Robert C. May
Telephone: (510) 987-0711
Email: robert.may@ucop.edu

Chair of the Assembly of the Academic Senate
Faculty Representative to the Regents
University of California
1111 Franklin Street, 12th Floor
Oakland, California 94607-5200

March 27, 2019

JANET NAPOLITANO, PRESIDENT

MICHAEL T. BROWN
PROVOST AND EXECUTIVE VICE PRESIDENT

Re: Concerns Regarding the Use of Research Information Management Systems (RIMS)

Dear Janet and Michael,

At its March 20, 2019, meeting, the Academic Council reviewed the attached report and recommendations from the University Committee on Academic Computing and Communications (UCACC) and the University Committee on Library and Scholarly Communication (UCOLASC) concerning administrators' use of research information management systems (RIMS) to assess faculty.

The UCACC and UCOLASC report echoes concerns expressed by UCAP last year about campuses' use of Academic Analytics, one RIMS that collates quantitative data on faculty research productivity. The report details broad concerns about RIMS that relate to 1) the use of data analytics for faculty advancement, strategic priority-setting, and resource allocation; 2) issues with the quality, reliability, and transparency of the data and algorithms used by RIMS for evaluation; and 3) the continued encroachment of commercial third party systems into core university operations and data without well-defined policies governing their use and appropriate safeguards.

A survey (also attached) conducted by a UC Office of Scholarly Communication Working Group found that at least 16 RIMS are in use at individual campuses. However, the Working Group was limited in its ability to collect meaningful data, and is seeking the systemwide administration's help in facilitating a more exhaustive inventory of RIMS and their uses.

At the March 20 meeting, Council unanimously supported requests from UCACC and UCOLASC to endorse the report and its eight recommendations, and to request from the Provost "a system-wide review of all RIMS currently being employed by academic units and elsewhere across the UC, since at the present time we lack a comprehensive understanding of what products are in use, under what contractual terms, who are responsible for acquiring and implementing, and for what purpose."

We look forward to working with you on this issue. Please do not hesitate to contact me if you have additional questions.

Sincerely,

Robert C. May, Chair
Academic Council

Encl.

Cc: UCACC Chair Martone
        UCOLASC Chair Schneider
        Academic Council
        Senate Directors

UNIVERSITY COMMITTEE ON ACADEMIC COMPUTING
AND COMMUNICATIONS (UCACC)
Maryann Martone, Chair
maryann@ncmir.ucsd.edu

UNIVERSITY COMMITTEE ON LIBRARY
AND SCHOLARLY COMMUNICATION (UCOLASC)
Richard A. Schneider, Chair
rich.schneider@ucsf.edu

ACADEMIC SENATE
University of California
1111 Franklin Street, 12th Floor
Oakland, California 94607

March 12, 2019

ROBERT MAY, CHAIR
ACADEMIC COUNCIL

RE: Concerns Regarding the Use of Research Information Management Systems (RIMS)

Dear Robert,

On behalf of UCACC and UCOLASC, we are submitting a report and recommendations on the use of research information management systems for faculty analytics. This report was drafted at your request, and was prepared in consultation with members of the UC Office of Scholarly Communications RIMS Working Group.

Regards,

Maryann Martone, Chair
University Committee on Academic Computing and Communications

Richard A. Schneider, Chair
University Committee on Library and Scholarly Communication

Cc: Kum-Kum Bhavnani, Academic Council Vice Chair
     Hilary Baxter, Academic Council Executive Director

Encl.

Concerns Regarding the Use of Research Information Management Systems at the University of California

Report prepared by:

Maryann E. Martone, Chair, University Committee on Academic Computing and Communications; Professor Emerita, Department of Neuroscience, UCSD

Richard A. Schneider, Chair, University Committee on Library and Scholarly Communication; Associate Professor, Department of Orthopaedic Surgery, UCSF

Allegra Swift, UCSD, Chair, Lisa Schiff, CDL, Jennifer Chan, UCLA, and Christopher Shaffer, UCSF, UC Office of Scholarly Communication RIMS Working Group[1]

Catherine Mitchell, Operations Director, UC Office of Scholarly Communication; Director of Publishing and Special Collections, California Digital Library

CONTENTS

1. Background
2. The current RIMS/research analytics landscape
3. Issues and concerns regarding the use of analytics and supporting systems
4. The RIMS Census and Survey by the UC Office of Scholarly Communication RIMS Study Working Group)
5. Recommendations
6. Concluding thoughts and an "ask" of Council

BACKGROUND

The University of California Committees on Academic Computing and Communications (UCACC) and on Libraries and Scholarly Communications (UCOLASC) held a joint meeting on Oct 22, 2018 to discuss mutual concerns regarding the use of research information management systems (RIMS) for faculty assessment. RIMS are being marketed by commercial entities to universities as a means for evaluating research productivity, and to provide profile and data management for

