Pushing on the Paywalls: Extending Licensed Resource Access to External Partners to Enhance Collaborative Research

Juleah Swanson
Presenter

Steven W. Brown
Presenter

Definitions of authorized users in license agreements not only dictate who is allowed to access licensed resources, but also define who can be considered part of an institution's user community. Researchers engage in collaborative research, sometimes holding multiple affiliations that, at times, may extend beyond a definition of authorized user. This paper examines how libraries play a role in supporting or inhibiting collaborative research by exploring a strategic partnership between the University of Colorado Boulder (CU Boulder) and the National Center for Atmospheric Research (NCAR) to further collaboration related to atmospheric research and climate studies. While the goals of the partnership sought to enhance research and collaboration through access to licensed resources, the authors found that the paywalled model of access through license agreements, authentication, and access and discovery methods has complicated the effectiveness of creating a collaborative research environment.

KEYWORDS affiliated users, authorized users, licensing, access and discovery, collaboration

BODY OF PAPER

Introduction

In our era of paywalls, access is granted based on who is and is not considered an authorized user. Yet, researchers may hold multiple affiliations, identities, and roles across institutions, universities, industries, governments, and even countries. When collaborative relationships among researchers extend beyond the definition of authorized users, the paywalled model of access for e-resources becomes a barrier to seamless and efficient collaboration and high quality research. Libraries can play a role in either streamlining or complicating access to e-resources for collaborative research.

For researchers holding multiple affiliations, choosing which affiliation to invoke may influence what they can and cannot access. A question received by the authors from a researcher who holds multiple affiliations, including credentials with the University of Colorado Boulder (CU Boulder), illustrates this complication.

> *What credentials should I use to register/log in to [Web of Science]? Do I use this email format: [xxxx]@colorado.edu and my standard Identikey password?*

This question demonstrates the confusing environment that libraries pose to researchers - primarily because of paywalled access to resources. The user is essentially asking, *Who am I supposed to be when I access Web of Science? My University of Colorado Boulder self or my other affiliated self?* It also raises questions on who researchers see themselves as, or which affiliation is most prominent in their minds. Should it matter what identity a researcher invokes? What role do libraries play in influencing this identity?

At CU Boulder, users who hold multiple affiliations might be designated as Persons of Interest (POI).[1] POIs are formally enrolled into the human resources system as a person with a

non-payment relationship with the university. POIs may include visiting faculty, visiting

researchers, those with dual appointments, and other formal affiliates. In license agreements,

publisher definitions impose a certain understanding of what it means to be a collaborator that

does not necessarily translate to the reality of collaborative research. Publishers define

authorized users to suit their individual philosophies of how broadly content access should

extend. For example, in the American Chemical Society license agreement, authorized users are

defined as those serving in the capacity of an employee, faculty, or teaching staff, but makes no

mention of visiting scholars, researchers, affiliates, or other formal members of an institution's

community.

> *"For purposes of this Agreement, "Authorized Users" means, those serving in the*
> *capacity of employees, faculty and other teaching staff, and persons officially registered*
> *as full or part-time students located at an Authorized Site."[2]*

Libraries are in the difficult position of having to translate between the publisher's definitions of

authorized users and the reality of collaborative research and the community of individuals who

comprise institutional users.

The complexity of collaborative research that takes place across campuses and at

institutions worldwide is further complicated by the paywalled system of access to resources

managed by libraries. This paper examines a specific example of a collaborative research

partnership between the National Center for Atmospheric Research (NCAR) and CU Boulder to

illustrate the tension between fostering collaboration among researchers, while grappling with

the many barriers a paywalled landscape imposes on access to and discovery of library resources.

Background

NCAR, a Federally-funded Research & Development Center (FFRDC) of the National Science Foundation (NSF), and CU Boulder have a long history of collaboration due to shared research interests and co-location of facilities in Boulder, Colorado. From 2005 to 2015, CU Boulder ranked as the top institution with co-authorship with NCAR, above prominent collaborators such as the National Oceanic and Atmospheric Administration and the University of California Berkeley. Similarly, from 2007 to 2018, twenty-five percent of all papers produced by NCAR included CU Boulder co-authors.[3, 4]

In 2016, the Dean of University Libraries at CU Boulder and the Director of the NCAR Library discussed ways both entities could further their existing collaboration. These discussions included leveraging resources that related to shared research areas including data services and data storage provided by the NCAR Computational and Informational Systems Lab (CISL) and access to relevant licensed content from CU Boulder Libraries. From these discussions, a memorandum of understanding (MOU) was signed that outlined mutual efforts to collaborate on data services and storage, and access to licensed content. The MOU helped to formalize existing collaborations, strengthen dual affiliation relationships, and acknowledge shared resources both entities sought to leverage. While a significant portion of the collaborative relationship between NCAR and CU Boulder is related to data services and data storage, this paper only focuses on the portion related to shared access to licensed resources, as the authors were not involved in the data services and data storage portion of the MOU.

