over a 6 year period

Dryad current project development phase



Post project Dryad



- Operational development**
- Capacity developments**



6. SCENARIOS FOR DRYAD

6.1. INTRODUCTION

The following three scenarios have been developed with input and guidance from the Dryad project team to frame discussion of potential revenue options that would provide financial sustainability. Each scenario focuses upon different weightings of the three methods identified in the table in Section 4.4 that have been assigned the highest priority because they align most closely with stakeholders and have the most potential to raise funds. These are: journal usage based upon a per-paper deposited price; annual subscriptions paid by journal; and a one-off journal joining fee. These scenarios should be seen as a core.

Recommendation 13: Other revenue options such as angel donors or sponsors should be pursued in addition but as these are unpredictable sources of revenue, any receipts from such donations should be used to support discrete activity, transitioning the service, or adding to operational reserves/creating endowment funds.

Scenario one outlines an approach that places the emphasis on revenue raised from per paper deposited fees, using annual subscription fees to compliment this main source.

Scenario two reverses these and places the emphasis on raising revenue from subscription fees and reducing the return from per paper deposit fees.

Scenario three looks at what happens when you give these two revenue sources equal weighting. Each of the three scenarios factors in the potential impact of revenue raised through one-off joining fees and discusses the nature of these.

“One-off joining fees” for the purposes of these scenarios would be mandatory and aimed at covering Dryad for journal integration and start-up costs.

The preliminary indication from the Dryad Board is that an acceptable price per paper deposited with supplementary data would be approximately \$30. However, to cover all the operating costs and to break even with this price will take considerable growth in Dryad only anticipated over a period of time, possibly a few years. It would be counter-productive at the start to set the price to cover all operating costs as this would result in a high price that journals would be unwilling or unable to meet. It would also act as a deterrent to any potential new journals joining Dryad. Setting the price at the accepted level will reflect the current partner journals limits and thereby facilitate their engagement. It will also be much less of a barrier to new journals joining Dryad. What this approach does result in is a shortfall in revenue as compared against operating costs. This funding gap will need to be addressed until sufficient growth has been achieved to fully cover operating costs. The following scenarios look at the effects of combining different revenue streams to raise enough funds to achieve a sustainable repository over time.

6.2. ASSUMPTIONS AND CAVEATS

- a) An operating cost baseline of \$320,000 is used, which is based upon the Eakin-Richards cost model assuming a volume of 10,000 papers annually and a corresponding per paper cost of \$32. Based on preliminary analysis by Todd Vision of average number of journal articles yielding supplementary data, it is estimated that this number of papers could be generated by 50-70 journals. This work has been funded separately from the development of the sustainability plan and we have incorporated it as directed. We cannot provide any warranty or accept any liability in relation to that work. It should be stressed that these figures therefore are estimates only and that they are used in the scenarios below to provide a picture of what could be expected. It is recommended that Dryad continues to review and build up its data on costs and develop detailed revenue models on which its business plan can be implemented.
- b) Inflation is not factored in.
- c) Per paper costs would need annual review and possible adjustment to any pricing scheme that is developed.
- d) For the purposes of the scenarios, annual subscription charges are assumed to be for the “average” journal. A pricing model would consider the merits of differential pricing for subscriptions based upon tiers or banding possibly based upon size of organisation or turnover. This is discussed further in section 1.5 below.
- e) The scenarios assume full member journals only.
- f) Dryad Curation level 2 is assumed as the standard.
- g) Early financial commitment by partner journals and the development and implementation of a pricing scheme will be necessary to form part of the transition funding case to NSF or other funders.
- h) Operating costs are not the same as “total” costs and alternative income will be sort to cover development etc.

6.3. SCENARIO ONE – PER PAPER CHARGES DOMINANT

In this scenario, Dryad aims to recover in the region of 80% of its operating costs via a use fee to journals based upon papers deposited with supplementary data. The remaining 20% is covered by an annual subscription to Dryad charged to journals. This places the emphasis upon a journals’ actual use of Dryad so a journal with few papers will pay significantly less than one with a great many. An annual subscription charge provides an upfront commitment from members and a predictable income to Dryad over 12 months. Keeping this relatively low reduces the hurdle for journals but will present a risk to Dryad in that it becomes dependant upon journals depositing papers to the level that will provide enough revenue to break even. If a journal experiences unexpected increases in the

number of papers being deposited in Dryad they will either end up exceeding their budget forecasts, or they may hold back on depositing papers for a time. Neither of these outcomes would be satisfactory. If use based fees are retrospective then a lead-in time is needed for budgeting i.e. potentially an 18-24 month period before a 12 month retrospective billing can be made. So, for an “average” journal producing annually in region of 175 papers with supplementary data to be deposited in Dryad the following could apply where the price per paper is set at \$25 and the annual subscription at \$1100:

No. of total journal members in Dryad	Total annual subscription revenue (\$)	Total annual per paper revenue (\$)	Total annual revenue (\$)	Estimated annual operating cost (\$)	Funding gap (\$)
60	66,000	262,500	328,500	319,400	+9,100
45	49,500	196,875	246,375	300,300	- 53,925
30	33,000	131,250	164,250	281,200	-116,950
15	16,500	65,625	82,125	272,100	- 189,975

Table 10: Possible revenue projections in Scenario One.

Note that operating costs will rise as the number of member journals and contributed papers increases. At these prices the aspiration of achieving 60 journal members and reaching anticipated levels of papers deposited in Dryad is crucial. It will be desirable to identify them in any transitional support application that is made to funders to ensure the funding gap can be dealt with and narrowed swiftly. Grants to address the revenue gap and support interim development of Dryad are unlikely unless there are already commitments from a critical mass of partners.

Recommendation 14: A realistic estimate of the possible increase in member journals, along with a strategy and implementation plan to achieve the necessary increase needs to be undertaken.

If a large journal joins Dryad with the expectation of depositing papers significantly above the average rate this will not only be positive in terms of increasing the repository but also financially. Those journals depositing most papers will pay more in total. This scenario relies upon maximising the number of papers deposited to break even.

Additional revenue from one-off joining fees can be used to off-set the costs associated with journal integration and set-up within Dryad in the first year of membership. Receipts are wholly dependant on the number of new journal members in any year. At this point the pace of the rate of growth is conjecture. However, assuming a three year horizon beyond the current project one might assume possible the following:

Current journal members	1 year beyond project	2 years beyond project	3 years beyond project

15	30	45	60
----	----	----	----

This would provide joining fees from 15 new member journals in each year for three years beyond the current project. Based upon the target of 60 journals a joining fee of \$5,000 would yield \$75,000 in each of the three years. During the transition period to full sustainability this revenue will help to address the gap in funding identified in the earlier table. Once a break even position is reached a reassessment could take place in consideration of the joining fee, possibly resulting in either a reduced subscription in the first year, or a joining fee could be used to off-set paper charges in the first year of membership. The above observations on the joining fee are also applicable to scenarios 2 and 3.

6.4. SCENARIO TWO – ANNUAL SUBSCRIPTIONS DOMINANT

In this scenario Dryad aims to recover in the region of 80% of its operating costs via an annual subscription charged to journals. The remaining 20% is covered by a use fee to journals based upon papers deposited with supplementary data. This places the emphasis upon a journals' committing the greater part of their payments to Dryad up front. It would offer a degree of budgetary certainty and predictability for journals. It would also lower the per paper price considerably thereby possibly encouraging greater use of Dryad by subscribing journals. For Dryad it would guarantee revenues covering the majority of its' operating costs and reduce the risk associated with fluctuating rates of deposit of papers by journals. Being less dependant upon retrospective billing for use fees (as mentioned in scenario one) Dryad should experience improved cash flow and easier management of its budgets.

So, for an "average" journal" producing annually in region of 175 papers with supplementary data to be deposited in Dryad the following could apply where the annual subscription is \$4500 and the price per paper is set at \$5:

No. of total journal members in Dryad	Total annual subscription revenue (\$)	Total annual per paper revenue (\$)	Total annual revenue (\$)	Estimated annual operating cost (\$)	Funding gap (\$)
60	270,000	52,500	322,500	319,400	+3,100
45	202,500	39,375	241,875	300,300	- 58,425
30	135,000	26,250	161,250	281,200	-119,950
15	67,500	13,125	80.625	272,100	- 191,475

Table 11: Possible revenue projections in Scenario Two.

In having the emphasis upon annual subscription revenue Dryad is very dependant on increasing the number of journal members and is less dependent on the number of papers deposited with Dryad. If journals experience a significant increase in the number of papers

they are depositing in Dryad then it is possibly to their advantage to have the emphasis upon the annual subscription and may create less of a disincentive for journals to encourage or require archiving if they do not have a steep per paper charge. To what extent this is true will be dependent on the pricing model (yet to be formulated) and also whether differential pricing is implemented and how. Section 6.5 below discusses differential pricing further.

