State of the IR 2016

June 2016 | AY 2016-2017

Andrew Weiss
Digital Services Librarian
I. Abstract
This report examines the collections, policies, general procedures and projects comprising the CSUN ScholarWorks Open Access Repository (SOAR) for the AY 2015-2016 and looks ahead to the coming AY 2016-2017 and beyond. The report begins by providing a breakdown of the repository as it is currently organized and populated. Overall the repository added approximately 6,800 items from April 2015 to June 2016, growing from 10,149 to 17,015; over 800,000 items were downloaded and 1.7 million records were viewed during this time.

Repository indicators of success are also examined and analyzed for their levels of effectiveness. The indicators of success examined include: mandates, planning integration, funding models, proximity to digitization centers, interoperability (including linked data), measurement, and promotion strategies.
The report also provides 10 future directions to improve the long-term viability of the repository. Rough timelines for short-, mid- and long-term planning goals are outlined and proposed.

Recommendations and goals for the coming academic year include the following:

- **Complete ETDs CQF project** (Phase IV completed by summer 2017);
  - Increase alumni involvement for larger open access ETD collections (write & submit CQF to fund student for this project for AY 2017-2018);
- **Continue outreach & collaborative efforts to reach faculty** participants;
  - Including targeted advertising strategy and design SOAR paraphernalia (pamphlets, postcards, etc.);
  - Improve Scholar Spotlight outcomes (in progress) using 1Scence’s oaFoldr citation data aggregation service;
- Collaborate with Chancellor’s Office to test next-generation repository platforms;
- Propose and begin initial development of a scalable and sustainable ‘Center For Digital Scholarship’ to address the multiple, growing needs of the CSUN faculty and to advocate for Open Access and library-centric digital publishing; [See Appendix A for brief proposal] focusing especially upon the following:
  - Design & implement a “Services Portfolio” comprising:
    - Repository Services (i.e.: Document Archive, OA Fund, Scholar Spotlight);
    - Data Management Planning (DMP) Services;
    - Faculty Publication Services (i.e.: OA journals & monograph publishing);
    - Scholarly Communication Copyright/IP Consultation Services;
    - University Archive Services;
  - Collaborate with and support other centers on campus, including MetaLab, Center for Digital Humanities (and others), for customized and prototype services and tools development.
II. Overview of Communities, collections & ongoing projects

A) The purpose of this document
This document is intended to provide details of the current state of the ScholarWorks Open Access Repository (SOAR) for the AY 2015-2016 and to provide room for speculation on future development. The document explains the various policies, ongoing projects and collections, and the overall digital ecosystem and infrastructure necessary to sustain a viable repository. The document also aims to examine current repository practices and projects at the university and at CSU system level, including a look at various stakeholders in the repository and an analysis of the CSU/CSUN relationship. Ultimately, the direction SOAR takes will be impacted by how successful projects are leveraged into new initiatives and how new policies affect participation behaviors. In particular, successful grant and CQF projects in tandem with campus and CSU system-wide mandates or resolutions will help to drive repository growth. This document aims to identify such potential projects as much as identify future trends.

B) The current state of the repository in raw numbers
The repository has existed at CSUN since the CSU Chancellor’s Office first offered the service in 2007. The first several years through 2011 showed low repository growth, but from Spring 2012 through the current spring 2015 the growth has reached consistently higher levels. As seen in the chart below, the repository has increased from 132 to over 17,000 items (Figure 1). A month by month analysis shows this growth as various projects have come into effect during this time (Figure 2). Along with the growth of collection sizes, there has also been an increase in the amount of internet traffic during the five-year period from 2011-2016. The repository has reached several milestones during this period. In August 2014, the repository reached 1 million downloaded files; in March 2016, the repository reached 2 million downloads (with some recent months reaching as high as 106,000 downloads) and 3.4 million page visits. (Table 1) SOAR currently averages nearly 70,000 downloads per month, though this is subject to somewhat dramatic swings. (Figure 3)

Figure 1: Archive total from May 2011 to June 2016. The repository has grown from a few hundred items deposited to approximately 17,000 during this period of time.
Figure 2: Items archived by month from May 2011 through June 2016.

<table>
<thead>
<tr>
<th>Date Range</th>
<th>Total Items archived</th>
<th>Total Bitstream views</th>
<th>Total item views</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 2010-March 2014</td>
<td>4,644</td>
<td>734,794</td>
<td>852,061</td>
</tr>
<tr>
<td>April 2014 – April 2016</td>
<td>12,371</td>
<td>1,416,831</td>
<td>2,796,170</td>
</tr>
<tr>
<td>TOTALS:</td>
<td>17,015</td>
<td>2,151,225</td>
<td>3,648,186</td>
</tr>
</tbody>
</table>

Table 1: Bitstream and item views from various time ranges starting from March 2010 to present time.

Figure 3: ScholarWorks item views (metadata records) and bitstream views (downloads) per month from May 2011 to June 2016.
Overall, the CSU system as a whole has collected over $136,000$ items in repositories across 22 participating campuses, 54% of which reside in DSpace (12.5% in CSUN SOAR). In comparison to the repositories of other California State University campuses, Northridge ranks third in total size (Figure 4). As of June 2016, Cal Poly San Luis Obispo ranks first at just below 32,000, while San Jose State ranks second at 23,200. Cal Poly and SJSU’s platform is bePress’ Digital Commons, a full-service, proprietary platform with certain outreach and tech support advantages for the price. Their combined collections account for 40% of the CSU system’s collected works.

