

Dryad Cost Model Report

Submitted in support of Dryad Cost Model, Version 1.0

Model created by Lorraine L. Eakin-Richards, December 12, 2009

Final Report submitted, February 8, 2010

Introduction and Goals for the Cost Model

Dryad desires to answer a primary question: “What will be the per article costs of the repository captured in the operational budget?” The cost model described here is geared to provide an answer to that question.

Dryad personnel are beginning to build a revenue model and are most concerned with ensuring that they develop revenue capture techniques that will support the ongoing sustainability of Dryad. This requires ensuring that those costs which are not likely to be supported via grant funds be supported by a sustainable mechanism for generating revenue.

Some revenue models will be more easily implemented at low per article costs than at higher costs. For example, if one wishes to charge a per article deposit fee, it will be easier to negotiate a fee that covers all operational costs if these costs are relatively low than if they are high. Thus, a related question follows from the primary question noted above: “What ingest levels are required in order to achieve various per article costs?” For the purposes of this model, the most desirable per article cost is assumed to be \$20. The Dryad cost model assesses what levels of ingest would be required to achieve per article costs of \$20, \$50, \$75, \$100, and \$150. In fact, at the lowest service level, all costs except \$20 are almost immediately achievable. At the highest Service Level, \$20 per article is virtually unachievable for any realistically achievable ingest level.

What This Model Is and What It Isn't

In general, it is desirable to create a “full costing” model when attempting to assess repository costs. Such a model will include all cost elements, whether they are actually covered via direct operating costs (and therefore written into the operating budget) or not. For example, for many repositories sponsored by an institution, various utilities, facilities charges, and frequently, servers and other networking infrastructure, are provided as cost share and are not included in the repository operating budget. Not infrequently, even some typical operating costs (such as personnel time) may be provided as part of institutional cost share. To understand the full costs of the repository, one would include all of these costs, whether they are direct or indirect.

Ideally, one includes both direct and indirect costs in order to mitigate risk: any time a repository depends upon cost share elements, the potential exists for institutional support to “dry up” in lean economic times or when the institution’s priorities change. If the repository monitors full costs, it can assess the degree of risk and when appropriate, develop contingency plans in the event that some of the cost share items suddenly need to be covered by the operational budget or additional grant funds. In short, keeping track of all costs enables better planning and management of all sources of funding.

That said, this version of the Dryad model is not a fully costed model. *The reason for this relates to the immediate goal of the model. It was developed on a very short time line and the key question it attempts to answer relates very specifically to the costs that need to be covered by the operational budget. The model*

attempts to understand the level of those operational commitments. Thus, at this point, capturing the actual outlay of funds is most appropriate.

For Dryad to manage and plan for long-term costs, however, it will be desirable to begin to delineate all potential costs. This will allow it not only to assess potential funding risks more accurately and to create a risk plan for them but also to allow Dryad to create a more accurate predictive cost model.

Requirements to Upgrade to Full Costing Model

- *Assess and add all hardware, software and networking infrastructure costs over time. Describe which costs are currently (and projected to be) cost share items and which must be captured via revenue capture techniques. (Currently only the latter costs are included in this model.)*
- *Discount the costs over time (i.e., calculate the net present value of the stream of costs over time) to make intertemporal comparisons valid for long-term strategic planning.*
 - *Determine the appropriate discount rate to use.*
 - *Determine the desired time frame for discounting (e.g., this is often from 5 to 15 years, with 8-10 years the typical discounting period for capitally intensive projects).*
- *Important missing elements:*
 - *Software licensing fees for format migration. For repositories that do not have strict format requirements on ingest, this can be a significant cost. Since Dryad is accepting all formats and currently appears to be planning to support them, licensing fees to maintain the accessibility of formats over the long run could become a significant cost. This should be assessed and included.*
 - *Software fees for dealing with naming authorities and other possible cataloging operations have not been included here. Further research to determine likely levels is necessary and should be assessed and included.*
 - *Hardware and system administration service fees.*

Structure of Model

The Microsoft Excel cost model contains a number of worksheets geared to answer the two questions above. The “default” portions of the model are associated with particular “default” assumptions and contain the following worksheets:

- *“Costs With Default Assumptions”*

- *This worksheet is a roll-up worksheet. It allows the viewer to select one of the three potential service levels via a drop-down list and to see the overall and per article costs associated with each of these three service levels.*
- *This sheet has a place holder for calculating the net present value of the discounted stream of costs over a five year time period. In this (the first) version of the Dryad cost model this functionality is disabled. If Dryad chooses to engage in full economic costing, this functionality can very easily be re-enabled.*
- **“Ongoing Costs”**
 - *This worksheet has a detailed categorization of all cost elements currently available to Dryad. It can be updated to include additional cost elements as needed, by inserting additional rows in the appropriate higher level categories.*
 - *Currently, this sheet automatically extracts some values calculated on other worksheets from those sheets, for example, all personnel salary and benefit costs. These costs are directly dependent upon the size and growth assumptions for the repository. In order to keep track of the fact that this area will grow as the repository grows, and to allow for changing salary bases, benefit loads, and annual salary growth, I created a separate “Personnel Assumptions” worksheet. Salary expenditures are populated into “Ongoing Costs” from this sheet.*
 - *This sheet does include a number of costs that can be directly inputted, however. These direct input line items are recognized by virtue of being highlighted in yellow. Each sheet that contains potential input cells has the relevant cell(s) highlighted in yellow and a note at the top that says, “DATA IS ENTERED IN YELLOW AREAS ONLY.”*
 - *Items that are currently manually populated in this sheet have had an annual inflation adjustment added. These adjustments are noted in the Assumptions section below.*
- **“Size and Growth Assumptions”**
 - ***This is a key cost driver for the model as a whole.*** *The model is built such that the required curatorial FTE’s are dependent upon two things: the annual ingest level (i.e., total number of data packages/articles ingested per year) and the time it takes to ingest a data package. Three service levels are included. Assumptions are given in the Assumptions section below.*
 - *As the default assumptions about the likely size of the repository change over time, the yellow cells on this sheet can be updated, and the required curatorial FTE’s will be automatically updated to reflect these new ingest volumes. Likewise, all costs will be automatically updated and propagated to the “Costs With Default Assumptions” worksheet, providing updated total and per article costs.*
- **“Personnel Assumptions”**
 - *This sheet contains a list of all projected repository personnel types. It also shows the hourly salary rate, the benefit load utilized in the model, and the projected FTEs over time for each personnel type.*
 - *The curatorial FTEs are shown in green highlighted cells on this sheet, since curatorial requirements are dependent upon and calculated from the ingest level.*
 - *Currently, one professional curator (or a portion of one FTE of a professional curator) is assumed as the baseline. As curator FTEs increase beyond one, all additional curation personnel are student personnel. Thus, there are two curatorial line items, to reflect the different skill category and salary rates associated with them.*

- *The other personnel types are included as part of the overall repository staffing plan and volumes must be input into the FTE requirement (yellow highlighted) cells.*
- *“Curation Assumptions”*
 - *This sheet provides information about processing times for various ingest levels. It takes the default ingest level given in the “Size and Growth Assumptions” sheet and applies to it the number of minutes required for each service level. It calculates the baseline minutes required to curate that volume for each of the three service levels and calculates the raw FTE requirement for that amount of processing. It then adds on a “FTE Add-On,” which is a flat value added to the processing time to reflect the fact that curators do more than just process packages one after the other. There is time between each package. There are sick and vacation days. They may have other duties besides performing curatorial functions. Currently, I have added ¼ FTE to reflect these times. I believe this to be quite conservative; most staffing plans add about .18 FTE to reflect sick and vacation days alone. If the curator engages in other activities such as thesaurus generation or finding aids, Dryad will want to account for this additional time by increasing the FTE Add-On value.*
 - *The final values on the far right of this sheet reflect the total anticipated curatorial FTEs with the FTE Add-On included. This is the number of FTEs that is used on the “Personnel Assumptions” sheet.*

In addition to the “Default” portions of the model, there is a sheet that is specifically designed to calculate the required ingest level for given per article costs. Ingest levels and curatorial FTE requirements are (mathematically) co-determined. Thus, in the default cost calculation sheets, one can set the desired ingest level as a management strategy and see how many curators are required in order to provide that ingest level for each level of service.

However, when answering the question “What ingest level is required to achieve a per article cost of \$20?” it is possible that a variety of levels of curatorial FTE could be used to do this. I have used the minimum level of FTEs required to achieve each service level and calculated the minimum ingest level required, given the minimum FTE level, in order to reach various per article costs. Worksheets associated with this are:

- *“Ingest Needs With FTEs Given”*
 - *Currently this sheet assumes from ¼ FTE up to 100 FTEs, incremented by ¼ FTE increases. It then shows for each service level the possible ingest level, the curatorial costs, the total costs, and the per article costs associated with each FTE level.*
 - *This sheet currently only shows results for Year 1. This could be upgraded in future versions of the model to reflect all years and to include a continuous increase in FTEs (rather than stepping up at ¼ FTE increments), but time constraints did not allow this for this version of the model.*
- *“Graph – Per Art Costs Svc Lvl 1”*
 - *This graphs the results of “Ingest Needs With FTEs Given” for Service Level 1. It shows the relationship between per article costs and ingest level per year.*
- *“Graph – Per Art Costs Svc Lvl 2”*
 - *This graphs the results of “Ingest Needs With FTEs Given” for Service Level 2. It shows the relationship between per article costs and ingest level per year.*
- *“Graph – Per Art Costs Svc Lvl 1”*

- *This graphs the results of “Ingest Needs With FTEs Given” for Service Level 2. It shows the relationship between per article costs and ingest level per year.*

Other sheets exist in the model, but are used either for calculations within the model (e.g., “Formula Data”) or to test the formulas used throughout the model (e.g., “Testing Assumptions & Formulas”). They exist purely for development & testing purposes, or to allow easier automated operation within the model.

Assumptions

General Assumptions for Model Development

- *The model is currently built using a 5-year timeframe. This reflects an assumed 5-year hardware refresh cycle.*
- *This timeframe should be increased with future versions. With a 5-year hardware refresh, the fifth or sixth year will include potentially sizable migration costs and hardware, software, and infrastructure expenditures that need to be tracked. Over time, the calculation of net present value will be more valuable if a longer time frame can be introduced to the model, say 10-years. Because of the preliminary nature of this version and because many of the key cost driving decisions are still being made, it would not have been particularly useful to include such a long time frame at this point.*
- *There is a place holder for calculating the discounted stream of costs over time (i.e., “net present value”). When the model is more fully costed (i.e., including both direct and indirect costs) this function will become more useful and can be enabled.*

