University of California
Planning for Investigation of the Institutional Costs of Gold Open Access
Final Report, March 2014

Submitted by:
MacKenzie Smith, University Librarian
University of California, Davis
One Shields Avenue
Davis, CA 95616
Project Reference No. 41300106

Table of Contents

Summary

In the fall of 2013, the Andrew W. Mellon Foundation granted the University of California, via the UC Davis campus and in collaboration with the California Digital Library, funding to develop a plan for an extensive and rigorous investigation of the economic implications of the "Article Processing Charge" (APC) funding model for Open Access scholarly communication. The APC model is a variant of the so-called Gold Open Access model in which publishers are compensated for their effort by authors or their proxies (e.g., their institution) rather than by charging subscription fees for access. The Gold Open Access model also includes journals that have other sources of funding and do not charge authors, but our study is focused on APC-funded scholarship because we believe that this is likely to be the predominant open access funding model for the majority of journals in which authors at our institutions publish. We recognize that other forms of scholarship, e.g. monographs, are beginning to shift to APCs as well, but journals are much further along in the transition and represent a much larger proportion of library budgets so we are focused on journals in this study.

We are also focused exclusively on large, North American research universities since one of the main motivations for this work is the observation that a relatively small number of U.S. research universities, such as the University of California, generate a disproportionately large percentage of published research articles. There have been some related studies done on this topic but so far these have been focused on European research institutions that are primarily government funded and have much stronger government mandates to convert their publishing activities to an Open Access model (Finch 2013).

The outcome of our planning effort is a proposed year plus-long project to build a set of financial scenarios, or models, depicting the financial implications an APC-based system of scholarly journal publishing for the conversion of the current system of scholarly journal publishing to an APC-based system, for large research institutions. What follows is a description of our process to-date, proposed project, and plans for the future.

Review of Work Completed

The planning project that is the subject of this report was conducted over the four months from November 2013 through February 2014, and has produced the framework for a project that we are now developing a detailed proposal to conduct. The project team consisted of the PI at UC Davis, the co-PI and Director of Collections from CDL, and an external consultant who developed the planning project framework and coordinated our work. We conducted our work with weekly conference calls and monthly in-person meetings, together with conference calls and emails with the UC University Librarians and each of the proposed partners.

The project we have designed with this planning effort builds on work initially done by the California Digital Library to assess the implications on our budget of emerging APC levels and proposes an investigation to help us understand:

1. The overall cost to publish scholarly journals and the value of different components to editors and authors,
2. What the emerging APC model would cost UC and similar large research institutions under a variety of rigorously-modeled scenarios, and
3. Reasonable financial models for publishing research that cover basic costs and ongoing innovation.

The components of the project we envision are described below, including an overview of its goals and working methodology, the partnerships we have established to conduct the work, a preliminary budget for the project, and our intentions for completing the proposal and moving it forward over the rest of this year.

Proposed Project Overview

The key question that the proposed project asks is whether a large-scale conversion to Open Access scholarly journal publishing funded via Article Processing Charges (APCs) would be

viable and financially sustainable for large North American research intensive institutions, whose faculty currently author a significant percentage of the world's research. We define 'sustainable' in this context as costing these institutions roughly no more than, and ideally considerably less than, current subscription costs for comparable content today, with a rate of growth that will be possible for these institutions to support over time. We consider 'viability' to further encompass the willingness of authors to publish under such a model and the likelihood that they will do so if the option is available and sufficiently congenial to them, as well as the additional likelihood that a scholarly publishing infrastructure optimized for open access will allow research and scholarship to flourish. Our immediate interest is not in the cost of the scholarly system as a whole – although we may be able to draw inferences about overall system costs – but rather the impact on representative large research institutions that, as key funders of the scholarly publishing marketplace, need to understand and help define any new financial model that is proposed to replace the current system.

The project we have planned will be conducted over 14 months, beginning in the fall of 2014 and concluding at the end of 2015. Its aspects, goals, partners, and tentative budget are described below.

Quantitative Analysis

Several prior analyses have suggested that a publishing system relying on APC-funded publications would not be sustainable for large research institutions, given their large share of the overall research output, and this has led many observers to discount the viability of broad-based adoption of the APC funding model for journals (see for example Davis et al 2004; Báscones Dominguez 2006; Walters 2007). However, these analyses have often relied on simple calculations that fail to take into account the multiplicity of factors at play in an APC-funded publishing model, such as co-authorship patterns and the availability of sponsored research funding. Nor have they taken into account what authors want and believe, in general and among different disciplines. Recent studies commissioned in the U.K., where a transition to Gold OA is a stated (if controversial) policy objective, conversely suggest that an APC-funded publication model would be less expensive for research institutions if these elements are considered (e.g. Swan 2010; Swan and Houghton, 2012). Our goal is to expand and stress-test these assumptions by building a rich model that takes into account at least the following factors:

1. The current state of APCs as they are evolving today, including recent trend data,
2. The attitudes and behaviors of authors in various disciplines, including their willingness to publish under an APC model,
3. Changes in the global publishing environment that might reasonably be expected to lower the cost of publication, such as new technologies enabling lower overhead, elimination of cost categories such as print publication and access management infrastructure, and the likely range of APCs that might result from such changes,
4. The role of research funders in contributing to the cost of publication
5. Levels of co-authorship in many disciplines, which are likely to spread the cost of publication across multiple institutions,
6. The introduction of healthy market forces that can moderate pricing, and