Figure 4: Comparison of IRs across the CSU. CSUN currently ranks third in the entire CSU, but first among those campuses using the CO’s DSpace platform, comprising roughly 12.2 % of the overall CSU tally and 23% of the DSpace collections.
The items within SOAR are mainly related to the ETDs projects (including new ETDs and retrospective scanned manuscripts). Currently the largest communities in SOAR are ranked as follows:

1. Electronic Theses and Dissertations (N=12,069) 73%;
2. Open Access articles (N=1736) 11%; (includes Faculty Publications [N=824] & CSUN OA Journals and Newsletters [N=916])
3. CSUN University Archives (N=1,650) 10%;
4. CSUN Centers and Institutes (N=888) 5%

*Note: we are currently working through Physics department faculty CVs; the number of physics papers will increase (by 2x or 3x) in the coming months, given the typical rates of publication and the long-standing practice of publishing in repositories like arXiv or in open access journals.*
C. Growth factors
There are currently six main factors driving the growth of CSUN’s repository:

1. Staffing
2. Funding
3. Open access mandates
4. Targeted project development
5. Dedicated scanning center
6. Discipline-specific ‘normative’ values

First factor: Staffing
SOAR currently has three full-time staff / faculty working on the repository (Digital Services Librarian, Web Services Coordinator, and ScholarWorks Assistant). Each person focuses on several important aspects of the repository, including project development, content uploading, outreach and advocacy. Along with the full-time employees, there are 5 CQF-funded student assistants assisting with scanning, abstract transcription, and metadata entry; 1-2 catalogers in CAMS provide services for metadata alignment between the thesis catalog records and SOAR ETDs in Dublin Core; finally, several volunteers in various departments and centers across campus upload content for their respective groups. Overall roughly 15-18 persons work on the repository at various times, not including Chancellor’s Office IT staff.

In examining the staffing levels of other repositories in the CSU, it becomes clear that there is at least an association between the number of work hours dedicated to a repository and its overall size and long-term growth. The number of people working on CSUN SOAR across campus and in the library suggests that the continued development of the repository will depend upon having sufficient staffing numbers to handle project loads to reach established benchmarks.

However, to address upcoming project development (including data management planning, open access publishing, university archives, OA initiatives) larger numbers of participating faculty and staff will be necessary. The problem of scalability and sustainability of project levels is beginning to make its impact felt. In order for these various projects to grow, a larger staff will eventually be necessary. Once the ETD CQFs are complete, it might be prudent to hire one or two SOAR student assistants on a permanent basis.

Second factor: Funding
Repository growth has benefited from three Campus Quality Fee projects aimed at digitizing the CSUN master’s thesis collection. The first CQF, covering Phase I of the project, scanned 4,500 manuscripts during the AY 2013-2014. The second CQF, covering Phase II of the project, digitizing 450-500 at-risk analog media works, including VHS cassettes, audio cassettes, and reel-to-reel tapes. Phase III has scanned 5000 more and as of June 15, 2016 is 99% complete. A fourth CQF, covering Phase IV of the project, will digitize and archive the remaining 4,000 thesis manuscripts from the years 1989-1999 as well as 300 manuscripts for the National Leadership Training Program (NLTP). The funding covers the costs of unbinding books and cutting pages for high-speed scanning as well as the hourly wages for student assistants.
Although these CQF projects have been successful, a new approach to secure funds may be necessary in order to develop other pressing projects. External grant agencies may provide the funding necessary to start up important open access, digital publishing and data management planning initiatives. Funding may be necessary to start up a center on campus. An open access fund, for example, would also require money to implement, though the source of this could be from somewhere else on campus.

**Third factor: Open Access mandates and resolutions (on campus & beyond)**
The CSUN ETD mandate, which requires all master’s students writing a thesis to place a copy in ScholarWorks, has spurred repository growth. New theses comprise 11.5% of SOAR. The CSUN Faculty Senate Open Access Resolution has helped to spur more awareness of ScholarWorks as well. CSUN’s President Harrison is also a signatory to the Berlin Declaration on Open Access, while California also recently passed AB 609, a mandate requiring research that uses funding from the California State Department of Public Health to be published in open access repositories. Certainly, more must be done to leverage these mandates and resolutions, but the all-important foundation for open access has been established.

**Fourth factor: Targeted project development**
Multiple projects have helped to increase the repository’s robustness. The aforementioned ETD projects, both current and legacy, account for roughly 73% of the collection. The University Archives collections will improve the development of the repository as well. The gathering of archival documents such as standing committee meeting minutes, agendas, and reports stands to increase dramatically over the next year or two.

Finally, Scholar Spotlight remains a viable source of content, though the problems of participation hinge on copyright clearance as well as securing pre-prints and post-prints of work from faculty directly. Partnering with Metalab on the faculty app (adding ORCIDs) may help raise the importance of ScholarWorks & open access works. Also, relying on third-party vendors to help aggregate CSUN-related open access publication citations will improve participation rates in the Scholar Spotlight program. *oaFoldr* should help to gather a list of citations SOAR staff can add to the repository.

**Fifth factor: Scanning Centers**
The development of collections has also been driven by the use of high-speed scanning devices. The ETD retro digitization projects use high-speed copy machines for scanning hundreds of thousands of pages. A separate copy machine was secured from the Physical Plant Management / Quick Copies for dedicated high-speed scanning projects. Upcoming University Archives projects will provide us with content in the coming years. Examining the issues related to mass digitization will be essential to increasing the repository’s size and impact.

**Sixth factor: discipline-specific “normative” values**
Every discipline has a sub-set of values that are considered ideal for constituents to follow. These will differ across the humanities, social sciences, mathematics the so-called “hard” sciences, and even the “dismal” science of economics. STEM fields (Physics and mathematics, in particular, and Biology to a lesser extent) have already embraced open access as an ideal on a national and international scale. This has helped SOAR to develop these collections to a greater extent than, for example, history or other
humanities-based disciplines that still rely more on the monograph for scholarly communication. Collection growth is driven by a wider philosophy and the degree to which it has filtered into the normative values of a specific discipline. As evidenced in basic work with Scott Kleinman and the Center for Digital Humanities, DH promises to break down some of the print-based traditions of the humanities fields; collaboration and exploration in new models will mark future growth in DH-library endeavors. A spirit of openness should drive much of the future repository growth in these cutting-edge disciplines.

D. Growth projections
The growth of the repository will mainly depend upon the factors listed above. In terms of collection numbers it is expected that by the end of 2016 or early 2017, SOAR will surpass the 20,000 item level. With the development of other projects including the University Archives, the increase in faculty participation (via Scholar Spotlight and 1Science’s oAFoldr), and the development of online journal publishing the repository could reach the 30,000 to 40,000 level in the next several years. SOAR will rival some of the larger single-campus repositories at R1 institutions in the United States by this point.