The risk for Dryad with this scenario is if journal members do not increase as expected, thus returning fewer subscriptions, and at the same time if the number of papers per subscribing journal should increase significantly, yielding limited revenue due to the lower rate. This may result in an inability for Dryad to cover operating costs dependant on the balance between papers deposited and the number of member journals. In this eventuality a review of the pricing policy would need to be quickly expedited.

See Section 6.3 Scenario One for discussion of the addition of a joining fee.

6.5. SCENARIO THREE – PER PAPER CHARGES AND SUBSCRIPTIONS BALANCED

In this scenario Dryad aims to recover 50% of its operating costs via an annual subscription to Dryad charged to journals and 50% covered by a use fee to journals based upon papers deposited with supplementary data. This places an equal emphasis upon both subscription and use fees.

So, for an “average” journal producing annually in region of 175 papers with supplementary data to be deposited in Dryad the following could apply where the annual subscription is \$2,750 and the price per paper is set at \$15:

No. of total journal members in Dryad	Total annual subscription revenue (\$)	Total annual per paper revenue (\$)	Total annual revenue (\$)	Estimated annual operating cost (\$)	Funding gap (\$)
60	165,000	157,500	322,500	319,400	+3,100
45	123,750	118,125	241,875	300,300	- 58,425
30	82,500	78,750	161,250	281,200	-119,950
15	41,250	39,375	80,625	272,100	-191,475

Table 12: possible revenue projections in Scenario 3.

With an even income from both sets of fees there will be a need for Dryad to see both the increase in subscribing journals and the growth in papers deposited to ensure reaching a break even point.

The risks associated with both scenarios one and two can to some extent apply here. However, the advantages that could result from the emphasis being upon either subscriptions or use fees are unlikely to materialise when equal emphasis is placed upon

them. By placing equal emphasis on both subscription and per paper fees each is at a level that may be considered more acceptable, so that neither one individually presents a barrier. See Section 6.3 Scenario One for discussion of the addition of a joining fee.

6.6. DIFFERENTIAL PRICING

Each of the scenarios above assumes an “average” journal with all journals paying the same subscription, joining fee and use fee. In order to provide incentives and to recognise differences in journals that join Dryad, differential pricing could be considered. It does introduce a level of complexity but it is an approach used successfully by other services to encourage take-up and to respond to what might be perceived as fair by potential members.

The approaches below are examples of what might be considered when developing the pricing policy for Dryad. A mix of options is possible but the pricing policy should be clear and should be perceived as being “fair” by members and other stakeholders.

6.7. BANDS OR TIERS FOR SUBSCRIPTIONS

The differences in participating journals can be acknowledged through the setting of subscription bands or tiers. Organisations such as Portico use total revenues as a measure, others might use the size of an organisation measured by the number of full time equivalent employees. For example, Portico publishes³ nine tiers based on a publisher’s total journals revenues (print and electronic subscriptions, licensing, and advertising), with archive contributions to Portico ranging from \$250 for the smallest to \$75,000 for the largest. CrossRef also uses total publishing revenue on which to base the tiers for its annual fees. Again, this is nine tiers with fees ranging from \$250 to \$50,000.⁴

A similar approach could be used by Dryad to categorise publishers into different tiers dependant upon total publishing revenue. The Portico bands are more graduated at the lower end thus taking into account more difference at the smaller publisher level. This may be a fit for Dryad given the anticipated profile of future members. However, before rates are set it will be important, as part of the work to estimate and identify potential member journals, (as mentioned in the recommendation in section 1.3), that an analysis of these journals is undertaken to broadly indicate their distribution across tiers. This information can then be used as an input into developing a pricing policy that will have at least the potential to return sufficient revenue from subscriptions to contribute to breaking even.

Recommendation 15: An analysis of potential Dryad members is undertaken to assess how they might fit within tiered fees and use this as input into the development of a pricing policy.

³ <http://www.portico.org/digital-preservation/join-portico/for-publishers/>

⁴ http://www.crossref.org/02publishers/20pub_fees.html#background

A variant on this that Dryad might consider is to offer a discounted subscription rate for those journals where two or more from the same Publisher or Society join Dryad.

6.8. DISTINGUISHING TYPES OF MEMBER

As Dryad moves beyond the current project it is suggested that there is a review of the type of membership available and how this might be differentiated. At present Dryad has partners (those who adopt the Joint Data Archiving Policy), and associates who have expressed an interest but not yet signed up to the Joint Data Archiving Policy. Dryad currently will also accept the deposit of data from authors as long as the data files are associated with a publication, even from non-partner journals. Beyond the current project phase it is suggested that there is just one classification of “partner” or “member”. A review of the collection policy around the acceptance of papers and data from non-member journals should take place prior to the end of current NSF funding. This would require an adjustment of the collections policy to align it with an approach to pricing that would reflect the costs of non-members depositing and the fact that they would not be contributing an annual subscription.