7. Disciplinary differences in all of the above areas that must be factored into any well-developed model.

To understand the relationship of current institutional expenditures to potential costs under an APC model, the data analysis and modeling phase of this project will assemble comprehensive data about current subscription expenditures at a representative set of major North American research-intensive institutions, and will marry that with data about the journal publication output at those same institutions and in the world at large.   This will be accomplished in four overlapping phases of work, described below:

I.   Gathering Institutional Expenditure Data
II.   Gathering Publication Data
III.   Gathering Current APC Data
IV.   Scenario Development and Modeling

Institutional Expenditure Data

For this phase, we will collect comprehensive annual subscription and licensing expenditures for scholarly research journals from each library partner for the five-year period from 2009 to 2013. A rigorous methodology will be employed to eliminate costs likely to remain unaffected by a conversion to APC-funded publication, such as archival subscriptions and secondary full-text aggregations. Five years of data will allow us to perform trend analysis, i.e. to compare the results of the APC modeling to overall journal expenditures and to the growth rate of those expenditures over time. At the end of the project we will be able to answer the question "How much more or less expensive is a given APC funding model likely to be overall, both now and over time, compared with what research institutions are spending as whole for scholarly journals now?"

Each partner institution is asked to provide information about two categories of journal expenditure:

1) Total serials expenditure, both print and electronic, in each of the 5 years, excluding from the total expenditures for non- journal content and for some journal content that might be presumed to continue on a subscription basis under an APC model
2) Expenditures for open access publication that may exist outside of the traditional serials budget (e.g. open access memberships and campus publication funds)
3) Serials expenditure data will be collected according to the following methodology:
    a. Exclusions: The following categories of serial expenditure will be excluded from the expenditure totals used for modeling purposes:
        i.   Databases, ebooks, and other non-journal content that may be funded via the serials budget
        ii. Secondary aggregator databases (EBSCO, ProQuest, Gale, Factiva, Lexis-Nexis, etc.)
        iii. JSTOR retrospective collections and other archival purchases or subscriptions
        iv. Newspapers and other news sources

> v. Popular and trade journals and magazines
> vi. Other content categories that may be identified

b. Total expenditures for each year will ideally be broken out by print vs. electronic in order to make it possible to analyze the impact of retaining or discontinuing print in the subsequent models. Combined print and electronic subscriptions will be reported separately or counted as electronic, to be determined in consultation with the partners.

c. Expenditures for a small number of known journal packages (e.g. ACM Digital Library, IEEE Xplore Digital Library) may need to be adjusted to account for non-journal content in those packages that would continue to be licensed on a subscription basis if the journal content were converted to open access. To control for this factor, the publishers will be asked (under the auspices of our partnership with ALPSP) to provide the percentage of cost attributable to the journals in those packages; each library partner will be asked to calculate the cost of the journals in its package using that percentage.

4) Analyzing journal expenditures by discipline

The institutional partners would like to understand not just the potential impact of APCs on their bottom line journal expenditures, but also how expenditures by discipline might change under an APC-funded model and how this aligns with disciplinary culture and funding patterns. However, tying journal expenditures to disciplines is not straightforward due in large part to the bulk pricing structure of large publisher packages. We are evaluating several methodologies for estimating expenditures at a disciplinary level, including sampling a subset of data at one or more partners, and will select a methodology in consultation with the library partners that can be implemented with relative efficiency and reliability.

Publication Output and Authorship Data

The publication data used to construct our models will be obtained via a partnership with Scopus. Scopus has been selected as the project's primary data source because it contains the largest number of scholarly journals of any comparable database, including some 20,000 peer-reviewed journals, and offers more comprehensive coverage in the humanities and social sciences than the Web of Science.

Two sets of publication data will be acquired from Scopus for modeling purposes: general data about publication output, by discipline; and publication data specific to each institutional partner. Broad-based publication data will allow us to discern disciplinary differences in publication practices such as publication volume and patterns of co-authorship and will serve as a baseline to determine whether the publication characteristics of the partner institutions are broadly representative or exhibit unusual traits. It will also allow us to make inferences about the percentage of worldwide production accounted for by large North American research intensive institutions and, thus, the financial burden such institutions might assume in an APC-funded system. The specific data to be acquired include the following:

1) *General Publication Data*: By Scopus discipline, by year of publication 2009-2013, for all research journal and proceedings content – excluding trade publications, books, etc.:

    a. # of research articles

    b. # of non-research articles

    c. Histogram of the distribution of research vs. non-research articles across all journals in that discipline (to discover if significant non-research-article content tends to be concentrated in a small number of journals within a discipline, or is broadly characteristic of the discipline). Since non-research-article content must be supported in any business model, understanding the distribution of this content will be important in suggesting how that content might be funded in an APC-funded system

    d. # of research articles that are sole-authored vs. multi-authored

    e. # of research articles that have co-authors from multiple institutions

    f. Average and median # of co-authors per research article

    g. # of research articles that are the result of sponsored research

    h. # of open access journals

    i. # of open access articles (sampling techniques will be required in order to estimate the number of OA articles in so-called 'hybrid' journals)