However, considering repository development from a consortium perspective, collection growth is strong for us but wildly uneven across the whole CSU. In combination with the other 22 campuses using a repository of some sort (DSpace or Digital Commons), the CSU system has surpassed 136,000 items (currently CSU DSpace repositories contain 75,000 items). CSUN comprises more than 12% of the overall CSU collection. But many peer campuses (i.e. Fullerton, CSU Long Beach, CSULA) lag in terms of collection development, staffing and funding.

This is not merely their problem since our success hinges upon the success of other repositories. I believe that greater interest in open access across all CSU campuses will improve our own local participation rates. As a result, the Chancellor’s Office as well as individual campuses should be making concerted efforts to market the CSU repository system to a much wider audience and create cross-campus repository initiatives and projects on the same level of complexity as the California Digital Library. They should consider stronger open access outreach initiatives by embracing a system-wide open access and copyright policies (i.e. faculty deposit mandates), by indexing all CSU ETDs in NDLTD or similar service, and by pushing for a system-wide open access journal/monograph publishing platform.
III. Indicators of institutional repository success

This report also outlines seven major indicators of success for an institutional repository. Each segment will examine what Oviatt Library is currently doing well and discuss what needs to be worked on to improve these indicators.

Indicator 1: Policy mandates and resolutions

Overview: The two most significant policies on campus are the CSUN ETDs mandate from May 2012 and the Faculty Senate Open Access Resolution (FSOAR) from November 2013. These two policies help to bring more content to the repository. CSUN president Dianne Harrison also has signed the Berlin Declaration on Open Access, pledging university support for the movement.

Positives: The ETDs mandate is incredibly successful. Due to the nature of the mandate and the requirement for students to submit a digital file for graduation, the ETDs mandate has created sustainable long-term growth for the repository. FSOAR has helped to raise awareness of the problems associated with scholarly communication and the need for open access publishing on campus — especially in light of the problems associated with rising journal costs.

Caveats: Because FSOAR is an endorsement rather than a mandate for open access, it imposes little pressure on the adoption of open access faculty publishing practices. In contrast, the entire UC system has an OA mandate that is entirely opt-out. As a result it is more effective.

Recommendations: A next step at CSUN would be to develop a stronger resolution or mandate to require faculty participation, though this is admittedly easier said than done, and may be the subject of a fight for a later time. Another step would be to advocate for a wider CSU mandate or resolution on open access requiring all CSU faculty to submit published articles into a green OA repository, to which 22 campuses have at least some access.

In the meantime, outreach and content acquisition initiatives would be useful. These would include: development of participation incentives such as department rewards (including swag, travel money, etc.) and an Open Access Fund; creating more CV mining techniques, including a periodic check for Open Access publications related to CSUN through new vendor 1Science's oaFoldr; and creating CSUN repository profiles with full citations and links to SOAR (especially in tandem with Metalab Faculty App and faculty ORCID registrations).

Other areas for mandates or clear policies could include: Course syllabi posting, University Archive development mandate (including standing committees, papers, etc. from all units on campus); open journal publishing mandates; data set publication mandates; better and more programmatic alignment with NSF/NIH/NEH/CA AB 609 mandates.

Indicator 2: Planning integration

Overview: Planning for SOAR requires a significant amount of policy alignment mixed with the development of new initiatives and plans. In order to do this, awareness of college and department needs as well as centers, initiatives and other “content generators” is imperative.
Positives: Currently SOAR provides services for the development of grants (working with the Office of Research and Sponsored Projects); ETDs support; retro ETDs digitization development (for increased ADA compliance and for format migration and digital preservation initiatives/projects). Overall the repository is both adhering to university policy as well as acting as a force for campus change.

Caveats: More can be done to improve SOAR’s alignment with university goals as well as to help shape policy itself, especially in the realm of “proof of concept” for such things as faculty profiles (similar to Digital Commons’ faculty profile pages), development of workflow for automated self-submission of materials to the digital university archive, development of journal publishing, increasing open data management plans, and sharing of learning objects/teaching tools.

Recommendations: Increase opportunities for discussions with campus administrators, including college deans, department chairs, center leaders, chairs of standing committees, and other units on campus. Develop an outreach strategy such as monthly/bi-monthly automated on ScholarWorks, as well as consistent lectures/discussions and sit-ins at various meetings. A greater sense of the practice and policy directions for the campus will surely arise from staying aware of the interests of the various campus stakeholders.

Indicator 3: Funding models
Overview: The CSU Chancellor’s Office provides financial and technical support for the DSpace repository. CSUN Oviatt Library also receives funds from various projects for digitization in the form of Campus Quality Fees and other grant-funded projects.

Positives: One of the main advantages to using the current Chancellor’s Office DSpace installation is the funding support. CSUN is not required to front money to implement the platform, maintain its development, or carry out upgrades. This allows the repository staff/faculty to focus on developing collections, providing services, and conducting outreach for policy and general repository awareness. The CQF thesis digitization projects have also been successful in digitizing over 10,000 CSUN masters’ theses. Nearly $75,000 has been awarded for the past 3 years; $24,000 is pending for the final CQF.

Caveats: Reliance on outside funding is dependent upon multiple factors that are not under the library’s direct control. Planning ahead across several years is required to ensure that projects are not interrupted due to sudden or even anticipated loss of funding avenues. Yet given the nature of the CQF model, the projects also must be completed within academic year cycles.

Recommendations: More funding could be secured for various projects, including Open Journal System development, Data Management Planning, University Archives, and faculty publications. Overall, a look at the ROI or costs per upload and costs per download are also necessary to further examine whether funding models are effective, sustainable and/or scalable. Additionally, it might be helpful to undergo a cost analysis of the current funding levels at Oviatt library dedicated to CSUN SOAR collections. This would provide important data to understand the optimal amount required for sustainable repository development.
Indicator 4: Proximity to digitization centers

Overview: Repositories that are integrated with digital processing equipment and workflows that generate and/or collect digital content tend to be better populated.

Positives: Over the course of the two ETD CQF projects the use of high-quality, high-speed copy machines has greatly facilitated the volume of scanning. This past year saw the removal of the ETDs scanning equipment to a more dedicated space in CAMS. Two digitization stations have also been set up in multipurpose cubicles in CAMS to help complete the project. As a result, the CQF Phase III is on pace to be completed in mid-May, roughly six weeks ahead of schedule.