Size and Growth Assumptions

- *The initial annual ingest levels are given on “Size and Growth Assumptions.” Currently they are assumed in the mid-thousands:*
 - *Year 1: 5,000 articles*
 - *Year 2: 5,500 articles*
 - *Year 3: 6,000 articles*
 - *Year 4: 6,500 articles*
 - *Year 5: 7,000 articles*
- *These assumptions can be changed at will to determine impacts on total and per article costs. The impacts will be seen on the front sheet, “Costs With Default Assumptions.” You will also be able to see the particular impacts on curatorial needs (“Curation Assumptions” sheet), overall personnel costs (“Personnel Assumptions” sheet), and ongoing costs (“Ongoing costs”) sheet.*

- *It is recommended that serious consideration be made regarding the potential impact on non-curatorial staff with repository growth.*
 - *Although not immediately intuitive, additional communication activities will be required, thereby having a likely impact on the repository manager, the communications specialist, and the administrative assistant when the repository grows.*
 - *Repository growth will also lead to increased expectations by users and funders regarding service offerings. The more successful Dryad becomes, the greater the demands for additional access services. This will increase necessary developer FTEs. **Currently it is assumed, however, that these demands will be able to be captured via grant funding for new development, so they are not reflected in the model.***

Personnel, Salary and Benefit Assumptions

- *Currently assumptions and rationale are:*
 - *Repository manager:*
 - *Because of the need for ongoing grant writing, communications with stakeholders, etc., ¾ of an FTE has been assigned for the first four (4) years. A full FTE is assumed in Year 5. This increase is due to continually increasing communication and repository coordination.*
 - *Administrative assistant:*
 - *A flat ¼ FTE has been assumed across all five (5) years. This person will be responsible for managing memos, sending emails, appointment management, etc.*
 - *Curators and Curator Lead:*
 - *Currently this is directly calculated, based upon the default annual ingest levels assumed in the “Size and Growth Assumption” spreadsheets. The Service Level processing times are currently given as*
 - *Service Level 1: 5 minutes*
 - *Service Level 2: 20 minutes*
 - *Service Level 3: 140 minutes*
 - ***The Service Level Processing times are a key cost driver for the model as a whole, since they determine the curation FTEs needed by the repository.***

- Each of these service level processing times can be updated on the “Curation Assumptions” spreadsheet, and the impacts to necessary curatorial FTEs, total costs, and per article costs are automatically generated within the model.
 - Currently, the initial curator is assumed to be a professional curator and the curation lead on the project. Once a full FTE has been committed, however, all additional curatorial personnel are assumed to be student curators and are costed at the lower student rate.
 - Currently at service level 1, it appears that for very low ingest levels there are negative curators on the staff. This is a mathematical function of the way the model is calculating salaries and should be ignored – **negative curator FTEs should be assumed to represent the fact that there are NO student FTEs and that less than a full curator lead FTE is used.** This can be seen on “Personnel Assumptions” under the FTE requirements category, where the cells on row 17 imply negative curator FTEs and less than 1 full curator lead FTE. **The curator costs on other sheets are zero and are appropriately calculated for the curator lead in those situations.**
 - Developer:
 - Currently, ¼ FTE developer is assumed across all five (5) years. This reflects the probable portion of developer costs that will need to be covered by the Dryad operating budget. It is likely that MORE than ¼ FTE will be needed to support Dryad growth and maintenance, but the model assumes that new development will be covered by grants and not by the operating budget, so only ¼ FTE is included here. When a full cost model is developed in future versions, Dryad will want to include all programmer time and separate the direct (that is, operationally covered) from the indirect costs. **(Note that this is not the way in which the terms “indirect costs” and “direct costs” are typically used in financial modeling literature. These terms are used here only to reflect whether the costs are captured directly in the operating budget or whether they are part of cost share.)**
 - Communications Specialist:
 - Because this person will manage correspondence with societies, journals, and depositors, as well as contracts and intellectual property rights (IPR) issues, the initial needs are relatively high. ½ FTE is assumed for the first four (4) years, and a full FTE in year 5.
 - As the repository grows, the requirement for this person’s time will also grow and it is not inconceivable that it will require more than one FTE eventually. In such a case specialization of duties will probably be most cost effective (e.g., an IPR person and an outreach/growth specialist).
- Other assumptions:
 - Salary Levels:

- *These are given on the “Personnel Assumptions” sheet and were provided by Todd Vision as initial estimates.*
- *Benefits Load:*
 - *This reflects the additional personnel costs over and above salary to cover items such as health insurance, vacations, internal training, etc.*
 - *With the exception of the student curators, for which load is not calculated, the current benefit load assumed is 20%.*
 - *Student curators, as part time and probably temporary personnel, have no load associated with them. If the repository includes items such as health insurance or tuition remission, however, this will need to be included as load.*

FTE Assumptions to Calculate Required Ingest Levels

- *To calculate ingest levels needed to achieve particular per article costs, curator FTEs had to be treated (quantitatively) as an “independent” variable. Curator FTEs were incremented by ¼ FTE from ¼ FTE (of a professional curator) up to 100 FTEs (i.e., 1 professional curator and 99 students). Although I recognize that this latter level of curation is highly unrealistic for the state of development of Dryad at any time in the foreseeable future, FTE levels were run up to this value simply for speculative purposes – to see if particular service levels, for example, could ever achieve a \$20 per article rate.*
- *The time to complete each of the three service levels is given in the “Curation Assumptions” worksheet and discussed above in the Curators and Curator Lead assumptions. These time commitments were provided by Dryad employee Ryan Scherle.*