2) *Institutional Partner Data*. For each institutional partner, by discipline, by year of publication 2009-2013

    a. # research articles with at least one author from that institution. From that set:

    b. # of research articles that are sole-authored vs. multi-authored

    c. # of research articles that have co-authors at multiple institutions

    d. Average and median # of co-authors per research article

    e. # of research articles that are the result of sponsored research

    f. # of open access articles (sampling techniques will be required in order to estimate the number of OA articles in so-called 'hybrid' journals)

    g. # of research articles whose Corresponding Author is from the partner institution vs. another institution

    h. For the most recent 1-2 years, a list of journals in which authors at the institutional partners have published, with unique article counts

3) *Data normalization.* There are a few known weaknesses in the Scopus data that the project may have to overcome. For example, 20% of Medline titles indexed by Scopus include the article first author only. More significantly, Scopus cannot identify articles resulting from grant-funded research with any reliability prior to 2013. We are evaluating the potential impact of this on our modeling and will seek to implement appropriate mitigation strategies for any weaknesses or anomalies in the data (e.g. if necessary, the project may decide to acquire relevant data from Web of Science).

Gathering APC Data

An analysis of current article processing charges as they are developing in the marketplace will serve as an important baseline for our modeling. Publication fees for hybrid OA journals are commonly estimated at $3,000 for example, yet some of the largest publishers are now differentiating these fees by journal and by discipline. Newer entrants into the open access publishing sphere with lower cost structures and thus lower publication fees are also radically altering the APC environment. Any modeling needs to take these current trends into consideration. Prior studies by Solomon and Bjork(Solomon and Bjork 2012; Bjork and Solomon 2014) extensively researched APC costs, but most of this information dates from 2011. The current project proposes to update this work to 2014 for the journals in which authors at the library partner institutions publish. This will not only provide a more current baseline keyed to the publishing behaviors of the large research institutions under study, but will also provide information about trends in APCs as they are developing in the laboratory of the marketplace. Solomon and Bjork will build on their prior work for this project and will correlate the data with Scopus disciplinary categories and impact measures to produce a map of likely APCs for the institutional partners differentiated by discipline and impact.

Scenario Development and Modeling

Working in consultation with a research economist (Mark McCabe) and under the direction of the core project team (Tananbaum, Anderson, Smith, and Farley), the project will incorporate all of the data inputs developed in this project – institutional expenditures, Scopus publication data, current APCs, potential future APCs derived from our research on publication costs, and information on author needs and preferences gleaned from author surveys and focus groups – to build a series of models depicting the social and financial impact of a largely APC-funded publishing landscape on the institutions participating in the study. A modeling consultant will work with us to ingest the data, develop, and iterate one or more scenarios under a variety of assumptions including levels of grant funding, co-authorship patterns, realistically envisioned APCs, and author uptake of APC publishing venues in different disciplines.

Cost Modeling

As previously mentioned, a key goal of the project is to construct a rich model that takes into account the current research output at large North American institutions, the ways in which a transition to a fully APC-based model might impact the journal publishing expenditures of these institutions, and how sustainable and desirable this fully APC-based model might prove for not only the institutions, but also for authors and for publishers.

A key component of this model construction is therefore the enumeration and costing out of key tasks associated with publishing scholarly journals. This is necessary to determine whether a APC-based model that is sustainable for large North American institutions and for authors is also viable for publishers.

We intend to develop a ground-up cost model for publishing in different disciplines. This modelling will draw from the experiences of scholars and departments from the partner institutions, including UC Davis neuroscientist Michael Rogawski, who is founding a new

journal of negative results called *Epilepsy Reports* with the Society for Neuroscience, and the California Digital Library's eScholarship unit, which manages a collection of 65 electronic journals. Other sources of this type of information will be consulted as appropriate, such as the directors of PeerJ, Ubiquity Press, and other new online publishers that have managed to dramatically lower costs for APC-based publishing. The model will also rely on available literature (e.g., 73 Things Publishers Do; "Launching (and Sustaining) a Scholarly Journal on the Internet: The International Journal of Baudrillard Studies"), and the project manager's direct experience in launching dozens of scholarly journals across multiple disciplines.

The goal of the ground-up cost model is to accurately reflect the full range of expenses that are necessary to run an APC funded journal. These expenses include salaries and benefits, editorial, technical, operations, sales and marketing, and administrative. The model must also address the notion of profits, and how these profits can be used to fuel growth and innovation.

We chose this approach for several reasons. We explored the possibility of acquiring actual line item expenditure data from both individual publishers and through publishing trade organizations but were unable to do so. Publishers are reluctant to share this information on the grounds that it is proprietary and might be misrepresented or misunderstood. In addition, discussions with a number of publishers revealed that the quality of the cost data they maintain is inconsistent and varies widely across publishers. These inconsistencies would make it difficult to draw accurate conclusions based on such data even if it could be readily obtained. Finally, existing publishers have legacy procedures and systems they may be keen to maintain, even in the event that the industry transitioned largely to a fully APC-based model. As such, any modelling built off of existing costs would encounter a number of legacy expenditures that would be difficult to weed out from essential operations. A ground-up model provides a more accurate window into what the true costs of a fully APC-based model would be. We will also compare components of this model with author attitudes and preferences about which publishing services they value.