[1] Allegra Swift, Scholarly Communications Librarian, UCSD
Lisa Schiff, Associate Director for the Publishing & Special Collections Group, CDL
Jennifer Chan, Scholarly Communication Librarian, UCLA
Christopher Shaffer, University Librarian, Assist. Vice Chancellor for Academic Information Management, UCSF

functions such as academic review. The Committee on Academic Promotion (UCAP) previously addressed a letter[2] to the Academic Senate rejecting the use of a specific company, Academic Analytics, that has been aggressively marketing in this space. However, they are not the only company, nor the largest. An overview of RIMS was provided by Allegra Swift, the Scholarly Communications Librarian for UCSD. Several of the major scholarly publishers, e.g., Elsevier, Springer, Pearson, who collectively control the majority of the research and scholarly output of colleges and universities, are turning their business models towards data analytics, using their considerable content assets to develop RIMS, online learning management systems (LMS) and associated tools for increasing productivity. These companies are also acquiring critical pieces of scholarly infrastructure and analytics companies, including pre-print servers and infrastructure, e.g., SSRN, the social sciences preprint repository, and bepress, both of which have been acquired by Elsevier[3]. These acquisitions are allowing such companies to develop end-to-end solutions for research and scholarly production, dissemination and analysis, while at the same time reducing competition in this space. Bepress, founded over 15 years ago by UC Berkeley Law and Economics professors, was the original system used for UC's open access repository and publishing platform (eScholarship). While eScholarship had already moved away from bepress technologies years ago, a great many other institutions (and their libraries and faculty) were surprised and dismayed to find that their local open access repositories were now owned by Elsevier.

Germane to this discussion, UCACC members were also given an update on the current negotiations with Elsevier regarding journal subscriptions and open access, and unanimously agreed to endorse the *Declaration of Rights and Principles to Transform Scholarly Publishing*[4] put forth by the UCOLASC to guide these negotiations and to promote transparency and open data more broadly.

As a result of this discussion, we believe that concerns about analytics regarding faculty activity (i.e. "faculty analytics") goes beyond issues of how these data are used in decisions about hiring, promotion and resource allocation, to a broader issue of

---

[2] https://senate.universityofcalifornia.edu/_files/committees/ucap/ucap-to-council-re-academic-analytics.pdf

[3] Kelty, C. (2016, May 18). It's the Data, Stupid: What Elsevier's purchase of SSRN also means. http://savageminds.org/2016/05/18/its-the-data-stupid-what-elseviers-purchase-of-ssrn-also-means/; Schonfeld, R. C. (2017, August 2). Elsevier acquires institutional repository provider bepress. https://scholarlykitchen.sspnet.org/2017/08/02/elsevier-acquires-bepress/

Posada, A and Chen, G. (2017, September 20) Preliminary Findings: Rent Seeking by Elsevier http://knowledgegap.org/index.php/sub-projects/rent-seeking-and-financialization-of-the-academic-publishing-industry/preliminary-findings/

[4] https://senate.universityofcalifornia.edu/_files/committees/ucolasc/scholcommprinciples-20180425.pdf

how 3rd party systems that may end up supporting critical core mission of the university should be evaluated and provisioned.

UCACC and UCOLASC are of the strong opinion that universities must exert control of the use of their data, which now includes not only scholarly outputs, but also data accrued across all aspects of the research enterprise from faculty and student activities. We see in the new area of faculty analytics both an opportunity and a threat. Such systems have the capacity to streamline cumbersome processes such as upkeep of academic CV's and preparation of promotion materials. At the same time, if we are not careful in the systems we choose and how they are used, we risk repeating some of the significant and unsustainable problems with our current scholarly communication systems. In particular, UCACC and UCOLASC urge extreme caution in entering deals with 3rd party vendors that would cede control and sale of even more of UC's scholarly and data assets without appropriate protections.

THE CURRENT RIMS/RESEARCH ANALYTICS LANDSCAPE

RIMS collect and store structured data about faculty research and scholarly activities for an institution.[5] Other terms include "profile system" or "networking tool" and variations thereof. In Europe such systems are more commonly termed Current Research Information Systems (CRIS). These systems are designed to perform various functions, such as giving an overall picture of the research and scholarly enterprise of an institution, and providing faculty tools for collaborating, publicizing their work, CV creation and maintenance, complying with policies (such as open access policies), and creating reports for faculty annual review or promotion and tenure. Current tools include:

- Converis (Thomson Reuters): (http://converis.thomsonreuters.com/)

- Pure (Elsevier): (https://www.elsevier.com/solutions/pure)

- Symplectic Elements: (http://symplectic.co.uk/products/elements/)

- Academic Analytics (https://academicanalytics.com/)

- Activity Insight (Digital Measures): (http://www.digitalmeasures.com/activity-insight/)

- Faculty180 (Data180): (http://www.data180.com/faculty180.php)

- Dimensions (Digital Science): (https://www.digital-science.com/products/dimensions/)

---

[5] This material was adapted from materials available from the American Library Association, http://www.ala.org/acrl/publications/keeping_up_with/rims

Most of these are proprietary systems, although some open source networking platforms exist, e.g, Harvard Profiles and VIVO (developed at Cornell) are designed for research networking and collaboration, providing a web-based snapshot of faculty scholarship in public, online profiles and tools for finding expertise within research areas. [6]

Recently, the California Digital Library (CDL), the UC Libraries, and the Office of Scholarly Communication (OSC) formed a working group that surveyed existing RIMS across the UC system. Their goal was to understand what products are in use and for what purposes. The working group faced many challenges in collecting their data because there is no formal mechanism in place for reporting on the use of RIMS, and a lack of consensus on which systems should be included in such a survey. The final report of the OSC working group is included at the end of this report.

ISSUES AND CONCERNS REGARDING THE USE OF ANALYTICS AND SUPPORTING SYSTEMS

UCACC and UCOLASC's concerns regarding RIMS are threefold:

1) We echo the concerns of UCAP and faculty at other institutions regarding the use of data analytics for faculty advancement, setting strategic priorities and resource allocation.