Hardware, Software, and Networking Assumptions

- *Currently, no costs are provided for this category. The model should include the fee for service charges for the system administration and networking provided by North Carolina State University (NCSSU), but these numbers were not available to me. They should be added when determined.*
- *Virtually all other costs within this category are currently cost share and for reasons given in the section “What This Model Is and What It Isn’t,” they were not included in this version. A full costing model needs to include them.*
- *Important missing elements:*
 - ***Software licensing fees for format migration.*** *For repositories that do not have strict format requirements on ingest, this can be a significant cost. Since Dryad is accepting all formats and currently appears to be planning to support them, licensing fees to maintain the accessibility of formats over the long run could become a significant cost. **This should be assessed and included.***

- *Software fees for dealing with naming authorities and other possible cataloging operations have not been included here. Further research to determine likely levels is necessary and should be assessed and included.*
- **Hardware and system administration service fees.** *This is discussed in the first bullet above. It should be assessed and included in the model.*

Detailed Enumeration of Important Cost Assumptions

- *Advisory Board travel and preparation costs are relatively high – in the range of \$20,000 to \$30,000 per year. The model currently assumed \$25,000 in the “Ongoing Costs” worksheet. A 2% annual inflation adjustment has been included here.*
- *All salaries are adjusted for inflation annually by 3%.*
- *All personnel have a flat 20% salary load, except the student curators for whom no load is added*
- *No costs have been included in the line item enumeration for “Research and Development” or for “User Documentation and Training.” The former is assumed to be covered by grant funding and for this version of the cost model has not been included. The latter is assumed to be covered by the Communication Specialist FTEs.*
- *Although in the future it may be desirable to include the Communication Specialist’s travel as a separate line item in the model (and a placeholder exists on the “Ongoing Costs” worksheet for this), all travel costs are currently included under the “Miscellaneous” category’s travel expense line item.*
- *No advertising costs are assumed. At this point, outreach and promotion of the repository is assumed to be provided via the project manager and communication specialist FTEs.*
- *No vendor and consulting fees are included at this point because they are assumed to be covered by grant funding. A full cost model would include them, and would attribute them as direct or indirect.*
- *Personnel training is assumed to be zero in this version, since it is likely to be covered by grant funding, per Todd Vision.*
- *Communication costs reflect miscellaneous telecommunications costs and a flat \$1,000 with 2% inflationary growth is assumed.*
- *Miscellaneous supplies (e.g., ink, printer cartridges, paper, etc.) is set at a very low \$500 per year initially, and has a 10% inflationary growth associated with it.*
- *No insurance is assumed.*
- *Because the nature of these estimates are already contingent upon managerial strategy and decisions, no additional contingency estimate has been included. As Dryad develops a strong strategy and policy, and thereby gains a better grasp on likely costs, a contingency factor can be included. This often ranges from*

15% to 25% of total costs, or it can be associated with particular line items that tend to be variable and difficult to predict.

Results

With the current assumptions regarding size, annual growth, and baseline costs, the three service levels achieve the following total and per article costs.

Service Level 1:

Service Level	1					
Data Package Estimates:	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL
Number of packages	5,000	5,500	6,000	6,500	7,000	30,000
Number of articles	5,000	5,500	6,000	6,500	7,000	30,000
Ongoing Costs:	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL
Repository Management	\$97,000	\$99,660	\$102,395	\$105,207	\$135,110	\$539,371
Administrative Support	\$7,488	\$7,713	\$7,944	\$8,182	\$8,428	\$39,755
Curation	\$21,615	\$23,254	\$24,972	\$26,772	\$28,657	\$125,271
Storage and Hardware	\$0	\$0	\$0	\$0	\$0	\$0
Infrastructure/Facilities	\$0	\$0	\$0	\$0	\$0	\$0
Research and Development	\$0	\$0	\$0	\$0	\$0	\$0
Repository Maintenance	\$24,000	\$24,720	\$25,462	\$26,225	\$27,012	\$127,419
Outreach and Promotion	\$48,000	\$49,440	\$50,923	\$52,451	\$81,037	\$281,851
User Documentation and Training	\$0	\$0	\$0	\$0	\$0	\$0
Outsourcing	\$0	\$0	\$0	\$0	\$0	\$0
Miscellaneous	\$4,000	\$4,120	\$4,246	\$4,380	\$4,521	\$21,267
Annual Costs:	\$202,103	\$208,907	\$215,942	\$223,217	\$284,764	\$1,134,933
Per Article Costs:	\$40	\$38	\$36	\$34	\$41	\$38

Service Level 2:

Service Level	2					
Data Package Estimates:						
	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL
Number of packages	5,000	5,500	6,000	6,500	7,000	30,000
Number of articles	5,000	5,500	6,000	6,500	7,000	30,000
Ongoing Costs:						
	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL
Repository Management	\$97,000	\$99,660	\$102,395	\$105,207	\$135,110	\$539,371
Administrative Support	\$7,488	\$7,713	\$7,944	\$8,182	\$8,428	\$39,755
Curation	\$48,000	\$49,440	\$50,923	\$52,451	\$54,024	\$254,839
Storage and Hardware	\$0	\$0	\$0	\$0	\$0	\$0
Infrastructure/Facilities	\$0	\$0	\$0	\$0	\$0	\$0
Research and Development	\$0	\$0	\$0	\$0	\$0	\$0
Repository Maintenance	\$24,000	\$24,720	\$25,462	\$26,225	\$27,012	\$127,419
Outreach and Promotion	\$48,000	\$49,440	\$50,923	\$52,451	\$81,037	\$281,851
User Documentation and Training	\$0	\$0	\$0	\$0	\$0	\$0
Outsourcing	\$0	\$0	\$0	\$0	\$0	\$0
Miscellaneous	\$4,000	\$4,120	\$4,246	\$4,380	\$4,521	\$21,267
Annual Costs:	\$228,488	\$235,093	\$241,893	\$248,896	\$310,131	\$1,264,501
Per Article Costs:	\$46	\$43	\$40	\$38	\$44	\$42