Qualitative Analysis

The shift to an APC funding model implies the introduction of a new socio-technical system[1] for scholarly publishing and will produce changes in workflow for numerous stakeholders. Literature from business management, informatics, science and technology studies, sociology, and anthropology indicates that the successful and sustainable introduction of any new socio-technical system requires the "buy-in" of these affected stakeholders (Steiner 2008). Buy-in is predicated upon a stakeholder's perception that the proposed changes required for the new system result in added "value" over that provided by the legacy system being replaced (de Waal and Batenburg 2014). In assessing this it is important to keep in mind that "value" is a multi-faceted and, often inter-subjective, concept that is directly linked to an individual's unique social and institutional context (Bijker et al. 1997, Mol 2002, Pinch 2003). Given this complexity socio-technical design scholars suggest that sound research methodologies elucidate the

---

[1] A socio-technical system is a social system (i.e., academic publishing) that is built upon a technical base (i.e., the technologies and technical workflows required for publication).

perspectives of stakeholders, their actions and motivations, and relationships between stakeholders (Sjostrom and Goldkuhl 2009, Scholl 2004, Whitworth 2009).[2]

Establishing a baseline understanding of stakeholders' "value" and potential "buy-in" to a different scholarly publishing systems is a necessary first step in our research plan. In order to understand potential barriers and facilitators in sufficient detail to accurately deploy these factors in our modeling, we will first need to identify the stakeholders[3] and then assess their perception of the value derived from the current publishing paradigm versus the potential value possible in an APC model. Key stakeholders in this assessment include researchers (faculty and graduate students who both author and read publications), scholarly institutions (universities, academic departments, and libraries), and publishers.

The data gathered from these key stakeholders will be used to inform and enhance the accuracy of the economic models we will be developing as the final product of this research. At this time there is a limited literature comprehensively addressing these matters with respect to a transition to APCs as this would specifically impact the North American research intensive university context. Moreover, these studies were conducted for other purposes and, while informative in designing this project, do not have the requisite detail and orientation required for our economic models[4]. Additionally, because the landscape of scholarly communication is changing so rapidly, the literature based upon data collected just a few years ago may not be appropriate to the current context and its use would negatively impact the accuracy of our models.

Project partners from the publishing world (ALPSP) will facilitate acquiring the data and analysis for this stakeholder group based upon information from their membership. Our intention is to survey ALPSP's publishing membership, which encompasses a diverse range of journal publishers. This will allow us to better understand how the growth in Gold OA and APCs has impacted publishers' operations, as well as their organizational planning. Such data will be useful in determining the appetite within the publishing community for exploring alternative models. ALPSP will collate and anonymize the data, as well as send out the survey under their imprimatur. This should ensure robust participation among the publisher community. For the

---

[2] Design research has been faulted for its emphasis on technological determinism without reference to social context resulting in high failure rates for newly introduced systems. Authors suggest that research should close the gap between social needs and technical performance or, in other words, what technology *does* and what communities *want* (Whitworth 2009).

[3] See work by Gerald Steiner (2008) for a discussion of stakeholder identification and how stakeholder perspective can be employed in designing sustainable innovations. An article by Achterkampf and Vos (2007) focuses on stakeholder identification in the context of civil society. 4 One large-scale survey, the EU-funded SOAP Project (Dallmeier-Tiessen 2011) surveyed author attitudes about open access, including funding experiences and barriers. Another similar survey-based study was conducted in 2013 and drew from author-respondents publishing in Taylor & Francis Group and Routledge journals (Frass et al. 2013). Finally, work by Diane Harley and colleagues (2010) has investigated faculty attitudes towards the changing landscape of scholarly communication across seven different disciplines. However, the research by Harley and colleagues does not address the economic feasibility of various publishing models.

project, Carol Tenopir will gather and analyze data for the researcher and institutional stakeholders using qualitative methods, such as focus groups and surveys.

Such a baseline would establish (1) How do researchers value the publication of their own work? (2) How do researchers and institutions value the accessibility of the scholarship that they generate? (3) How do researchers and institutions value the production and accessibility of other scholars' work? (4) How do these attitudes vary amongst different disciplines? And (5) How do these attitudes vary across different institutions? In particular, we want to assess how researchers and their institutions currently understand their role vis-à-vis one another in the current publishing paradigm (gated access), how they think this role might change in the event that the scholarly communication changes to an APC-funded model, and what perceived advantages and drawbacks might such a change produce. This will inform our understanding of the acceptability of APCs for various stakeholders and may provide initial clues as to how advantages could be highlighted and drawbacks addressed in a successful and sustainable model.

In particular, we will gather qualitative data from researchers in different disciplines to assess what types and forms of scholarly communication are important to each field, whether APC-financed OA journals is a feasible goal for these different forms, and if there is widespread familiarity and experience with both OA and APCs. For example, most health sciences researchers primarily produce article-length publications with multiple contributing authors and, given funding mandates of the NIH, are familiar with and have published in open access journals and databases (e.g., PubMed journals and PLOS). In contrast, researchers from the humanities and social sciences have significantly less experience in this matter because fewer open access publishing platforms are available and those that do exist do not have the requisite legitimacy to attract a desirable audience. In particular, many humanities disciplines consider book-length manuscripts to be the ideal vehicle for the dissemination of a fully developed scholarly argument and few influential works have been successfully produced through open access mechanisms at this point. Therefore, there is more at stake for a humanities researcher since it would require him/her to publish the culmination of several years' work in an untested, and possibly unsuccessful, manner.