DSpace allows for decentralized content uploads as well. This is helpful for such projects as the new ETDs, the Hybrid Math Repository, the Office of Assessment and Program Review, and Faculty Senate standing committee documents. It should be considered axiomatic that the more content collection points there are, the faster and steadier the IR growth will be.

Caveats: Scanning and submission projects are managed in various places within the library and across campus. This can contribute to miscommunication, and/or lack of awareness of what others are doing. Decentralization can sometimes be double-edged.

Recommendations: Digitization centers are moving at various paces and with varying levels of commitment. It might be necessary to coordinate and manage efforts more closely as the number of participants in adding content to the repository increases. Periodic meetings and tutorials are necessary. It might be necessary to create a “ScholarWorks User Tutorial Packet”, to help refresh users’ memories.

Indicator 5: Interoperability

Overview: The DSpace platform allows various external organizations to harvest the metadata and provide it openly online. Google, Yahoo, and other search engines are able to access the information. Additionally, open access registries such as ROAR and DOAJ are also able to access the metadata.

Positives: Currently SOAR records appear in the open web and are easily accessible. DSpace logs of record and item views provide a look into the amount of traffic that finds its way to SOAR.

Caveats: More can be done to increase the profile of CSUN publications and documents in ScholarWorks.

Recommendations: The possibility of creating linked enhanced data should be considered. Interoperable identifier systems such as ORCID (Open Researcher and Contributor ID) should be investigated for possible integration with SOAR/Metalab. For increased exposure for our ETDs, we should contact Networked Digital Library of Theses and Dissertations (NDLTD) to have them harvest our metadata for inclusion in their massive digital library of over 4.4 million records.

Indicator 6: Measurement

Overview: This area examines the usage and growth of the IR and provides quantitative data in order to identify viable and desired future services for users. This would include the gathering of usage statistics
(as evidenced in this report’s various figures and tables) as well as the creation of usable feedback loops for users, especially those that have submitted content and who would like to see the specific details of its use.

**Positives:** The statistics being collected can be used to promote the IR to faculty and departments across campus. Those interested in the development of ETDs, for example, can see the impact that a digital medium provides. Furthermore, with the increasing application of alt-metrics SOAR can be used as another tool to provide a more accurate understanding of a scholar’s overall impact.

**Caveats:** The availability of statistics in DSpace for end users is still limited. Currently, a faculty member or student interested in finding out the number of downloads or access statistics must request it from ScholarWorks staff. This is time consuming and causes delays for the users.

**Recommendations:** Advanced Google Analytics techniques should be applied to more easily collect information about user accesses of specific types of content and content of individual authors. Apps or widgets should be developed that facilitate visualization of GA and other content usage data. Most faculty would be curious to see the number of downloads for their works, from where they were accessed, and so on. Creating a digital download/statistical dashboard could generate greater interest in ScholarWorks itself; it would also provide opportunities for faculty on the tenure track to improve their PIFs and to demonstrate alternative metrics for scholarly impact. SJSU’s data visualization widget with Digital Commons is a good example. (link: [http://scholarworks.sjsu.edu/](http://scholarworks.sjsu.edu/)), though other models and applications, such as those related to Google Analytics, should be investigated.

**Indicator 7: Promotion & project facilitation**

**Overview:** Various strategies have been developed to make the IR and open access practices better known on campus. The Scholar Spotlight service, ETD (current and retrospective collections) orientation and the Open Access week event, as well as recurring copyright summits, are a few of the initiatives in place that have helped to promote the IR.

**Positives:** The ETD project has put SOAR on the map among campus entities. It generates good will among faculty and administration based on its stable online presence and unique service. It helps to streamline the thesis process as well as promote the variety of projects that CSUN faculty and students are working on. Scholar Spotlight has also provided faculty with a chance to share work. Working on University Archives has also helped to improve the profile of SOAR. As the university moves dynamic web content into Web-One and revises the content of its pages, the repository is providing a stable online location for digital documents and other assets.

**Caveats:** Some problems with ETD have come to light both from the technical standpoint – inaccurate records, glitches in the system, etc. – as well as from the administrative side, including embargo issues, dislike of publishing low-quality content, and ongoing S-factor funding issues. These problems need to be addressed in discussions with faculty, staff and administrators. Scholar Spotlight also runs into issues when faculty publications cannot be added to ScholarWorks because of publisher copyright restrictions. A few faculty members have objected that their SOAR profile/collection contains only a few items. Some also balk at the thought of submitting early drafts of works (pre-prints and post-prints).
Recommendations: Outreach needs to be conducted carefully. False promises may be worse than no promises at all. A review of the ETD policies and workflow is warranted. Working directly with the owners of the ETD workflow application in the Matador Emerging Technology and Arts Laboratory (META Lab) to improve project and system organization would be worthwhile.
IV. Future directions/conclusions
The following recommendations are proposed:

Direction 1: Develop a center for digital scholarship
It is believed that a center for digital scholarship can help to meet the increasingly complex and growing needs of 21st-century scholars. The center would provide support for a suite of services related to scholarly communication in the digital era, including:

- Repository support services (especially self-archiving & green OA);
- Data Management Planning services;
- Publication services (i.e. providing a university press);
- Copyright / IP consultation services;
- University Archives Services;
- Student digital scholarship services (esp. thesis support)

The key priorities of the center would be an emphasis on scalability (how to address larger numbers of clients) and sustainability (how to provide services for extended periods of time). It is believed that centralizing these services and organizing them as an interlinked, interdependent portfolio of services may help to streamline the process and provide more efficient results. [A more detailed rationale, outline, and discussion of plans appears in Appendix A.]

Direction 2: Increase outreach
Outreach needs to be implemented in a more organized manner. Perhaps patterns of outreach can be established as part of an overall whole. The guiding principle should be face-to-face discussions of SOAR and its potential to serve all faculty members on campus. Outreach can be implemented in a stepped fashion, such as group meetings with college deans, then with faculty chairs, then with faculty on an individual, consultative level. A two-year plan to meet with various departments should be drawn up and implemented, focusing on the STEM disciplines and then on social sciences and finally with humanities. These meetings would be supported with handouts and slides provided to all participants.

Direction 3: Advertising and paraphernalia
Along with a schedule for outreach efforts regarding ScholarWorks, advertising and “swag” should be developed to give faculty a greater awareness about SOAR. In particular postcards for all faculty as well as buttons or coffee mugs should be created to give to faculty for reminders of SOAR. The costs should be relatively small, perhaps $500-$1000 per year, but will provide important reminders as well as increased awareness.