Service Level 3:

Service Level	3					
Data Package Estimates:						
	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL
Number of packages	5,000	5,500	6,000	6,500	7,000	30,000
Number of articles	5,000	5,500	6,000	6,500	7,000	30,000
Ongoing Costs:						
	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL
Repository Management	\$97,000	\$99,660	\$102,395	\$105,207	\$135,110	\$539,371
Administrative Support	\$7,488	\$7,713	\$7,944	\$8,182	\$8,428	\$39,755
Curation	\$193,769	\$216,914	\$241,273	\$266,899	\$293,844	\$1,212,699
Storage and Hardware	\$0	\$0	\$0	\$0	\$0	\$0
Infrastructure/Facilities	\$0	\$0	\$0	\$0	\$0	\$0
Research and Development	\$0	\$0	\$0	\$0	\$0	\$0
Repository Maintenance	\$24,000	\$24,720	\$25,462	\$26,225	\$27,012	\$127,419
Outreach and Promotion	\$48,000	\$49,440	\$50,923	\$52,451	\$81,037	\$281,851
User Documentation and Training	\$0	\$0	\$0	\$0	\$0	\$0
Outsourcing	\$0	\$0	\$0	\$0	\$0	\$0
Miscellaneous	\$4,000	\$4,120	\$4,246	\$4,380	\$4,521	\$21,267
Annual Costs:	\$374,257	\$402,567	\$432,243	\$463,344	\$549,951	\$2,222,362
Per Article Costs:	\$75	\$73	\$72	\$71	\$79	\$74

Additional Findings – Required Ingest Levels

Service Level 1 – 5 Minute or 15 Minute Processing Times

Currently the model assumes that Service Level 1 will require approximately 5 minutes per article to process. Dryad is not at a state of development to accommodate this and the estimate should probably be adjusted to reflect this. With 5 minutes per article for processing, the following graph (Figure 1) shows ingest levels required to achieve the various possible per article costs.

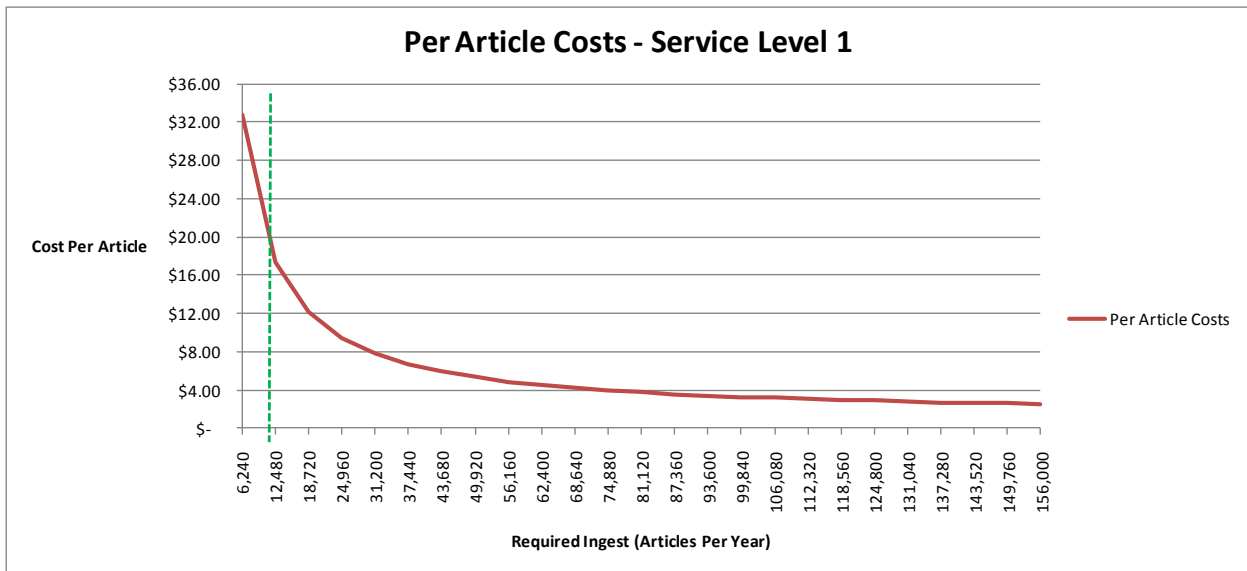


Figure 1 – 5 Minute Processing Time

If a more conservative estimate of 15 minutes per article is assumed, the ingest rates adjust as shown in Figure 2. The number of additional articles ingest is not huge, perhaps in the range of 1,000 to 1,500 more required annually.