A prior study on a similar topic has noted that author attitudes towards OA scholarly publishing is such a complex issue that it is difficult to conduct "meaningful research" in a single set of survey questions (Frass et al. 2013, 4)   Therefore, the qualitative arm of our proposed project on APC-funded models will proceed in two phases. In the first phase we will gather data through focus groups with researchers from different disciplines at the six partner institutions. These data will inform the second phase of the qualitative arm which will consist of surveys administered to the entire faculty of the six partner institutions. It is commons practice to use small, hand-selected focus groups to inform population-wide surveys when little understanding of a social phenomenon exists (Bernard 2006). Both the focus groups and surveys are described in greater detail below.

Focus Groups of Faculty Members and Graduate Students

Focus groups provide a way for candid conversations not only about current practices and beliefs, but about desirable or predicted changes into the future. They do not attempt to involve a representative sample of all viewpoints, but identify some current key behaviors and attitudes, and help bring out some strong beliefs and possible benefits or costs into the future. They help suggest questions to be asked in follow-up surveys. Focus groups will be limited to 8-12 participants in each stakeholder group, including representation from multiple subject disciplines. We will plan a total of 9 focus groups (two at three UC campuses and one at each partner institution) with an average of 10-11 participants in each, for a total of approximately 100 participants. Focus group questions will vary according to the group, but will include things such as:

- How does access to high quality scholarly publication venues help you as an author, a researcher, a teacher, a student, etc.?
- Which components of the publication process do you value most?
- What do you think are the best ways to pay for the costs of journal/article publishing? Who is the most logical party or parties to pay for these costs?
- Which models do you use/support now and which will you in the future?
- How can you/members of your stakeholder group influence cost models?
- Have you ever been asked to pay an author payment for an article at the time of publication? If so, what did you think about that? If not, would you be willing to?
- What are the most important criteria you use to determine where to publish and how do you see that changing in the future (if at all)?
- Is there a difference in quality between gold OA and subscription-based journals? How is pricing related to quality?
- How might a change in pricing mechanism from subscription to Gold OA help/hurt scholarship in your field, other researchers, your institution? Are there other viable alternatives?

Surveys of Faculty and Graduate Students

Surveys are a good way to reach large numbers of participants to get candid responses to specific questions about current behaviors, attitudes, and opinions. They can be analyzed to reveal statistically significant differences among sub-groups that can be extrapolated to populations and subpopulations and ensure that a range of respondents are heard. Survey instruments will be developed and tested by UTCICS and distributed by partners. Some of the questions will replicate those asked at other venues in the past (see, for example, Tenopir, et al., Nicholas, et al., Xia), other questions will be unique to this study.

Survey response rates for academics has been decreasing over time, but we can anticipate an approximate 10-20% response rate for faculty members, with slightly lower rates for students. To get a target of approximately 300 responses per campus to enable meaningful statistical analysis of results, surveys should be distributed to at least 2000-3000 faculty members and 3000-4000 graduate students. The surveys will seek to answer questions such as:

- How many articles did they publish in the last year, 2 years, 3 years? How many co-authors on average on each article?

- In what types of venues are faculty and students currently publishing? (i.e., are they currently publishing in OA venues or with APC-funded journals and, if so, for what % of their articles?)
- Have they ever or are they willing to pay APC fees? From grant funds? From their departmental or personal funds? From another source?
- What do they think is the most appropriate source of APC fee funding (if any)?
- What range of costs/fees do they consider reasonable?
- What are their personal opinions (if any) about the costs and prices of publishing?
- Do they envision changes in their practices and in publishing in the near-term future?
- Do answers vary by discipline, age, rank, or publishing productivity of academics? (for example, are faculty members who publish more frequently less likely to publish in APC-funded journals? are faculty members in the sciences more favorably inclined towards author payments than faculty members in the humanities? do younger faculty members have a lower threshold for a reasonable payment amount? are PhD graduate students more likely to envision changes in publishing than professors?)

Beyond the benefits described above, the inclusion of the behavioral and attitudinal data we will collect is important to our financial modeling for several reasons. It will allow us to understand the extent to which authors in specific disciplines have the financial wherewithal to contemplate a growing APC publishing market. How diverse are their funding sources? How flexible are these funding bodies in the application of grant monies toward APC charges? How are these considerations changing over time? And it will provide a snapshot of what authors in specific disciplines are actually doing today with respect to APCs. The behavioral and attitudinal data to be gathered by the project will help inform the models developed in the final phase of the project. They will provide important context for the pricing levels and funding diversity within these models. This will extend the project beyond the realm of theoretical exercise toward a set of prescriptions that have the chance of adoption within the authoring and publishing communities.

Proposed Project Partnerships

In addition to the core project team of the PI, co-PI, CDL consultant, and external consultant (the same team who performed this planning effort) we have identified an extensive set of partners necessary to conduct the work we envision for the project. These span the key stakeholder communities and expertise required to conduct the research we described and to build the models that will be the main outcome of the project. All of the partners listed below have agreed to participate in the project if it is funded and have worked with us to develop Statements of Work and budget to support their participation. While the Statements of Work are not included in this report, since they are still under discussion in a few cases, we have included the current budget requirements in the section below to provide sense of the financial scale of the proposed project.

University Partners
While the project is based in the University of California and will involve the ten UC campuses and the CDL, we feel it necessary and useful to include additional research institutions for several reasons: to validate the data that we produce for UC, to compare our faculty's attitudes

and behaviors to others, and to bring additional and complementary library and institutional expertise to the project. We sought out additional partner institutions that were both large (i.e., Carnegie RU/VH) research institutions with heterogeneous research programs that include humanities, social, life and physical sciences, and that had extensive experience with Open Access publishing (e.g., a faculty mandate or extensive library engagement). After due consideration of the many institutions that we could consider inviting as partners, we settled on three:

Harvard University, Cambridge, MA, USA
Ohio State University, Columbus, OH, USA
University of British Columbia, Vancouver, BC, Canada

Our focus groups will include these three partners and three of the ten UC campuses. The faculty survey will be conducted at those institutions and the remaining seven UC campuses.