2) Issues with quality and transparency of the data and algorithms employed by these systems for evaluation

3) The ceding of critical university operations and data to closed, 3rd party systems without well defined policies governing their use and appropriate safeguards for the data assets of the university.

1) **Data analytics for faculty advancement, setting strategic priorities and resource allocation**:

> "*As scientometricians, social scientists and research administrators, we have watched with increasing alarm the pervasive misapplication of indicators to the evaluation of scientific performance.* "-Hicks et al., 2015

The tendency to rely on simple numeric indicators such as citation counts, and scores such as the impact factor and h-index to evaluate research performance and impact has had a negative impact on scholarly communications (Hicks et al., 2015). Such scores are often developed initially as an "unbiased" aid for comparing outputs and are meant to be used in conjunction with other factors and human intelligence for making decisions. However, over time, these numbers become imbued with a power

---

[6] See Bryant et al., (2018) Practices and Patterns in Research Information Management: Findings from a Global Survey, Report prepared by Online Computer Library Center for additional examples

they were never meant to have and cannot achieve. The pernicious effects of such numbers on promotion, tenure, resource allocation and academic quality has led to concerted efforts to eliminate or reduce their influence, e.g, the San Francisco Declaration on Research Assessment (DORA)[7] and the Leiden Manifesto (Hicks et al. 2015). Despite these efforts, and "…a long history of scholarship that demonstrates the meaninglessness of most of these numbers in assessing actual impact or long-term value of an individual's contributions", the uses of these metrics continue to proliferate (Borgman, 2018)." As we are trying to move away from uncritical use of such numbers in evaluating scholarly communications, we must not let this practice gain a foothold in other areas of university business. Both UCACC and UCOLASC join UCAP and faculty at other institutions in rejecting commercial RIMS for faculty evaluation[8].

2) **Data/Algorithm quality and transparency**

"No one should accept a black box evaluation machine" (Hicks et al. 2015)

The current state of RIMS and the information sources on which they rely mean that they are often built upon a foundation of incomplete and unreliable data operated on by unvalidated and opaque algorithms. "Calculations of citations, h-index, and other indicators vary widely between common sources such as Web of Science, Google Scholar, and Scopus, due to differences in editorial coverage, algorithms, and methods used by those attempting to mine these databases." (Borgman 2018b). One of the chief complaints against RIMS is that their data are incomplete, as they are usually accumulated by relying on automated extraction from literature, grant or patent databases, or scraping of faculty profiles available on institutional web pages. Coverage for many disciplines (especially humanities and social sciences) or for particular publication types, e.g., grey literature such as lectures, technical reports, pre-prints, blogs or white papers, is either absent or poorly served by available automated data sources. Non-federal grant data is difficult to capture automatically and even federal grant data may be incomplete. Perhaps of most importance, much of the information an institution may want to capture about their faculty does not have a electronic source and requires manual entry. Some examples include honors and awards, students mentored, journal editorships, grant reviewing, creation of electronic resources such as databases, websites and software tools.[9]

---

[7] https://sfdora.org/. DORA has been signed by over 1100 organizations and 13,000 individuals as of Feb, 2019

[8] https://www.educationdive.com/news/ut-austin-faculty-joins-movement-against-employee-analytics-company/515706/

[9] Excerpted and adapted from the American Library Association, http://www.ala.org/acrl/publications/keeping_up_with/rims

Transparency of purpose, data and algorithms is one of the Leiden Principles guiding research evaluation: *Leiden principle* #4: *Keep data collection and analytical processes open, transparent and simple.*

Any system that purports to provide data and tools for faculty research should provide clear information on the type of data collected, how they are structured, the sources of the data and how they are extracted and updated. The same requirements should apply to any analytics calculated. Without transparency of both data and algorithms, there is no way to independently validate them using trusted sources, so that the reliability and significance of these numbers for a given purpose can be understood.

3) **The continued encroachment of commercial 3rd party systems into the core business of the university.**

> *Those individuals who recognize the value and opportunities in these data are not necessarily obligated to seek permission to exploit them. Third parties outside the university may be the first to recognize data opportunities, and approach individuals at any level of the university for partnerships. (Borgman 2018a)*

Academic institutions and scholarly societies long ago ceded the copyright and stewardship of the majority of their intellectual output to commercial publishers. A true competitive marketplace for the services they provide would have led to competitive pricing and advantages for the consumers. But as numerous reports and our own battles with Elsevier indicate, that is not what happened. Scholarly content is not interchangeable, where we can swap the contents of one journal for another. Rather, we ended up with a "bizarre triple pay" system, where the public sector pays for the research, pays the salaries of those who check the quality of the product, and then buys back the product (Buranyi 2017).

In the age of big data, the intellectual assets of the university no longer just include its scholarly output in the forms of articles and books. Rather, the activities of the university, whether in patient care, faculty activities or student learning, generate huge amounts of data that can and are being monetized. These areas are increasingly being targeted by commercial entities, some of whom also hold the rights to our scholarly output. Scientific, scholarly and educational publishers are repositioning themselves as data analytics companies (SPARC, 2019). These companies are combining their content holdings with additional types of data, and acquiring an increasing number of companies and products that are involved in the production, hosting and analysis of these data. The major players-e.g., Elsevier, and Digital Science-have or are considering product offerings in RIMS, as are several smaller companies. These companies are aggressively marketing these complex infrastructures and associated analytical services to universities, not just to libraries, but to provosts, IT services, chancellors, research offices, and others seeking tools to monitor productivity and evaluate core activities of the university. Governance

mechanisms to assure protection of privacy, academic freedom, intellectual property, information security, and compliance with regulations in the uses of such data are either non-existent, out of date or nascent at best (Borgman, 2018).