Recommendation 16: Classify members and non-members, and review the policy on the acceptance of papers from non-members prior to the end of the current funding. Adjust the Collections Policy to align with an approach to pricing that would reflect the costs of non-members depositing and the fact that they would not be contributing an annual subscription.

6.9. FOUNDING MEMBER INCENTIVES

As a good-will gesture and to acknowledge the commitment the founding members have made to Dryad the waiving of the one-off joining fee may be considered.

The approach taken by Portico was that participants joining in their first 2 years of operation (2005-07), were given a 10% discount on participant fees for 5 years.

6.10. OTHER INCENTIVES

Where journals are able to make multi-year commitments for subscriptions & per-paper charges during the phase when it will affect the repository's success in achieving transitional funding, the Dryad pricing policy should provide incentives to encourage journals to make this decision. This may take the form of fixing prices over either a 3 or 5 year period. There is of course a risk to Dryad in this approach but it also brings the advantage of being able to plan with some certainty over the period. It will also demonstrate to potential funders the faith and commitment members have in the repository.

Usage based pricing may be used as an incentive to encourage high deposit rates by subscribing publishers, where a discount is offered on per paper price after certain thresholds are reached. For example, using scenario one, if a journal deposits 320 papers in a year the following might apply:

Numbers of papers deposited	Price per paper	Revenue for Dryad
1-200	\$25	\$5000
201-300	\$20	\$1980
>300	\$15	\$285
	Total =	\$7,265

Table 13: Possible revenue projections from usage based pricing.

DRAFT

7. RISKS

This is an initial risk register for the sustainability of Dryad for your review and further development.

Recommendation 17: The risk register is developed and regularly reviewed by the Board.

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Problems in building a critical mass of data within the repository	3	5	15	Mitigated by populating with targeted datasets and links to other repositories.
No growth in numbers of partner journals	1 (3)	5	5 (15)	Assumes marketing effort is contained in minimum viable staffing (or higher risk applies). Demonstrate benefit of partnership via existing partners.
Breakdown of partnership	2	3	6	Maintain good communication levels and shared goals.
Technical problems with delivering core functionality and stable platform	2	5	10	Thorough analysis of requirements, options and capabilities throughout development. Working with beta test Journals and “handshaking” repositories to identify and iron out any problems.
Inability to meet expectations of stakeholders/partners	2	5	10	Manage expectations – be realistic about what can be achieved. Identify clear milestones.
Short fall in revenues	3	5	15	Review early and anticipate where renewed efforts need to be applied.
Underestimating total costs leaving Dryad in a deficit	2	4	8	Review regularly costs, using full cost model.
Authors resist depositing supplementary data	3	5	15	Clear mandate from Journals. Simplify deposit process. Data citation. Advocacy to authors on benefits to themselves.

APPENDIX 1: ADDITIONAL TABLES

	Archaeological Data Service (ADS)	KCL Centre for e-research
Repository type	National Specialist Subject Repository	Institutional research data repository and research support centre
Brief description	ADS makes data sets available to support research and teaching. It provides a catalogue of its holdings and a gateway to other collections. It also acts as the archive for supplementary data associated with articles published in the e-journal Internet Archaeology. Established in 1996.	The Centre for e-Research is maintaining the Arts Humanities Data Service Repository which stores editions of digital collections deposited in addition to its current institutional focus.
Organisation	Based at the University of York. Founded by a consortium including the Council for British Archaeology and several universities. Guided by an Advisory Committee consisting of representatives from all sectors of the discipline. Funded by Arts & Humanities Research Council, assisted through revenues obtained via a charging policy. ^{xiii}	KCL are developing central repositories to work with departmental facilities and discussing federated local data repositories for research data preservation combining services and skills from central and departmental repositories.
Collection type	The types of data involved include: text reports, databases (related to excavated contexts or artifacts, for example), images (including aerial photographs, remote sensing imagery, photographs of sites, features and artifacts), digitized maps and plans, numerical datasets related to topographical and sub-surface surveys and other locational data, as well as reconstruction drawings. 55 new collections released in year 2008/09.	Staffing at launch affords the ingest of about 30 collections each year, with an assumption that 10 will be simple collections (images or texts created to standard formats and metadata), and 20 complex collections, comprising more complex formats (such as 3D visualisation materials) and multi media. Once this limit is reached, a new Collections Officer (CO) is required to cope with the additional work.
Preservation	Collection Policy identifies preferred file formats for deposit to minimize future costs in archiving. Preservation copies of all data synchronized off site into a deep storage facility within the UK Data Archive at the University of Essex.	