Publishing Partners
On the publishing side, we have selected two key partners to help us collect data and explore publisher attitudes about APC funding for OA journals.

Scopus (Elsevier)
As a key source of data and analytics expertise, we will partner with Elsevier's Global Academic Relations (GAR) team, a non-commercial arm of the company that has extensive access to and knowledge of the Scopus database on which much of our research will depend. The University of California subscribes to Scopus and has access to this data directly, but we feel that the partnership with the GAR team will provide critical data analysis expertise that we can independently verify with our other partners. The GAR group has committed to partnering with UC Davis/UC on this project as an unfunded partner, with Dan Morgan as the lead coordinator.

Association of Professional and Learned Society Publishers (ALPSP)
ALPSP is an international association representing all types of nonprofit publishers and is the largest trade association for scholarly and professional publishers with more than 300 members in 40 countries, publishing scholarly content in many different ways. ALPSP membership includes a wide range of different types of publishers - journal publishers, book publishers, learned societies and professional bodies, database publishers, university presses and intergovernmental organizations. The variety and range of publisher members and associate members provides a unique network serving the interests of scholarly publishing worldwide.

Qualitative Analysis Partners

Carol Tenopir, University of Tennessee, Knoxville, TN, USA
Dr. Tenopir is a Chancellor's Professor at the School of Information Sciences at the University of Tennessee, Knoxville and the Director of Research for the College of Communication and Information, and Director of the Center for Information and Communication Studies. Her areas of teaching and research include: information access and retrieval, electronic publishing, and the information industry. She is the author of five books, including, Communication Patterns of Engineers, winner of the American Society for Engineering Education, Engineering Libraries

Division 2005 Best Publication Award, (IEEE/Wiley InterScience, 2004) with Donald W. King. Dr. Tenopir will lead the project's qualitative data collection and analysis, in collaboration with the UC Davis team and members of the quantitative data team (Bjork, McCabe, and Anderson).

Quantitiative Analysis Partners

Bo-Christer Bjork, Hanken School of Economics, Helsinki, Finland

Dr. Bjork is a Professor of Information Systems Science in the Department of Management and Organisation, Hanken School of Economics, Helsinki, Finland. Dr. Bjork possesses advanced degrees in systems science, economics, and construction management. From 1993 to 2000 he was professor of construction IT in the Royal Institute of Technology in Sweden, where he founded the *Electronic Journal of Information Technology in Construction*, an early OA journal. This led to an enduring research interest in the scientific research process which has been the focus of his work since 2000. Dr. Bjork chaired the FinnOA committee from 2003-2008 and was a member of the board of the Open Access Scholarly Publishing Association (OASPA) from 2000-2012. He has an extensive open access publication portfolio includes numerous commissioned research reports for organizations including the British Library, the Max Planck Society Library, the Wellcome Trust, and others. With David Solomon (below), Dr. Bjork recently co-authored Developing an Effective Market for Open Access Article Processing Charges 2014, a report commissioned by a consortium of research funders including Jisc, Research Libraries UK, Research Councils UK, the Wellcome Trust, the Austrian Science Fund, the Luxembourg National Research Fund and the Max Planck Institute for Gravitational Physics.

David Solomon, Michigan State University, East Lansing, MI

Dr. Solomon is a Professor in the Department of Medicine and the Office Medical Education Research and Development at Michigan State University, with thirty years of experience in social science/educational research and evaluation. In 1996 he founded *Medical Education Online* (MEO), a respected peer reviewed web-based journal in medical education now published on an open access basis by Co-Action Publishing. Dr. Solomon is the author of *Developing Open Access Journals, A practical guide (*Chandos Publishing), and with other colleagues founded the Open Access Scholarly Publishing Association (OASPA), for which he served as a founding board member. Since 2011, Dr. Solomon has focused much of his scholarly work with Dr. Björk and others researching the nature and growth of open access publishing, particularly APC-funded OA publishing.

Mark McCabe, University of Michigan, Ann Arbor, MI

Dr. McCabe is a Research Investigator at the University of Michigan's School of Information. His current research interests include industrial organization, competition policy and regulation, and information economics. Dr. McCabe is an expert on the economics of journal publishing. He has written several reports and articles on the topic, including, "Online Access and the Scientific Journal Market: An Economist's Perspective," a commissioned report for the National Academy of Sciences' Board on Science; and "A Portfolio Approach to Journal Pricing," in the book *Economics and Usage of Digital Libraries: Byting the Bullet*. Dr. McCabe has published numerous articles on various component of journal publishing, including open access, online availability and citation counts, and journal pricing. His work has been published in *American Economic Review, Nature, Rand Journal of Economics,* and *Journal of Academic Librarianship,*

among other leading journals. Dr. McCabe has also served as an adjunct associate professor and a visiting assistant professor at the School of Information, University of Michigan. He has been a visiting professor at SKEMA Business School, Boston University, and Georg-August-Universität Göttingen. Dr, McCabe has also held positions at Georgia Institute of Technology, the Antitrust Division of the US Department of Justice, and American University. He has received two Mellon Grants for his work on journal publishing – "Measuring the Impact of Digitization and Online Availability on Journal Citations," (co-PI with Christopher Snyder, Dartmouth College), and "Scholarly Journals," (co-PI with Daniel Rubinfeld, Aviv Nevo and Aaron Edlin, all at UC Berkeley). Dr. McCabe has also received grants on this topic from the Sloan Foundation, the Open Society Institute, and the American Association of Law Libraries.