Many in the scholarly communications field, particularly those that have been pushing for more open access to scholarly outputs, have grown increasingly concerned regarding the new directions of these companies, and their aggressive marketing practices (Posada and Chen 2018). We are concerned that the same mistakes that have brought us to the impasse with Elsevier will be made in other areas of university business. Towards this end, the Scholarly Publishing and Academic Resources Coalition (SPARC)[10], a community organization dedicated to open access to research and educational materials, commissioned a business analyst to compile a landscape analysis of these data analytics companies and their marketing practices to universities[11]. The report provides an in depth assessment of the commercial potential of this business model, outlines the potential benefits to the university but also the significant threat these systems pose if not regulated by strong data governance and contract negotiations by the universities. The report warns:

> *"The move by publishers into the core research and teaching missions of colleges and universities, with tools aimed at evaluating productivity and performance, means that the academic community could lose control over vast areas of its core activities. In addition, the collection of massive amounts of data about faculty and students poses a significant legal and reputational risk for institutions, along with potential privacy and security threats for individuals." (SPARC draft report, pg 6).*

Why do these systems pose a threat? It is not the commercial nature of these products per se, but rather their proprietary nature and the lack of protections for the university's intellectual property and the privacy of its individuals that pose the greatest concern.

Dr. Christine Borgman, the former chair of the UCACC, had a recent experience that illustrates some of the potential problems in the use of these systems. The original UC eScholarship Repository was powered, until 2012, by bepress, a software company originally founded by two academics as the Berkeley Electronic Press. Bepress also encouraged UC faculty to use its "Selected Works" product to maintain public, professional profiles, running a branded pilot with UC in 2008-09. When CDL decided, in 2009, not to contract with bepress for Selected Works because of technical faults in the system, bepress removed the UC logo from all of the profiles of

---

UC-affiliated faculty. In 2017, bepress was acquired by Elsevier for the reported price of $100 M. UC faculty, who invested time and intellectual energy in creating enhanced metadata for their works in "Selected Works", were not notified of this change. When Dr. Borgman requested that the data and metadata she had contributed be returned to her, bepress/Elsevier refused to provide that data.[12]

In other words, when UC faculty work to improve the accuracy, completeness and richness of the data within a system, we improve the data offerings of the commercial provider, which in turn improves the quality of their analytics. The commercial providers can then monetize the results of these faculty efforts to other parties. But if the contract ends, the system is sold to a 3rd party, or UC wants to change providers, we may lose rights to the data we have provided.

Even if UC does put protections in place to ensure faculty can export their data, the nature of data infrastructures can still lead to significant difficulties when trying to change providers. Proprietary companies tend to build customized solutions that ensure that data contained therein are not portable, so that algorithms that work with one platform do not work with another, even with similar types of information. If the data models, file formats and programmatic interfaces to these systems are proprietary, the data extracted may not be in a form where they can be easily ported to another platform. In this way, commercial systems seek to lock in customers by making it costly to change providers.

The FAIR Data Principles have garnered a lot of attention recently across academic disciplines and government agencies (Wilkinson et al. 2016). UCACC endorsed them in 2016. The FAIR principles outline a set of recommendations to ensure that any data produce is findable, accessible, interoperable and reusable, by both humans and machines. FAIR is usually discussed under the auspices of research data produced as part of research activities, but it also applies to data generated and maintained by the university around faculty and student activities. FAIR does not mean open, but it does mean that the data contained within systems must be structured in a way that conforms to open community standards and is accessible by computers. If data are FAIR, the cost of changing systems is lower, meaning that universities can more easily avoid third party lock in. If true competition exists in the marketplace, the negotiating position of the university for price and services is strengthened.

---

[12] Slide presentation given by Dr. Borgman to UCOLASC on February 23, 2018

THE RIMS CENSUS BY THE UC OFFICE OF SCHOLARLY COMMUNICATION RIMS STUDY WORKING GROUP

Institutions like universities that are highly decentralized are the most at risk from the current marketing practices because different segments of the university often make independent business decisions without consulting other concerned parties.

The UC Office of Scholarly Communication (OSC) has conducted a survey on the use of Research Information Management Systems (RIMS), Current Research Information Systems (CRIS), and related systems across UC campuses that are used to aggregate, organize, and present information on research activities. The goal of this survey, and follow-up interviews, was to create an inventory of systems currently in use in order to provide a baseline understanding of individual campus practices and reporting needs.

This survey included questions to identify which systems are currently in use at each campus, what data those systems contain, the purposes for which the data are collected, and whether access to the data is restricted. The information collected, while not exhaustive, helps us understand—as a system and at specific campuses— where we might be duplicating efforts and where there might be opportunities for efficiency or synergy.

A snapshot of the ecosystem at each campus contributes additional insights into:

- The systems in use and those being marketed to campuses.
- What the stakeholders need these systems to do.
- If these systems are helping faculty/researcher/campus.
- How they are being used to assist Promotion & Tenure review.

With this ecosystem in mind, UC is positioned to compare these different systems in terms of key features such as transparency, data models, open standards used, programmatic access, privacy and reuse policies, so that appropriate governance and contractual requirements can be instituted,

The survey was sent to the following campus stakeholders/offices: Academic Personnel Office, Grants/Contracts Compliance/Administration Office, Office of the Vice Chancellor for Research, Department of Educational Technology, Department of Information Services and Technology, the Library, or campus entities that would purchase or use a RIMS.