Direction 4: Financial and organizational incentives
Incentives for faculty participation should include department-level and individual incentives. On one end of the scale, departments with the largest number of participants could be awarded a certain amount of money to go toward open access fees or toward a department event/activity. At the individual scale, faculty could be provided with a certificate or an award (with a stipend of some kind). These incentives could spur more participation in the repository. Continuing the faculty CSUN Open Access Award would also be recommended as well as funding the Open Access Fund.
Direction 5: Data management planning
One of the more important developments in the past few years has been the increased demand for open data. It is recommended that a DMP working group be formed to better address this campus-wide need. The group should include interested library faculty, research faculty external to the library (especially Biology and other data-intensive disciplines), and the Office of Research and Sponsored Projects (ORSP). Some consultation with ORSP has already occurred. Development of policies, implementation of tools and the creation of best practices and channels of consultation should be established.

Direction 6: Increased & improved Scholar Spotlight services
Scholar Spotlight remains a vital aspect of ScholarWorks even as a few obstacles have prevented it from being as effective as it could be. CV “mining” has occasionally proved to be stymied by the copyright regulations of publishers that allow only pre- and post-prints of versions. To improve services, a more targeted approach may be warranted. Rather than seeking CVs, the approach should be to add articles in open access, create collections for CSUN faculty automatically, and then notify faculty. This approach will help to alleviate the problem of “false promises” that arises when copyright restrictions impede faculty from self-archiving their work. We will accomplish this by using a list of CSUN faculty open access publications generated by vendor 1Science with its service oaFoldr. The first test round was completed in May 2016 by adding 86 faculty publications taken from a list of 100 generated by oaFoldr. The number of CSUN faculty participating in SOAR increased from 84 to 130 by using this service. Full scale implementation will require significant time and resources, but would amount to a greater increase of faculty open access publications appearing in SOAR.

Direction 7: University Archives Development – full-scale project implementation
As part of the charge of the University Archives Group, SOAR has been gathering the minutes, agendas, and so on of the various Faculty Senate Standing committees. This is just a tip of the iceberg. A larger-scale attempt at gathering / uploading the materials will need to be undertaken. This will be done in partnership with SC/A. Full-scale implementation will require partnerships with other departments and their staff in order to facilitate large-scale uploading. An implementation group should be developed and the possibility of applying for funding from grants should be investigated.

Direction 8: ETDs: Phase IV and Alumni Outreach
The fourth and final phase of the ETD digitization project will involve theses from 1989-1999, which totals 3,985 manuscripts, most of which will be housed in our partial access collection. The success of the ETDs collections, however, hinges upon the amount of open items. We propose to work in tandem with the office of alumni to begin securing permissions for open access. A pilot for the year 2011 was conducted resulting in 10% participation (N=40; of which 90% said OK to OA). A CQF to fund a part-time student assistant might be worthwhile considering in order to pursue this further.

Direction 9: Journal publishing development – full-scale project implementation
Currently OJS and OMP remain in the testing phases. A few test journal issues have been examined, but a full-fledged examination as well as a testing of the back-end roles for online digital publishing journals has yet to be accomplished. A more directed open journals publishing development plan should be
implemented. The next few years should also focus on more digital content generation and publication, focusing on partnerships with current publications and with those who express interest in developing new ones. A time line and a campus-wide implementation group, which might include the Faculty Technology Center and other campus entities, should be established. The goal should be to secure grant funding for the development of CSUN Open Journals and a new CSUN university press.

**Direction 10: New repository platform investigation and development**

In the past year the Digital Sandbox group examined Hydra, Islandora and Fedora as possible future replacements for the ScholarWorks and CONTENTdm systems, as well as possible platforms on which to archive born-digital collections. Concurrently, the Chancellor’s Office is examining Islandora and Fedora with the aim to install a next-generation repository system by 2017. We will continue to provide support to the CO for implementing a comprehensive repository platform (one that can incorporate images, documents, multi-media formats, datasets, etc.).
V. Planning Implementation timelines: goals and benchmarks

Implementation 1: Short term goals and benchmarks (1-3 years out) | 2016-2018
1) ETD CQF Phase IV (completed by summer 2017);
2) Develop comprehensive Digital Scholarship Services Portfolio including:
   a. Data management planning service (DMP) tools and advisory group;
   b. University archives service (involving various admin units in partnership with SC/A);
   c. University publishing service (OJS/OMP/Open Conference ready by 2017; full-fledged journals by 2018; supported partly by seed grant proposals or CQF funding);
   d. Open access & copyright consultation service (for open access advocacy (including OA Fund) open access rights negotiation & general copyright assistance);
3) Propose & develop plan for Center for Digital Scholarship to help address scalability & sustainability issues stemming from the services portfolio listed above in 2);
4) Increase advertising & outreach: postcards, logos, mugs, etc.; email & face-to-face correspondence; include making rounds to all CSUN departments via department liaisons;
5) Collection benchmarks:
   a. 20,000 item level (estimated by spring semester 2017)
   b. Greater faculty participation: 200+ faculty; (includes all librarians; all Physicists; and a majority of faculty in the other STEM fields; ~25% of tenure-track/tenured faculty);
   c. Increased number of datasets added; serve as default repository for some CSUN centers

Implementation 2: Mid-term goals and benchmarks (3-5 years out) | 2018-2021
1) Center for Digital Scholarship fully established and seeking various avenues of funding for its long-term sustainability; focusing on:
   a. Funding for journal publishing – ongoing campus and external funding for CSUN Open Journals;
   b. DMP funding grants – ongoing campus and external funding for support of Data management;
   c. University Archive structured and organized; UA self-archive workflows and procedures;
   d. Financial incentives – ongoing campus and external funding for Open Access publishing (Gold fees and “Green stipends” or “Green OA Department Awards”);
2) Improved feedback loops: data visualization dashboard, faculty profiles (RSS feeds to/from SOAR), auto notification of faculty profile changes;
3) Collection benchmarks:
   a. 30,000+ item level
   b. 400+ faculty (~50% of tenure-track/tenured faculty)