Greg Tananbaum, ScholarNext, Berkeley, CA
Greg Tananbaum is a scholarly communications consultant with nearly 20 years of experience at the intersection of technology, content, and academia. Since 2006, Greg has run ScholarNext Consulting, working with clients on a broad variety of issues, including strategic planning, business development, management consulting, policy analysis, and understanding and developing new functionality. ScholarNext clients include publishers, technology companies, universities, libraries, non-profit organizations, startups, and other scholarly communication entities. Representative engagements include the development of new discovery services for scientific researchers (Microsoft); usability and design testing for a highly visible new life sciences publication (*eLife*); development and authorship of an open science policy for a major research funder (American Heart Association); and development of the Open Access Spectrum guide (PLOS). He has also served as a consultant to the Scholarly Publishing and Academic Resources Coalition (SPARC) since 2007 and as an advisor to the Association of Research Libraries for its SHared Access Research Ecosystem (SHARE) project, which offers a long-term vision for higher education to manage its digital assets. In addition to his consulting work, Mr. Tananbaum has served as President of The Berkeley Electronic Press (which he helped launch in 2000), as well as Director of Product Marketing for EndNote. He writes a regular column in *Against the Grain* covering emerging developments in the field of scholarly communication. Mr. Tananbaum has been as an invited speaker at dozens of conferences, including the American Library Association, the Society for Scholarly Publishing, the Association of Professional and Learned Society Publishers, and Online Information UK. He holds a Master's Degree from the London School of Economics and a B.A. from Yale University.

Conclusion and Future Plans

The model of Open Access for the dissemination of scholarly communication is growing rapidly and could possibly become the default in some disciplines within the next decade, particularly with recent mandates coming from governments and funding agencies in the U.S. and Europe. While higher education institutions have been closely involved in explorations of the "Green" model of OA, we have largely left unexplored those experiments with the "Gold" model. These models, driven mainly by publishers charging Article Processing Charges (APCs) to recover their costs, often maintain current profit levels for this group at the expense of other stakeholders. As the Gold model with associated APCs takes hold there is a very serious risk to higher education, in general, and large research universities, in particular, for losing even more

ground to commercial publishers that pursue high profits at the expense of research. OA is a laudable ambition for scholarly communication, but it must be achieved in a way that restores balance to the scholarly ecosystem and is sustainable for the long term.

Our project has accomplished a prodigious amount of background research and planning in a short amount of time. Given the complexity of the project we are designing in both methodology and partnerships, we are still working through various details of funding, timing, data processing, and formal agreements (e.g. with the Global Academic Relations team at Elsevier). We expect to complete these details over the next two months and to have a fully-specified proposal by June 1.

A version of this report that includes the methodology discussed above and other logistics for the planned project has been published online at http://icis.ucdavis.edu/?page_id=286 under a CC-BY license.

We would like to thank the Mellon Foundation for making this initial work possible and ask for your feedback on how well-aligned our plans are with the Scholarly Communications and Information Technology Program's priorities and whether there is interest in funding the project we are planning. This is the ideal time to engage with Foundation staff, while there are opportunities for adjustments to the project scope, methodologies, partnerships, process, and budget that we have described above. We look forward to your reactions and, ideally, to proceeding with the project later this year.

Bibliography

Achterkamp, Marjolein and Janita Vos 2007. Critically identifying stakeholders: Evaluating boundary critique as a vehicle for stakeholder identification. *Systems Research and Behavioral Science* 24(1): 3-14.

Anderson, Kent 2013. 73 things publishers do (2013 edition). *Scholarly kitchen* ((Oct 22, 2013). http://scholarlykitchen.sspnet.org/2013/10/22/updated-73-things-publishers-do-2013-edition/.

Aspesi, Claudio, Andrea Rosso, and Wielechowski 2012. Reed Elsevier: Transitioning to Open Access – Are the cost of savings sufficient to protect margins? Report. http://www.richardpoynder.co.uk/OAcosts.pdf.

Báscones Dominguez, M. 2006. Economics of open access publishing. *Serials: The Journal for the Serials Community*, *19*(1), 52-60.

Bernard, Russell 2006. *Research methods in anthropology: Qualitative and quantitative approaches* (4th ed). Oxford: Alta Mira Press.

Bijker, Wiebe, Thomas Hughes, and Trevor Pinch 1987. *The social construction of technological systems: New directions in the sociology and history of technology*. Cambridge: MIT Press.

Björk, Bo-Christer and David Solomon 2014. Developing an Effective Market for Open Access Article Processing Charges. Report. http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtp055910.pdf.

Cook, Joel, Daniel Hulls, and David Jones 2011. Heading for the open road – costs & benefits of transitions in scholarly communications. A report for the JISC. http://www.jisc.ac.uk/media/documents/publications/reports/2011/dynamicsoftransition.pdf.

Coulter, Gerry 2010. Launching and sustaining a scholarly journal on the internet: The International Journal of Baudrillard Studies. *Journal of Electronic Publishing* 13(1).