The final report is attached.

RECOMMENDATIONS:

We note actions and strategies that UC should adopt to mitigate the potential risks posed by the new breed of company and to leverage the potential benefits of improved efficiencies and services. These recommendations are based on the SPARC report, a checklist for scholarly infrastructures developed by Allegra Swift and David Minor, UCSD Library (Swift et al. 2018), Principles for Open Scholarly Infrastructures (Bilder, Lin, and Neylon 2015) and other efforts to put scholarly communications and both research data and research information data on a more open and sustainable path. These efforts encourage the use of open infrastructures and standards to enable competition rather than squelch it.

1. The university should endorse the DORA and Leiden Principles and use them as a basis for developing policies regarding the use of analytics for faculty assessment and the choice of any systems that may support it.

2. The data policies of the university need to be revisited and updated to deal with all types of digital data-not just data generated by faculty but data generated about faculty, addressing their ownership, collection and reuse and public-private partnerships.

3. Terms and conditions of contracts with commercial vendors should not be covered by non-disclosure agreements (i.e., open procurement)

4. The university should endorse the FAIR principles, and ensure that any systems used for faculty analytics adhere to them

5. Any results from using 3rd party systems should be portable (i.e., institutions maintain the right to historic output series in order to facilitate switching to other vendors).

6. Data may not be resold to third parties, at least without the consent of the university and subject to university governance policies, e.g., data sets that are deemed particularly sensitive should not be turned over to government agencies without first resorting to the appropriate court

7. At minimum, the appropriate committees in the Academic Senate must be consulted on any decisions made by ***any part*** of the university on systems designed to collect or evaluate data on faculty or students. These committees should include data scientists, librarians and other experts who have the expertise to review not only the services provided, but the underlying infrastructure.

8. The University should consider convening a cross-stakeholder review panel comprising administration, faculty, library, IT, both to develop a set of review criteria for such systems and proposed uses consistent with the values of the University of California, and which can provide rapid feedback on any proposed systems relative to these criteria.

CONCLUDING THOUGHTS AND AN "ASK" OF COUNCIL

The SPARC report concludes:

> *"We believe there is still time for the academic community to act, and now is the time to do it. By taking stock of the situation, asking the right questions, and choosing the right course of action, the academic community can prevent itself from winding up in a position where it is obliged to follow a path out of its control and harmful to its future."* pg. 41

We therefore propose that this report and a resolution on behalf of the Academic Senate that endorses the above recommendations be sent to the President .

We also propose that Council ask the Provost to commission a system-wide review of all RIMS currently being employed by Academic units and elsewhere across the UC, since at the present time we lack a comprehensive understanding of what products are in use, under what contractual terms, who are responsible for acquiring and implementing, and for what purpose.

REFERENCES CITED:

Bilder, Geoffrey, Jennifer Lin, and Cameron Neylon. 2015. "Principles for Open Scholarly Infrastructures." *Science in the Open.* https://cameronneylon.net/blog/principles-for-open-scholarly-infrastructures/.

Borgman, Christine L. 2018a. "Open Data, Grey Data, and Stewardship: Universities at the Privacy Frontier." *Berkeley Technology Law Journal / Boalt Hall School of Law, University of California, Berkeley* 33 (2): 365.

———. 2018b. "Text Data Mining from the Author's Perspective: Whose Text, Whose Mining, and to Whose Benefit?" *arXiv [cs.DL].* arXiv. http://arxiv.org/abs/1803.04552.

Buranyi, Stephen. 2017. "Is the Staggeringly Profitable Business of Scientific Publishing Bad for Science?" *The Guardian*, June 27, 2017. http://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science.

Hicks, Diana, Paul Wouters, Ludo Waltman, Sarah de Rijcke, and Ismael Rafols. 2015. "Bibliometrics: The Leiden Manifesto for Research Metrics." *Nature* 520 (7548): 429–31.

Posada, Alejandro, and George Chen. 2018. "Inequality in Knowledge Production: The Integration of Academic Infrastructure by Big Publishers." In *22nd International Conference on Electronic Publishing.* OpenEdition Press. https://doi.org/10.4000/proceedings.elpub.2018.30.

Swift, Allegra, David Minor, Elena Feinstein, Emily Frank, Vanessa Gabler, Robyn Hall, Claudia C. Holland, Allison Langham-Putrow, and Charlotte Roh. 2018. "Scholarly Communications Infrastructure Checklist." https://doi.org/10.6084/m9.figshare.7406849.v1.

Wilkinson, Mark D., Michel Dumontier, Ijsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3 (March): 160018.

Office of Scholarly Communication Research Information Management System Working Group: Final Report

March 1, 2019

Allegra Swift (Chair)
Scholarly Communications Librarian, UC San Diego

Chris Shaffer
University Librarian and Assistant Vice Chancellor for
Academic Information Management, UCSF