Implementation 3: Longer-term goals and benchmarks (5-8 years out) | 2021-2024
1) Center for Digital Scholarship fully funded from various funders and organizations (internal & external);
2) New repository platform development and implementation;
3) Collection benchmarks:
   a. 40,000+ item level
b. 600+ faculty (~75% of tenure-track/tenured faculty)

Implementation 4: Ultimate goals (8+ years out) | 2024 and beyond

1. *All* tenured/tenure-track CSUN faculty members registered and self-archiving with SOAR;
   a. Include adjunct faculty (a number that will likely grow in the next decade) to help improve their employment opportunities;
2. *All* faculty members participating in open access as a default rather than as an option via CSU system-wide open access mandate or other national law;
3. Inter-campus repository consortium— a master unified CSU archive & provider of repository services on a level comparable to the UC system’s *California Digital Library*. 
VI. Appendix A: Proposal for the development of CSUN Oviatt Library Center for Digital Scholarship

Part I: Introduction:
The proposal for a Center for Digital Scholarship stems from the development of a number of policies, mandates and practices originating outside of CSUN's campus, on the campus itself, and from within the library. Future speculation is always risky, but several trends have become self-evident and are resulting in practices no longer sustainable for a majority of institutions. Monopolistic price gouging from the large database publishers has instigated adaptations to the publishing model in universities and other public-serving organizations. The resulting state and federal open access mandates now require faculty to rethink their approaches to grant writing, scholarship and data dissemination. Forward-thinking libraries, in order to simultaneously avoid excessive pricing models (and alleviate tightening budgets) as well as to meet the grant-funding needs of their institutions, have branched out from the long-standing triad of information management (defined as to collect, preserve, and provide access), and are adopting a model that crosses into what was traditionally the domain of publishers, content distributors, and even scholars themselves. (See figure 1 below.)

![Figure 1: Parallel lifecycles: scholarly & information lifecycles (CDL 20xx)](image)

In response to these new developments, collection development and cataloging librarians are also moving from a stereotypically passive role in the information management lifecycle toward a more active service role by providing avenues for publishing. They are helping to bridge the gap of the information and scholarly lifecycles by providing and managing discovery and creative tools for scholars. Such types of tools and resources include digital archives and collections, open access repositories, open journals and monograph publishing, Open Educational Resources (OERs), data archiving and data management planning. Some are even providing “maker spaces” to facilitate the creation phase in scholarly lifecycle. Along with these new digital resource avenues, copyright and other intellectual property concerns have become essential for faculty members to understand in order to be active and empowered scholars in the digital age.
As these needs for services catering to digital scholarship issues continue to grow, the problem of scalability and sustainability looms in the background. The types of services now being offered by many libraries includes the following: ETD management and publishing, open access repository development, digital archive collections creation & distribution, data management planning, copyright and intellectual property consultation, and open access digital monograph and journal publishing. Yet it is becoming difficult to provide this full suite of services for large numbers of constituents. These interrelated services and projects are often offered in piece-meal fashion as the funding and time and demand for services warrants it. However, to more effectively manage and provide these services – and to provide them as a coherent bundle or portfolio of interrelated items – it will be necessary at some point to consolidate the management of them.

It is proposed, then, that a center for digital scholarship with the mission to provide these various services to faculty and students be developed to meet the needs and challenges outlined above. The benefits of this will be improved scalability to reach larger numbers of faculty and greater sustainability to make sure the services offered can be offered for the foreseeable future.

Case studies: precedents and peer institutional models
Increasingly, centers are becoming commonplace in the academic library (See part VI for a partial list). The Library Publishing Directory 2016 lists 105 library-centered publishing endeavors in North America; among these, 43 of these are placed squarely within centers for scholarly communication. The range of services provided by these centers varies, ranging from copyright consultation and typical scholarly communication issues (i.e. OA, IRs, etc.) to full-fledged digital curation and digital humanities research support organizations. Some centers have even incorporated media technology and learning commons functions.

This proposed center for the Oviatt Library, however, would focus purely on the types of services traditionally bundled under the topic of Scholarly Communication (copyright, data management planning, gold OA publishing support, green OA repository services, and library-centric digital publishing), rather than special collections and archives, the Creative Media Studio, or other aspects of digital services already covered by an existing area.
Part II: The "Portfolio of Services" approach

One of the precedents for this approach is the UC system’s California Digital Library. The CDL has led the way in providing digital scholarship services and open access advocacy for the past twenty years. Since 1997 their projects have grown from a platform-centric digital library to a suite of services essential for University of California faculty and students to thrive in the digital era. In particular CDL’s UC Curation Center, a set of repository and information curation services, provides a great model for this proposed bundle of services. In their planning documentation, they describe their actions as “micro-services,” focusing on meeting user needs rather than merely providing platform capabilities. In other words, the service determines the tools used, not vice-versa.

In the table below the types of micro-services provided for the repository are listed. In all but two, Characterization and Annotation, the ScholarWorks repository can handle each of these actions. The difference, though, is that UC3 provides these as a set of procedures and on-demand services rather than automated functions of a digital tool. Therein lies the important distinction. Instead of relying on the automations of a tool, UC3 relies on staff and people to provide that necessary service point.

<table>
<thead>
<tr>
<th>Identity (S)</th>
<th>Storage (S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixity (S)</td>
<td>Replication (S)</td>
</tr>
<tr>
<td>Inventory (S)</td>
<td>Characterization</td>
</tr>
<tr>
<td>Ingest (S)</td>
<td>Index (S)</td>
</tr>
<tr>
<td>Search (S)</td>
<td>Transformation (S)</td>
</tr>
<tr>
<td>Publication (S)</td>
<td>Annotation</td>
</tr>
</tbody>
</table>

Table 1: Services provided by UC3; the (S) indicates comparable system capabilities in ScholarWorks.

While it’s important to have a repository platform available to do such tasks, the scalability and sustainability of these actions may not be possible if relying on just software and a few staff members. The UC3/CDL hires a significant number of employees to handle each of these micro-services.