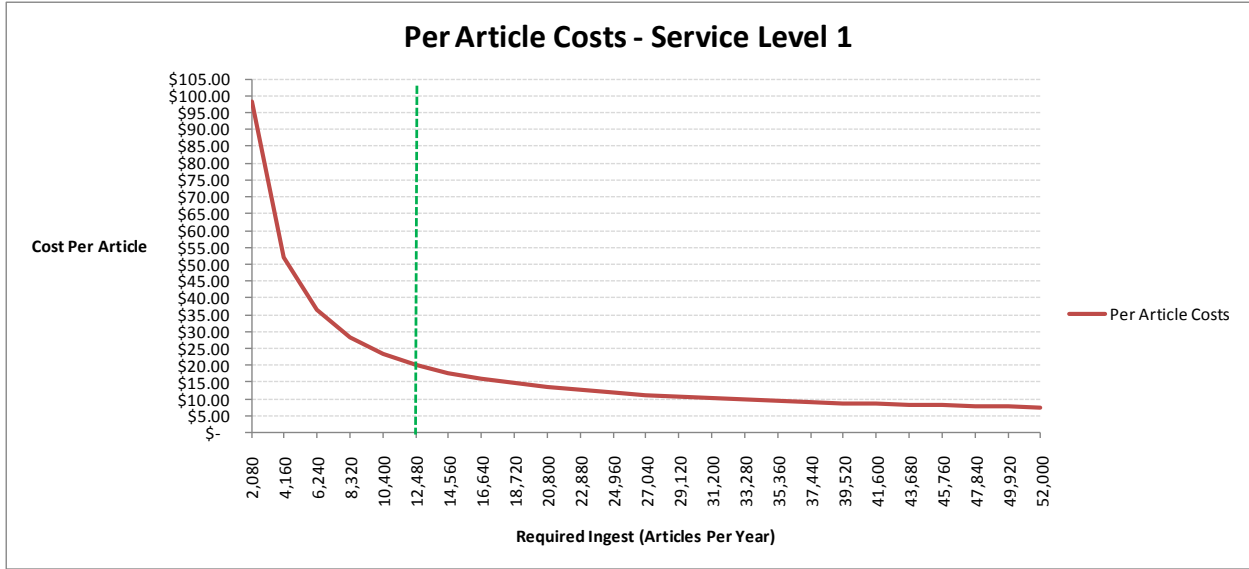


Figure 2 – 15 Minute Processing Time

Service Level 2 – 20 Minute or 45 Minute Processing Times

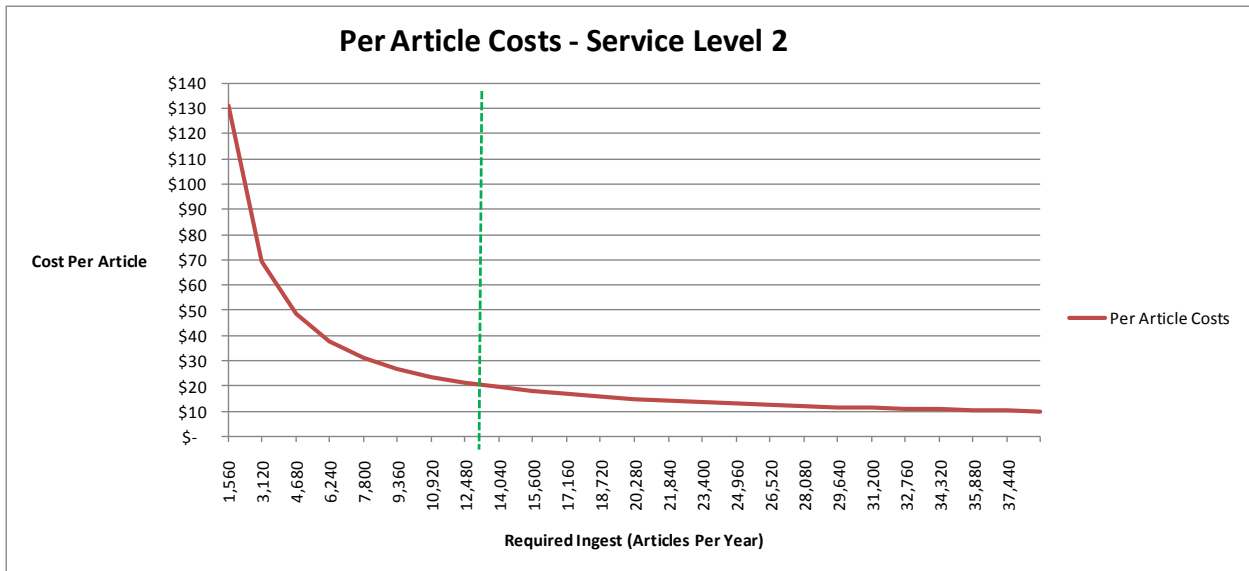


Figure 3 – 20 Minute Processing Time

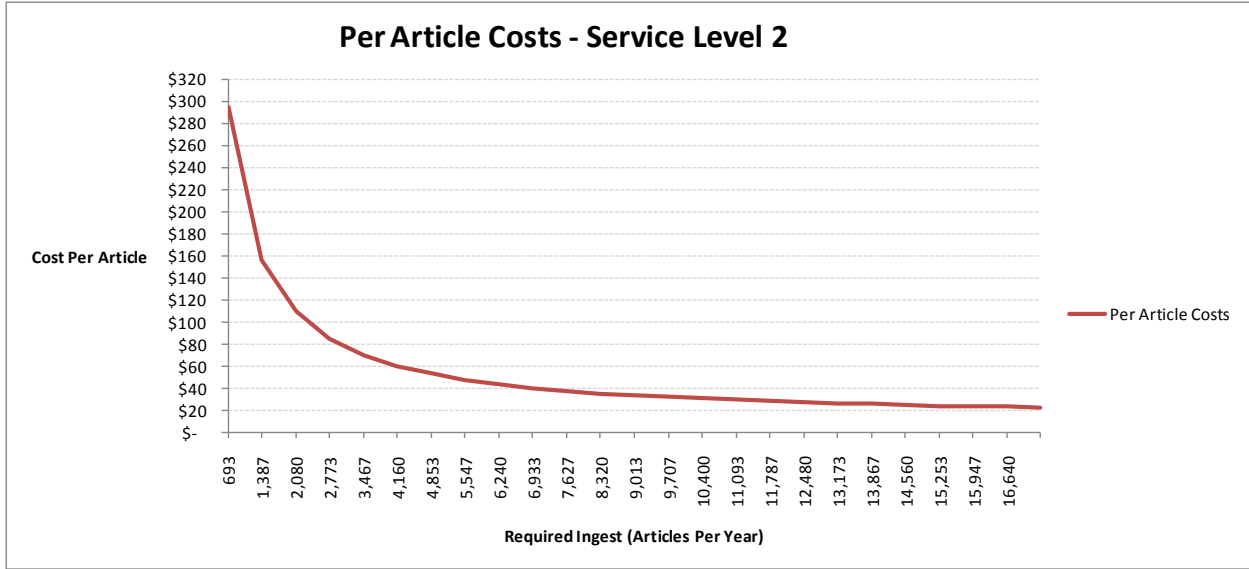


Figure 4 - 45 Minute Processing Time

Service Level 3 - 140 Minute and 22 Minute Processing Times

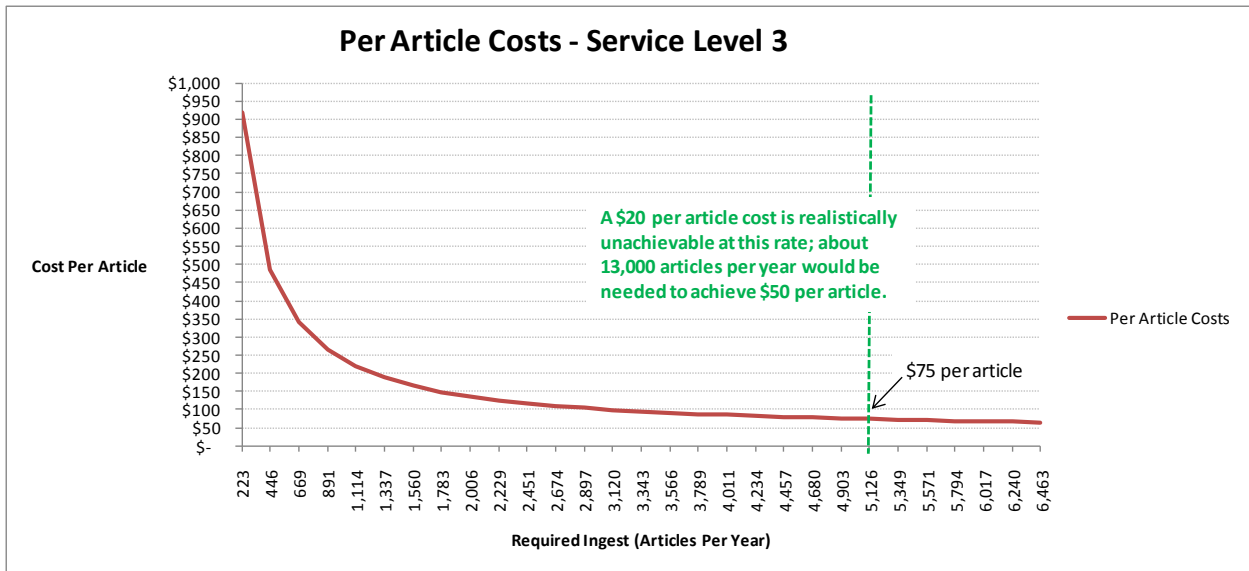


Figure 5 - 140 Minute Processing Time

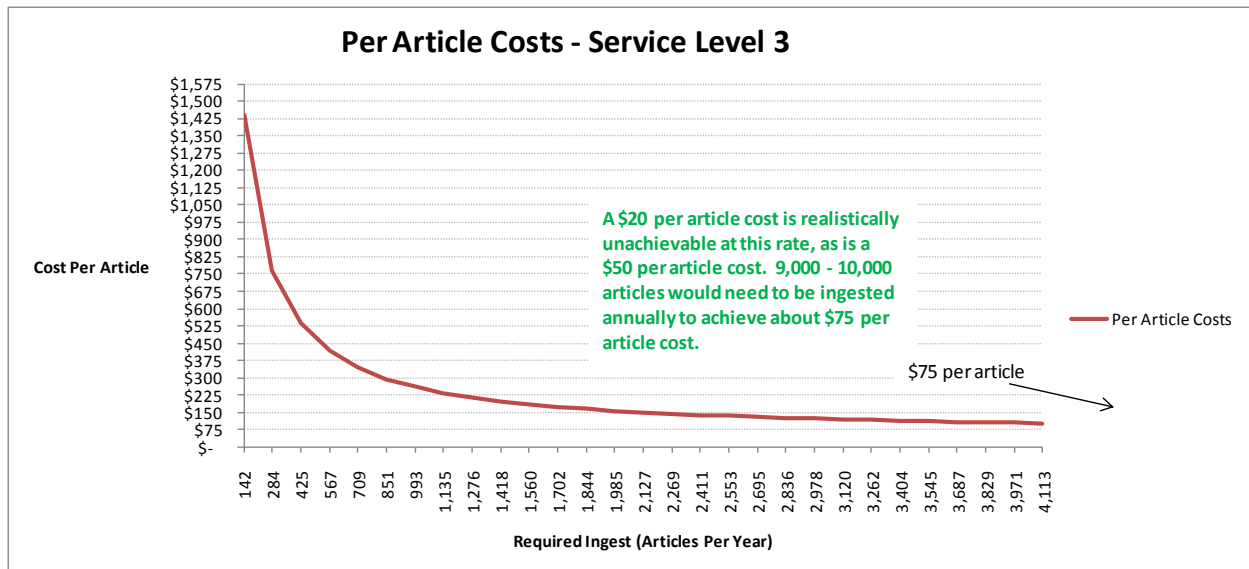


Figure 6 - 220 Minute Processing Time

Ingest Levels and Per Article Costs

Service Levels		Assumed Cost Per Article				
		\$20	\$50	\$75	\$100	\$150
Level 1	5 minutes	c. 10,000 - 12,000 articles (0.5 - 1 curation FTEs)	< 5,000 articles (< 0.5 curation FTEs)	Easily achievable	Easily achievable	Easily achievable
	15 minutes	c. 12,500 articles (1.5 - 2 curation FTEs)	c. 5,000 articles (c. 1 curation FTE)	c. 3,000 articles (0.5 - 1 curation FTE)	< 2,000 articles (c. 0.5 curation FTE)	Easily achievable
Level 2	20 minutes	c. 12,500 - 14,000 articles (2 - 2.5 curation FTEs)	c. 18,000 articles (c. 1 curation FTE)	c. 10,000 - 12,000 articles (0.5 - 1 curation FTE)	c. 5,000 articles (0.5 - 1 curation FTE)	Easily achievable
	45 minutes	c. 20,000 - 21,000 articles (7.5 - 8 curation FTEs)	c. 4,900 - 5,400 articles (2 - 2.5 curation FTEs)	c. 2,800 - 3,300 articles (1 - 1.5 curation FTEs)	c. 2,500 - 2,700 articles (1 - 1.5 curation FTEs)	< 1,300 articles (0.5 - 1 curation FTE)
Level 3	140 minutes	Not achievable	c. 12,000 - 12,200 articles (13.5 - 14 curation FTEs)	c. 5,000 articles (5.5 - 6 curation FTEs)	c. 3,100 articles (3.5 - 4 curation FTEs)	c. 1,600 articles (2 - 2.5 curation FTEs)