Jennifer Chan
Scholarly Communication Librarian, UCLA

Lisa Schiff
Associate Director
Publishing & Special Collections, CDL

Table of Contents

Executive Summary

In 2017, the UC Office of Scholarly Communication established a Research Information Management Systems (RIMS) Working Group and charged it with developing an inventory of RIMS in use across the UC campuses. The Working Group distributed a survey and conducted interviews to accomplish this task. Significant difficulties were encountered in collecting data (as described in the "Data Collection Challenges" of the full report), thus the final data set is quite limited. Despite this small response rate, the Working Group was able to discover that a variety of RIMS (sixteen were uncovered by the Working group's efforts, which is highly unlikely to cover the extent of RIMS and related systems) are in place across the system, each of which is typically used by just a few campuses for discrete purposes in specific units on those campuses, disconnected from other entities. The exceptions to this are the Symplectic Elements RIMS, implemented to support the UC Academic Senate Open Access Policy and used by all campuses, and the UCSF Profiles system, in place at the five medical schools. Most respondents did not answer the questions around data ownership, data privacy, interoperability between related systems, and in one of the interviews we learned that these issues were not considered. Apart from the inventory, the Working Group identified several findings, two of the most significant being that: 1) the library is not well situated for conducting such an exploratory effort outside of higher level campus conversations about the need for generating such information; and 2) opportunities to leverage investments of funds, time and effort in RIMS implementations have yet to be surfaced and realized.

Overview

Research Information Management Systems (RIMS), or Current Research Information Systems (CRIS) as they are referred to in Europe, aggregate, organize, and present information on research activities. Practices vary across institutions and there is often a costly duplication of efforts as departments or offices contract with commercial products or develop systems of their own that may have shelf life for the duration one-time funding.

As part of the University of California Office of Scholarly Communications (OSC) 2017-2018 Work Plan, a working group was charged with developing an inventory of

RIMS and their related systems across UC campuses in order to help us understand—as a system and at specific campuses—where we might be duplicating efforts and where there might be opportunities for efficiency or synergy.

Charge

Our working group was charged with creating an inventory of systems currently in use in order to provide a baseline understanding of individual campus practices and reporting needs. This foundation information can additionally be used by the campuses to determine if there are opportunities for interoperability between systems or consolidation for efficiency and maximizing data reuse for the benefit of satisfying multiple information needs such faculty with compliance/reporting activities and administrators with demonstrating the impact of research at UC.

Goals

The goals of the inventory were to identify:
- The units or offices using RIMS systems on the 10 campuses (Note that the labs and other affiliated research centers were not included in this inventory.)
- The specific systems in place,
- The purposes associated with those systems.

Working Group Members

The RIMS Working Group was comprised of individuals from within and outside of the Office of Scholarly Communication:

- Allegra Swift, Scholarly Communications Librarian (Chair) at UC San Diego;
- Chris Shaffer, University Librarian and Assistant Vice Chancellor for Academic Information Management at UCSF;
- Jennifer Chan, Scholarly Communication Librarian at UCLA;
- Lisa Schiff, Associate Director for Publishing & Special Collections at the California Digital Library.

Methodology

Survey

The Working Group developed a [survey](#) to gather the information presented in this report. It included questions to identify which systems are currently in use at each campus, what data those systems contain, the purposes for which the data are collected, and whether access to the data is restricted.The survey was distributed to the UC Scholarly Communications Common Knowledge Group (SC-CKG) to send to stakeholders on their individual campuses, with the following offices suggested as examples for contact:

- Academic Personnel Office
- Grants/Contracts Compliance/Administration Office
- Office of the Vice Chancellor for Research
- Department of Educational Technology
- Department of Information Services and Technology
- The Library,
- or other campus entities that would purchase or use a RIMS.

Responses were gathered between September 13, 2018 and October 12, 2018.

Follow-up Interviews

Despite the SC-CKG's distribution efforts, only eleven individuals responded to the survey, a number much lower than hoped for and insufficient from which to draw any conclusions. In addition, while analyzing the data the Working Group discovered that as constructed, the survey tool does not allow us to firmly connect respondent identity to submitted answers, making it impossible to be certain which responses are associated with a specific campus and unit. The group decided to address these limitations by conducting follow-up interviews with individuals in various departments either likely or known to have a RIMS.

Our process included an initial contact with the campus individual originally requested to send out the survey in order to: 1) confirm that the survey had been sent out and to whom; 2) solicit suggestions for additional offices to get in touch with; and 3) invite the campus contact to participate in phone calls with these additional individuals. Phone interview questions were the same questions included in the survey.

Data Collection Challenges

In addition to the low number of survey responses, the Working Group identified a number of other difficulties in collecting meaningful data, the most significant of which are enumerated below:

- **Insufficient institutional knowledge:** Nothing even approaching a knowledge base on UC institutional RIMS activities exists from which to gather contacts. While the absence of such a knowledge base was in itself one of the motivations for the formation of the Working Group, operating in this vacuum was essentially a bootstrapping exercise.
- **Limited cross-campus connections**: In general, the library does not have direct relationships with all units and/or individuals who either are or might be running RIMS. Thus, identifying the right staff members across a campus to speak with was a challenge, as was actually setting up conversations even when those people were identified.
- **Ineffective data gathering approach:** While distributing a survey seemed the most efficient way to gather a large amount of data from a broad spectrum of people, because of the limited connections to other campus units and the complexity of the topic, the survey format did not turn out to be an effective data collection method. Additionally, complete survey results reveal that the survey questions were often not well understood, partly due to the idiosyncratic nature of individual RIMS implementations. By contrast, the follow-on interviews resulted in deeper responses and would be the recommended approach for any future efforts, however this strategy would require a significant investment in identifying and being introduced to the appropriate campus units.
- **No/low perceived value:** The Working Group suspects that individuals may not have perceived sufficient value for their own work in answering the survey. Responses generally demonstrated that the scope of current RIMS activity is narrowly tailored to the needs of each individual group running a RIMS and shared systems might not necessarily help such efforts.
- **No incentives:** No incentives were provided to individuals who completed the survey, which may have contributed to the low response rate.