Along the same lines it will be important to develop a suite or portfolio of services that emphasize meeting client needs rather than relying primarily on the tools to do the heavy lifting. As has long been the case when working with siloed systems such as CONTENTdm, DSpace, and many others, once we begin to rely on specific tools for completing our work, they often start to dictate policy and workflows based on their system capabilities. The process is somewhat backwards as a result because policy sometimes becomes shaped to meet the needs of software and not the needs of constituents. Platforms are only as effective as the support staff managing them and the policies guiding them. UC3’s example demonstrates that to be effective on larger scales and meeting longer-term needs we must to be meeting user-driven needs with actionable services, independent of the tools and their capabilities.
Part III: The “portfolio”
The services offered by the Center for Digital Scholarship would include the following:

- Repository services: including digital document archiving & preservation, OA Fund, Scholar Spotlight, Department OA Awards, &c.;
- DMP Services (Developing greater data management planning initiatives and establishing a cross-campus DMP work group);
- Faculty Publication Services (OJS/OMP); Santa Susanna Press; cross campus group; [report forthcoming on development from DPIG group]
- Copyright / IP consultation services on scholarly communication;
- University Archives Services (in collaboration with SC/A);

A. Repository services:
This set of services is related to the repository and the scope is within the areas of Open Access, especially green open access self-archiving, support for gold open access, and the organization and dispensation of annual open access awards.

The specific services provided to faculty would include: Scholar Spotlight, which mines faculty CVs for citations; OA fund, which provides funds for APCs in select Gold OA journals; tutorials and outreach demos for using ScholarWorks; OA Awards committee, which would organize the annual OA award for CSUN faculty, staff or campus organizations showing the most support for the OA movement; digital preservation, which would ensure long-term viability of archived digital documents.

The staff needed for this are mostly in place currently. There might be need of adding interns or student assistants to handle any influx of CVs or in updating the repository with the latest publications. For other partnerships, such as the University Archives, designated representatives in departments or colleges would need to be identified as submitters and trained.

Funding is sufficient to handle all but the OA Fund, which would depend on money coming from as-yet-undetermined sources on campus – and might be handled outside the scope of the center.

B. DMP services:
This set of services is related specifically to providing assistance to faculty for creating grant-mandated data management plans. Establishing a DMP work group; a partnering w/ Developing greater data management planning initiatives and establishing a cross-campus DMP work group

The services provided will offer DMP drafting and provision of documentation for improving plans and writing plans tailored to specific grant funding agency requirements. Currently documentation and web guides have been designed to help facilitate data management planning. ScholarWorks exists to provide an accessible archive for the data.

Current staff devoted to this are two librarians and others in the ORSP. More are needed and training and outreach materials need to be developed to increase outreach as well as support of DMP development.
The amount of funding would be related to the demand for DMPs. Likely this is to go up over time. At the same time it would be prudent to partner with ORSP to find cost-effective ways to approach this.

**C. Faculty OA publishing services:** This set of services concerns “born-OA” publishing through open access publishing platforms, including Open Journal Systems, Open Monograph Press and Open Conferences; the service would be aligned with the redevelopment of the Santa Susanna Press as a library-based university press.

Services provided will offer platform space & tutorial works for starting CSUN open access journals. We will also provide display space for currently published journals and their back-issues.

Staff might be a big problem for this one. We would need people to provide back-end support for the OJS/OMP/OC platforms. We would need to provide instruction on how to use the editorial functions for those wanting to develop a full-fledged journal. We would need to hire student assistants to help with content uploading and issue designing. The running of specific journals would need to be handled by faculty/groups and would likely be outside our area of expertise. People necessary would include: editors, peer reviewers, proof readers, issue designers and so on.

Funding would need to be secured for long-term sustainability. This would support administrative staff of the platforms and any student assistants. A campus quality fee proposal might be worthwhile if a suitable student-centered journal publication could be recruited for collaboration.

Details are forthcoming in a report from the Digital Publishing Implementation Group for developing this service. The report will examine various necessary aspects to implement a full-service library press from initial development phases to proof-of-concept to final official implementation.

**D. Copyright & IP consultation services** Copyright and IP consultations are a necessary aspect of scholarly communication. The university and the library need to provide these essential support services to help scholars maintain their rights for scholarship published.

This set of services would provide, as it is currently done with the ©-Team, information and consultation on copyright and IP issues related directly to scholarly communication: especially publishing rights and transfer, negotiation of open access clauses, use of materials in classroom settings, articles, and other publishable media.

Staff are essentially in place and aligned with the Copyright Team.

Some funding would be necessary aside from annual/bi-annual hosting of copyright events for faculty and students. Funds would help to bring in speakers prominent in their fields.

**E. University Archives services** University colleges and departments are mandated by the CSU to retain documents. Many have not been doing so. As a result a large amount of uncollected documents lies hidden across CSUN’s campus. This service would help to gather them and place them in ScholarWorks. Currently being employed on a
small scale with Faculty Senate standing committees. The issue will be increasing the scale and scope of the project.

The services provided would include outreach for demonstrations and training, helping to designate department/college submitters, and checking and updating metadata of submitted articles. A retention schedule and strategy for developing the collections would also be provided.

Staffing would include Archivist from SC/A, and SOAR staff. Much of the work would be done by designated outside of the center. A customized online submission form (as designed by the CO) might streamline the process.

Little funding would be required for this scale up. It would depend mostly on recruiting designated submitters and implementing a suitable bulk submission form.
Part IV: The four-tiered approach to digital scholarship services

Vinopal and McCormick (see Figure 2 below) propose developing digital scholarship services by focusing on tiers of research services. In the diagram below, they propose that university campuses and their service centers provide the bulk of Tier 1 and Tier 2 digital scholarship research services. Tier 3 and 4 services are defined as the cutting-edge, proof-of-concept, custom-designed and research-centric tools meant to meet very specific campus and scholar needs. Often these are aligned with grants for R&D or for very specific projects within a campus. The proposed Center for Digital Scholarship would attempt to focus on maintaining and scaling up these standardized Tier 2 services while simultaneously collaborating with other entities on campus to develop newer Tier 3 or Tier 4 services.

Currently the university and the library already provide for a number of faculty and student digital scholarship needs at the first and second tiers, including LMS (Moodle via ATC), web pages (Drupal via Web Services & Web One), multimedia production (CMS), online exhibits (CONTENTdm), profile pages (Drupal), streaming services (Video Furnace, et al.), file storage and email (IT), and e-book publishing (ATC). A few examples of previously developed Tier 3 and 4 services would include the ETD online submission form which started as a master’s thesis project and is now incorporated into Tier 2 as part of SOAR; the batch file uploading system for ScholarWorks developed at the Chancellor’s Office; and the Faculty App, currently in development at MetaLab. In these cases Tier 3/4 projects were requested or supported by Tier 2 providers and eventually incorporated into this standard portfolio of services.