	220 minutes	Not achievable	Not achievable	c. 9,100 articles (16 - 16.5 curation FTEs)	c. 4,400 articles (8 - 8.5 curation FTEs)	c. 2,000 articles (3.5 - 4 curation FTEs)

Figure 7 - Required Ingest Levels and Curation FTEs For Given Per Article Costs

Some Key Cost Drivers

- Selected service levels are the key cost driver, since the majority of costs are associated with personnel in most repository cost models.
- Because Dryad is accepting all file formats, software licensing costs and migration could become unwieldy. It is recommended that Dryad personnel pay serious attention to these potential costs and strategize the long-term steps needed to ensure that future access needs and the repository's capacity to provide long-term preservation be well formulated. Further research into likely format submissions should also occur.

- *Dryad management should consciously consider the potential costs of implementing particular revenue models. Communications and outreach costs can be heavy for some types of revenue models, particularly those that require buy-in and participation among a large number of stakeholders. The process for ensuring proper communication and training should be planned in advance and cost estimates for doing this should be created. Travel and training of potential funders and Dryad communications personnel will be included in this.*
- *Hardware refresh cycles also can represent a key cost driver. How often hardware is refreshed and what types of architecture will serve the repository will be key. For example, some repositories are currently testing cloud storage and grid structures in order to significantly reduce capital costs over time. If Dryad anticipates that these costs will be provided as cost share from a parent or partner institution, they should create a risk management plan associated with the possibility that this institution (or institutions) may cease providing this cost-share item in the future.*

Recommended Next Steps

- *Create a comprehensive fully costed cost model, including those elements discussed earlier in this report (“Requirements to Upgrade to Full Costing Model”) and link this to a risk management plan that can serve as a discussion point with advisory board members for developing a strategy to handle potential losses in financial support or potential cost-related issues that are likely to crop up in the future (e.g., excessive migration and licensing costs associated with Dryad’s desire to preserve all submitted formats).*
- *Include a time discounted factor to the model, to ensure that communications with potential funding sources can provide accurate long-term assessments of their opportunity costs.*
- *Choose levels of curation and re-validate expected levels of work needed to perform those service levels. Examine the impacts of having curation performed by generalist curators vs. subject matter curators and professional vs. student curators. This may have a strong impact on overall costs. (E.g., specialist curators may have a higher hourly rate but may also perform curation on specialized literature more quickly. Likewise, if personalized communications with data submitters is expected, the overall time to provide training to the data submitters may be reduced with the use of subject specialist curators, but at the cost of higher per hourly rates.) Examine the impact of these factors on re-work due to error, which has been shown to be quite costly to repositories.*