Findings

Despite the data collection challenges, the Working Group was able to gather a sufficient amount of information for a simple, yet still useful, analysis. Presented below

are our high level findings, followed by more detailed information on a per campus basis.

Themes

- **Heterogeneity:** Apart from the systemwide use of Symplectic Elements in support of the UC Academic Open Access Policy, the other 15 RIMS mentioned in the data are typically in place at only one or two campuses.
- **Narrow use**: Where they are in place, most RIMS are used by a single unit for a specific purpose, without plans or strong interest in connecting with other RIMS, data sources, or initiatives.
- **Isolated library "location":** Libraries are not as tied to processes and decisions related to RIMS as we might wish; they aren't well connected to the units that have invested in these systems and sometimes aren't even aware of them.
- **Constrained library "role":** Questions about RIMS are not expected to come from the library; units that are running RIMS for the scoped purposes of their work don't necessarily see the connection to and value of the services and skills the library has to offer.
- **Overburdened staff:** Individuals responsible for maintaining and using RIMS are often overburdened and can't prioritize the time to answer surveys or interviews not obviously beneficial to their work.
- **Localized views:** RIMS administrators often have a narrow view of what functions a RIMS can serve, based on their local use case. They don't think of them as multi-purpose tools, but rather as "the system that lets us do X."

Campus Responses

This table shows at-a-glance the RIMS at various campuses that were found via the survey and interviews (others undoubtedly exist). Detailed responses follow below it.

	CDL	B	D	I	LA	M	R	SD	SF	SB	SC
Symplectic Elements (UC PMS)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Harvard / UC Health Profiles			✓	✓	✓			✓	✓		
Interfolio				✓	✓			✓			
SciVal (without PURe)				✓							
PURE (formerly Elsevier SciVal Experts)			✓					✓			
VIVO								✓			
Academic Analytics											✓
Cayuse (Formerly E-Visions)						✓					
Dimensions				✓					✓		
UCLA Opus					✓						
Digital Measures						✓					
UCSD SIO Research Profiles								✓			

UCOP Research Enterprise Management System (REMS		▨						✓		▨	
Coeus		▨						✓		▨	
UC R e-File		▨					✓			▨	
Kuali Research		▨						✓		▨	
UCSF Advance		▨							✓	▨	

Legend:
✓ Confirmed active
. Formerly active
Greyed Survey Received No Response

- **California Digital Library (CDL)**
  - CDL provides system wide access to Symplectic Elements, initially implemented in support of the UC Academic Senate Open Access policy. That implementation will be expanded in 2019 to include a phased roll-out of the Presidential OA policy will begin later in 2019.
  - The UC instance of Elements is primarily populated through automated queries against publication data sources, such as PubMed Central and Web of Science. Users can also add records manually.
  - Elements is accessible only by individuals with login credentials or systems connecting via the API and an API key, however the data captured therein can be shared with other systems via an API that also requires credentials.
- **UC Berkeley**
  - The survey was sent to the UCB VCR's office by the University Librarian but there was no response to the survey or subsequent requests for an interview by phone.
- **UC Davis**
  - Schools of Medicine, Nursing, and several other health departments use Elsevier Pure.

- Health sciences researchers are also included in the UCSF-hosted UC Profiles.
- Davis is investigating a campus-wide RIMS. The Library is taking a lead role in investigating options and doing a landscape analysis. Academic Senate will be establishing a governance model. They expect to go to RFP in mid-2019.
  - **UC Irvine**
    - The Office of Research is interested in understanding the research and publication practices of its faculty and to support this work had a license to Dimensions for one year,but have now switched to SciVal instead.
    - Health sciences researchers are also included in the UCSF-hosted UC Profiles.
  - **UC Los Angeles**
    - While anecdotally aware of various RIMS activities taking place at UCLA, we were only able to obtain one response to the OSC survey courtesy of the UCLA Academic Personnel Office. The response centered on the implementation of [Opus](), which is a bridge system between various Faculty Information System data sources and [Interfolio](). The various internal and external data sources from which Opus draw include arXiv, ORCiD, PubMed, SSRN, Web of Science, and others.
    - Health sciences researchers are also included in the UCSF-hosted UC Profiles.
  - **UC Merced**
    - The Academic Personnel Office uses Digital Measures for faculty review purposes, a system that requires faculty to upload a reference and either a publication or a link. A BioBibb can then be generated and submitted to Interfolio, which is Merced's review system. Academic Personnel has also used Digital Measures to understand which journals Merced faculty publish in and have occasionally evaluated keywords to determine research areas. This is a closed system.
    - The Office of Research and Economic Development doesn't have a formal RIMS, but does have a system called Cayuse (formerly E-Visions) that is used to manage grant proposals, submissions and human subjects protocols. Faculty are responsible for uploading relevant information, which can include prior publication information in support of a grant request or human subjects approval, however this is not required. This is a closed system and in order to ensure confidentiality there are no integrations into or out of it.
- **UC Riverside**