On a final note, Tier 2 services should be scalable to accommodate ever growing numbers of users. It could be argued currently that very few of our services for the repository meet this scalability requirement, with the exception of the ETD project and the CO Batch file uploader. Other projects remain smaller in scope and may require enhancement to meet larger numbers of potential participants/users – especially Scholar Spotlight.

Figure 2: Proposed model for digital scholarship services (from Vinopal and McCormick).
This multi-tiered approach should be considered essential for the development and organization of the center and the long-term sustainability of these services. The center would provide Tier 2 Standard Research Services to all faculty within this limited scope of scholarly communication (i.e. the portfolio of services suggested above), while actively seeking to develop Tier 3 and Tier 4 services in collaboration with various entities on campus and at the Chancellor’s Office to ensure scalable services.

We should also look to tools developed by the CDL to see if they can be either adopted (for a fee, such as the DMP Tool) or copied and modified for local purposes. This would ensure the efficacy and impact of the center. It will also limit our scope and prevent undue mission creep. This will provide flexibility as well. The center could focus on Tier 2 efforts while also suggesting directions for potential research partners to take, as in the case with ETD submissions and the collaboration with MetaLab, or in the case of developing the Faculty App, where we can suggest ways in which to integrate current standard services with enhanced services or applied R&D.
Part V: Strategies for Development

In order to develop this center a number of things must be done first, including an examination of peer institutions and other centers for modeling, an analysis of CSUN faculty needs regarding digital scholarship, a literature review to understand the wider needs of faculty nationally and internationally, and an investigation into the funding and staffing necessary to create such a center.

Examination of peer institutions and model centers

Examining past and current models for digital scholarship would be an important initial step. Examining what services are essential and what strategies are necessary for development must be considered.

Gap analysis of faculty needs:

Another initial step to developing a center would be to assess faculty needs regarding digital scholarship. A gap-analysis of faculty needs in the area of digital scholarship might be conducted. The survey would try to examine all of the actions and important software along with faculty perceptions of what might be missing from their workflows. It would focus on the broad areas of: copyright consultation, green open access, gold open access, grant funding, open data and data management needs.

Literature review

Ancillary to this would be an analysis of the literature examining faculty needs in digital scholarship. We can surely extrapolate local needs based on the results of needs found in other studies.

Investigation and analysis of funding and staffing needs:

Finally, a more detailed analysis of funding needs for each part of the Service Portfolio would also be in order. Investigation of past models at other institutions is essential. Staffing needs, research and development to meet faculty expectations, and implementation of Tier 2 services will be examined. However, it appears that most centers and digital publishing endeavors are funded primarily through their libraries’ existing budgets. According to the Library Publishing Directory 2016, 74 of the 105 (70%) library operating budgets fund at least 90% of the center/publishing costs. Only 9 institutions received grant funding. Among the roughly 40 actual centers for scholarly communication (as opposed to projects based within a traditional technical services or digital collections department), only 5 (Iowa State, Illinois Wesleyan, Simon Fraser University, SUNY Geneseo, and the University of New Orleans) received funds greater than 50% from somewhere other than the library. It should be noted that Simon Fraser is the developer of Public Knowledge Project’s Open Journal Systems, Open Monograph Press and Open Conference; and SUNY Geneseo has a strong grant-funded program (75% grant-funded). These examples, though, are extreme outliers among the typical model of funding.

Basic timeline for development:

- June 2016-December 2016: in-depth examination of peer institutions and model centers, literature review and gap/needs analysis of CSUN faculty;
- January 2017: service development and investigation of funding and staffing needs;
- Fall 2017: Pilot services;
- Spring 2018 (or after): Launch services;
• Ongoing: consolidate & meet developing demands for services. Devising ongoing strategies for scaling up.

Part VI: References & resources

Further reading:
• ACRL Scholarly Communication Toolkit Publications http://acrl.ala.org/scholcomm/?page_id=48
  o New-Model Scholarly Communication: Road Map for Change (University of Virginia).
• University of California Curation Center (UC3) Merritt https://merritt.cdlib.org/

Select centers on digital scholarship:
• Brown University: Center for Digital Scholarship http://library.brown.edu/cds/ *(looks like a good model)
• Columbia University: Scholarly Communication Program http://scholcomm.columbia.edu/  
  o Columbia University: Center for Digital Research and Scholarship http://cdrs.columbia.edu/cdsmain/?q=index.php
• Emory University: Emory Center for Digital Scholarship http://digitalscholarship.emory.edu/
• Fresno State University: Technology Innovations for Learning & Teaching (TILT) http://www.fresnostate.edu/academics/tilt/
• IUPUI: Center for Digital Scholarship http://www.ulib.iupui.edu/digitalscholarship
• Johns Hopkins University: Scholarly Communications Group (SCG) http://guides.library.jhu.edu/scholcomm
• North Carolina State University: Copyright and Digital Scholarship Center https://www.lib.ncsu.edu/cdslc
• Northwestern University: Center for Scholarly Communication & Digital Curation http://www.library.northwestern.edu/services/faculty-graduate-students/scholarly-communication
• Notre Dame: Center for Digital Scholarship http://library.nd.edu/cds/
• Rutgers University: Scholarly Communication Center http://www.libraries.rutgers.edu/scc/about
• Simon Fraser University: Public Knowledge Project Publishing Services (PKP/PS) 
  https://pkp.sfu.ca/
• Stony Brook Center for Scholarly Communication
  http://library.stonybrook.edu/locations/center-for-scholarly-communication/
• SUNY Geneseo: Technical Services (Scholarship and Publishing Services)
  http://www.geneseo.edu/library/publishing
• University of California: Office of Scholarly Communication http://osc.universityofcalifornia.edu/
• University of Kansas: Center for Faculty/Staff Initiatives & Engagement
  https://lib.ku.edu/people/center-faculty-staff-initiatives-engagement
• University of Oregon: Digital Scholarship Center (DSC)
  https://library.uoregon.edu/digitalscholarship