Table 1: External Comparators – The Archaeology Data Service (ADS) University of York and Centre for e-Research King’s College London.

ADS Staff 1996 at launch	KCL Centre for e-research at launch
---------------------------------	--

ADS Staff 1996 at launch	KCL Centre for e-research at launch
Part-time Director: responsibility for the day-to-day running and project finances, and lead role in liaison with external organisations. (0.1 FTE)	Director of Centre: time allocated to management not specified – costed by KCL as separate budget line (unknown)
Data Coordinator: responsible for approaching grant holders, negotiating with depositors and acquiring access to collections. Manages the project Web site and is editor of ADS-News, the project newsletter. Responsible for monitoring and dealing with enquiries to the Service. Represents the ADS on a number of committees and (with the Director) at liaison meetings. (1FTE)	Archives Manager: co-ordinate activities (1 FTE)
Collections Manager: responsible for the accessioning, mounting, and cataloguing of the ADS collections. Responsible for the development and maintenance of the metadata indices and also (with the Data Coordinator) for liaison over the development of metadata indices to external holdings. Concerned with the validation of new collections and the preservation and migration of all ADS holdings. In liaison with staff of York University Computing Service, the Collections Manager will take charge of the development and maintenance of the ADS computer systems. (1 FTE)	Collections Officer: develop and implement appropriate workflow and standards for the curation and preservation of research data (1 FTE)
Part-time Administrator: responsible for maintaining the ADS mailing list, general project administration and maintaining the ADS accounts. (0.5 FTE)	Systems Administrator to install and manage hardware and software (0.5 FTE)
	Administrative Officer (0.5 FTE)
Total 2.6 FTEs	Total 3 FTEs

Table 2: External Comparators – staffing at launch for the Archaeology Data Service (ADS) and Centre for e-Research.

ENDNOTES

- ⁱ Serco, (July 2008), *UK Research Data Service Feasibility Study, Interim report*, <http://www.ukrds.ac.uk/resources/download/id/17>
- ⁱⁱ Ashley, K. (1999). *Digital Archive Costs: Facts and Fallacies*. *DLM Forum '99*. <http://ec.europa.eu/archives/ISPO/dlm/>
- ⁱⁱⁱ Barton, Mary R.; Waters, Margaret M (2004), *Creating an Institutional Repository: LEADIRS Workbook* URI: <http://hdl.handle.net/1721.1/26698>
- ^{iv} Beagrie, N., Chruszcz, J. and Lavoie, B., (Joint Information Systems Committee 2008), *Keeping Research Data Safe: a cost model and guidance for UK Universities*, <http://www.jisc.ac.uk/publications/publications/keepingresearchdatasafe.aspx>
- ^v Beagrie, N., Lavoie, B. and Woollard, M., (July 2009), *Keeping Research Data Safe2 ("KRDS2")* http://www.beagrie.com/KRDS2_ActivityModelUpdates.pdf
- ^{vi} Case study observations from reports previously undertaken including "Keeping Research Data safe"[*ibid*]
- ^{vii} Informed by: *Keeping Research Data Safe: a cost model and guidance for UK Universities*, [*ibid*]; *UK Research Data Service Feasibility Study (UKRDS): Project Interim Report*; [*ibid*] and *Motivating Online Publication (BioScience)* May 2009 / Vol. 59 No. 5.
- ^{viii} Open Society Institute, (July 2003) *Guide to Business Planning for Launching a New Open Access Journal (Edition 2)*, http://www.soros.org/openaccess/oajguides/business_planning.pdf
- ^{ix} <http://www.ithaka.org/> . JSTOR has over 5808 participating institutions world-wide from which they levy per collection a one off Archive Capital Fee and an Annual Access Fee. These differ according to the JSTOR classification of the participating organization.
- ^x http://www.jisc-collections.ac.uk/about_collections/coll_memandarts?keywords=Memorandum+and+Articles+of+Association. It is a company limited by guarantee and not having a share capital and bound by its memorandum of association.
- ^{xi} It is supported through grant funding as a JISC service see <http://www.jisc.ac.uk/aboutus.aspx>.
- ^{xii} <http://www.ncbi.nlm.nih.gov/pmc/about/pmci.html>
- ^{xiii} Archeology Data Service, Charging Policy 4th Edition, <http://ads.ahds.ac.uk/project/userinfo/charging.html>