- ○ Staff from the Academic Personnel Office and Information Technology Solutions responded to the survey.
- ○ The UCR eFile system was described as maintaining "academic record including research, service, awards, teaching, etc. It is also used to create review files for merit and promotion academic actions." Some metadata is entered manually and ingested from other campus systems (i.e. teaching load and evaluations).The UCR eFile System can be accessed by campus and affiliate accounts. Faculty have control over the data that is entered in the system.
- ○ When asked what features were desired in a RIMS, several features were checked as extremely important.
  - ■ Automatically collect citations/metadata from a variety of sources (not just PubMed or STEM related databases) such as ORCiD, BASE, Institutional repositories, preprint servers, etc.:
  - ■ Ease of user entry
  - ■ Faculty/researcher control of their data
  - ■ On-the-fly CV and biobib creation
  - ■ Institution has perpetual data access and transparent exit/export strategies
  - ■ The respondents also sought "ease of feature updates" and "full integration with other campus systems."
- ● **UC San Diego**
  - ○ The three respondents were from the Library's Research Data Curation Program (RDCP), Vice Chancellor's Office of Academic Affairs and Electronic Research Administration Program (ERAP).
  - ○ Academic Affairs is responsible for Interfolio Faculty180 and gathering faculty research information for the purposes of internal reporting solely for use by the academics. However, "a formal Senate task force is being created to consider opening access to senior administrators to run reports on the data collected in the system."
  - ○ The library reported campus use of Elsevier SciVal Experts (now PURE), VIVO,
  - ○ UCSD Health uses UCSD [Harvard] Profiles
  - ○ Scripps Institute of Oceanography uses UCSD SIO Research Profiles
  - ○ VCR ERAP uses the UCOP Research Enterprise Management System (REMS) for managing lifecycle activities related to research projects, Coeus, and Kuali Research .. The research compliance systems are used to insure research activities/projects remain in compliance with various federal, state and local policies.

- ○ Material Transfer Agreement Request System (EMTA) manages requests to transfer research materials in/out of the university. eDisclosure is used to manage intellectual property invention disclosures.
- ○ Several features were checked as extremely important but not currently included in RIMS;
  - ○ Automatically collect citations/metadata from a variety of sources (not just PubMed or STEM related databases) such as ORCiD, BASE, Institutional repositories, preprint servers, etc.
  - ○ Ease of user entry
  - ○ On-the-fly CV and biobib creation
  - ○ Institution has perpetual data access and transparent exit/export strategies

- ■ **UC San Francisco**
  - ● The CTSI, an NIH-funded grant office, runs UCSF Profiles, based on Harvard Profiles. Harvard Profiles is very health oriented, and most data is harvested from PubMed, ClinicalTrials.gov, and NIH Reporter (grants). Researchers can claim publications harvested from CDL's Symplectic Elements. They can also add/edit/correct other information in their profiles. Faculty can export to Advance, the UCSF promotion and tenure system.
  - ● They also run a UC Health Profiles at https://profiles.ucbraid.org/ that includes health sciences researchers from Davis, Irvine, Los Angeles, San Diego, and San Francisco.
  - ● The Research Development Office (RDO), which supports major grants and team science, recently subscribed to Dimensions. They plan to use it to support team science activities. It hasn't been activated yet, so nobody is quite sure exactly how it will be used or what will be available to end users.
- ■ **UC Santa Barbara**
  - ● The survey was distributed to multiple campus units, however, no responses were received. Subsequent follow up attempts similarly received no response.
- ■ **UC Santa Cruz**
  - ● The Office of Institutional Research, Assessment, and Policy Studies has licensed Academic Analytics in order to understand at an aggregate level how UCSC faculty activity compares to that of peer campuses, looking at publications, grants, and awards. The purpose is to provide strategic context, not for tenure or personnel decisions. For instance, if a

department receives all of its grants from NSF and peers are getting them from NIH, IRAP will share that information with the department so that its faculty can broaden their grant requests. Departments also request IRAP to provide them with this type of peer-context. .

Conclusions and Possible Next Steps

To develop a landscape view of the existence and use of RIMS across UC, the OSC RIMS Working Group developed and distributed, in partnership with Scholarly Communication CKG members, a relatively brief survey intended to efficiently surface information about RIMS on the campuses. Survey questions were constructed in a nuanced way to learn not only which systems were being utilized, but to what end. This nuance, however, may have appeared as complexity to respondents, as the survey produced minimal results. Fortunately follow-on interviews conducted to augment that limited data provided sufficient information from which to draw some very simple conclusions:

- Sixteen different RIMS were discovered to be in place across UC, most used by one or two campuses. (Note that other RIMS likely exist, but were not found via this process.)
- Symplectic Elements is used by all ten campuses, making it the most widely implemented. However its use is narrowly focused on the open access policies.
- Five campuses use UCSF/Harvard Profiles (limited to health sciences schools and units), making it the second most frequently used RIMS.

The data also reveal a narrow understanding, or perhaps minimal interest, in the broader use of individual products, which in turn may indicate a lack of understanding of the possibilities to be gained by integrating and/or sharing RIMS across a given campus. Such opportunities hold the promise of accruing benefits to multiple stakeholders, providing richer, more complete reporting while reducing effort and increasing control for individual researchers--but only if those responsible for running such systems also recognize that promise.

The findings presented earlier in this report do not claim to be an exhaustive inventory of all RIMS activity at each campus in the UC system, but they do represent a starting point should there be a desire to address this issue in the future. Given the somewhat isolated location of the library on campuses and the fast-developing nature of the RIMS space, establishing consensus on the utility of such an inventory among high-level unit heads (e.g. between ULs and Vice Chancellors for Research) in advance of such an

effort would be advised. To that end, determining whether there are prioritized specific goals and follow on actions would be helpful in communicating the value of contributing to such an effort.

Appendix A: Survey Instrument

The survey instrument has been exported to a [PDF](#) for viewing.