Dallmeier-Tiessen, Suenje, et al. 2011. Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing. arXiv:1101.5260.

Davis, Philip M. and Cornell University Library. Task Force on Open Access Publishing 2004. Calculating the Cost per Article in the Current Subscription Model. (spreadsheet) http://hdl.handle.net/1813/236.

de Waal, Benny ME and Ronald Batenburg 2014. The process and structure of user participation: a BPM system implementation case study. *Business Process Management Journal* 20(1): 107-128.

Edgar, Brian and John Willinsky 2010. A survey of scholarly journals using open journal systems. *Scholarly and research communication* 1(2).

Finch Group, Accessibility, sustainability, excellence: how to expand access to research publications 2013. http://www.researchinfonet.org/publish/finch/.

Fisher, Julian 2008. Scholarly publishing re-invented: Real costs and real freedoms. *Journal of Electronic Publishing* 11(2).

Frass, Will, Jo Cross, and Victoria Gardner 2013. Open Access survey: Exploring the views of Taylor & Francis and Routledge authors. Report for Taylor & Francis Group and Routledge. http://www.tandf.co.uk/journals/pdf/open-access-survey-march2013.pdf.

Harley, Diane, Sophia Krzys Acord, Sarah Earl-Novell, Shannon Lawrence, and Judson King 2010. Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines. Center for Studies in Higher Education, UC Berkeley. http://escholarship.org/uc/cshe_fsc.

Houghton, John, Bruce Rasmussen, Peter Sheehan, Charles Oppenheim, Anne Morris, Claire Creaser, Helen Greenwood, Mark Summers, and Adrian Gourlay 2009. Economic implications of scholarly publishing models: Exploring the costs and benefits. A report for the JISC. http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf.

Mol, Annmarie 2002. *The body multiple: Ontology in medical practice.* Durham: Duke University Press.

Natriello, G and M. Rennick 2004. The middle way: The very low cost model for scholarly publishing. EdLab Research Publication. edlab.tc.columbia.edu/files/EdLab_Natriello4.pdf.

Odlyzko, Andrew 2013. Open Access, library and publisher competition, and the evolution of general commerce. Preprint. http://www.dtc.umn.edu/~odlyzko/doc/libpubcomp.pdf.

Oudshoorn, Nelly and Trevor Pinch 2003. *How users matter: The co-construction of users and technology.* Cambridge: MIT Press.

Pinfield, Stephen 2013. Medical research charities and open access. *Learned Publishing* 26(4): 285-302.

Research Information Network 2008. Activities, costs and funding flows in the scholarly communications system in the UK. Report.

Scholl, Hans 2004. Involving salient stakeholders: Beyond the technocratic view on change. *Action Research* 2(3): 277-304.

Shapiro, Lorna 2005. Establishing and publishing an online peer-reviewed journal: Action plan, resourcing, and costs. OJS project report. http://pkp.sfu.ca/files/ OJS_Project_Report_Shapiro.pdf.

Sjostrom, Jonas and Goran Goldkuhl 2009. Socio-instrumental pragmatism in action. *Handbook of research on socio-technical design and social networking systems* (ed, Brian Whitworth and Aldo de Moor). Hershey, PA: IGI.

Solomon, David and Bo-Christer Björk 2012. A study of open access journals using article processing charges. *Journal of the American Society for Information Science and Technology* 63(8): 1485–1495.

Solomon, David and Bo-Christer Björk 2012. Publication fees in open access publishing: Sources of funding and factors influencing choice of journal. *Journal of the American Society for Information Science and Technology* 63(1): 98-107.

Steiner, Gerald 2008. Supporting sustainable innovation through stakeholder management: a systems view. *International Journal of Innovation and Learning* 5(6): 595-616.

Suber, Peter 2012. *Open Access.* Cambridge: MIT Press.

Swan, Alma 2010. Modelling Scholarly Communication Options: Costs and benefits for universities. A report for the JISC. http://eprints.ecs.soton.ac.uk/18584/

Swan, Alma and John Houghton 2012. Going for Gold? The costs and benefits of Gold Open Access for UK research institutions: further economic modelling.   Report to the UK Open Access Implementation Group. *http://repository.jisc.ac.uk/610/2/Modelling_Gold_Open_Access_for_institutions_-_final_draft3.pdf*

Van Noorden, Richard 2013. Open access: The true cost of science publishing. *Nature* 495(7442): 426-429.

Walters, William H. 2007. Institutional journal costs in an open access environment. *Journal of the American Society for Information Science and Technology*, *58*(1), 108-120.

Ware, Mark and Michael Mabe 2012. The stm report: An overview of scientific and scholarly publishing (3rd ed). Report. http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf.

Waltham, Mary 2010. The future of scholarly publishing among social science and humanities associations. Report. http://www.marywaltham.com/JSPfulltextarticleApril2010.pdf.

Whitworth, Brian 2009. The social requirements of technical systems. *Handbook of research on socio-technical design and social networking systems* (ed, Brian Whitworth and Aldo de Moor). Hershey, PA: IGI.

Withey, Lynne, Steve Cohn, Ellen Faran, Michael Jensen, Garrett Kiely, Will Underwood, and Bruce Wilcox 2012. Sustaining scholarly publishing: New business models for university presses. Report. http://www.aaupnet.org/images/stories/documents/aaupbusinessmodels2011.pdf.