For the portion of the MOU that related to extending access of licensed resources to dual affiliates from NCAR, there were three primary goals that shaped the solutions CU Boulder and NCAR sought:

1. Enable dual affiliated researchers' access to CU Boulder licensed resources at NCAR facilities;

2. Integrate discovery and access of CU Boulder licensed resources into existing system and tools used by the NCAR Library;

3. Maintain secure, authenticated access to resources.

Licensing

To extend access, the first step in the project involved investigating existing CU Boulder license agreements to understand coverage of affiliates as authorized users. The basis of extending access centered on the dual affiliate, or POI status, that researchers at NCAR held with CU Boulder. While most CU Boulder license agreements allow for on-site, or walk in user access, not all licenses extend remote access to affiliated persons. Because the needs of the project were to enable access from non-campus locations, this required a license-by-license review of the definitions of authorized user for inclusion of affiliated persons and remote access. In cases where it was not evident or when affiliated users were explicitly not included in the definition of authorized users, discussions with publisher or vendors were necessary.

We engaged in conversation with each publisher in question in order to seek a resolution to the issue. In some cases, additional payment was required in order to modify the license agreement's authorized user language, increase the number of users covered, or add an additional

site to the license. In other cases, publishers were agreeable to the change in authorized user language without additional payment, as the total number of additional users and the level of usage was insignificant in comparison to the number of users and usage that was already occurring from CU Boulder students, faculty, and researchers. Yet, the problem with allowing license agreements to rigidly define who authorized users are is that the institutions' own definitions of users are not recognized by publishers, as is the case with CU Boulder's POIs. By embarking on this investigation and determining whether definitions of authorized user include affiliated researchers such as those from NCAR, we found that our license agreements, at times, are not reflective of the university community, particularly for those engaging in collaborative research and holding multiple affiliations.

Secure Access, Authentication, and Discovery

To enable on-site access at NCAR facilities through secure authentication, an instance of EZproxy was set up. Because access was facilitated through EZproxy, local IP access at NCAR facilities was terminated. Thus for NCAR researchers, their pathway for accessing articles and journals changed. NCAR researchers were expected to login with their username and passwords via a proxied URL, such as through the NCAR Library link resolver, instead of being able to navigate to a resource via a bookmark or directly from the publisher website. However, in some cases both methods of access via local IP and EZproxy persisted for resources, either due to local IPs remaining active, due to perpetual access rights to previously purchased content, or in some cases having access to Open Access articles that were not clearly differentiated from paywalled content. Anecdotally, NCAR researchers found this confusing and were unclear about why former methods of access worked some of the time.

The MOU between CU Boulder and NCAR led to an astounding increase in access to e-resources at NCAR from approximately 150 e-journals in 2017 to over 6,000 by 2018. This increase required the NCAR Library to integrate a significantly larger number of titles into their existing discovery systems and tools. To integrate the resources into systems and tools maintained by the NCAR Library, titles were added into the NCAR knowledgebase and link resolver. One challenge in integrating access resulted from the differing systems in use by CU Boulder and NCAR. Both libraries used different link resolver and knowledgebase systems with CU Boulder using 360 Knowledgebase, while NCAR uses SFX. As a result, title matching was an imperfect process, highlighting the problem of lack of integration and de-centralized knowledgebases from vendors.

Assessment and Usage

Beginning in 2019, the NCAR-CU Boulder library team began to explore ways in which to measure impact, value, and success of the project since implementation in January 2018. We expected to see a correlation between the increased availability of resources and an increase in usage, even with only a little over one year's worth of data.

To evaluate usage, Counting Online Usage of Networked Electronic Resources (COUNTER) and link resolver usage reports were downloaded for each publisher. NCAR-specific COUNTER Journal Report 1 (JR1) reports, measuring the number of successful full-text article requests by month and journal, were downloaded for seven publishers, while combined NCAR and CU Boulder COUNTER JR1 reports were downloaded for eighteen publishers when separate reports were unavailable. Combined NCAR and CU Boulder COUNTER Book Report 2

(BR2) reports, measuring the number of successful book section requests by month and title, were downloaded for publishers with e-books included in the subscription. In addition to the COUNTER reports, NCAR's SFX and CU Boulder's 360 Link click-through reports were downloaded for several publishers to analyze usage through link resolvers.

To assess whether or not an increase in titles correlated with an increase in usage, we examined usage from the American Institute of Physics (AIP). Researchers at NCAR gained access to an additional twenty-four titles through the project. We found that usage for titles that NCAR had access to in 2017 remained steady in 2018, but the newly accessible titles were virtually unused through 2018. Only two titles saw a single use each, while the other twenty-two titles went unused. Though we were only able to evaluate one year's worth of data, the near-zero usage of the twenty-four new titles was surprising and raises questions about the project generally, beyond AIP journals, regarding implementation, outreach, and impact. Why is there near-zero usage of the twenty-four new titles? Is additional outreach and training needed to highlight the additional titles? Are discovery systems optimized to allow for discovery of newly added content? Or, are the added titles simply not relevant to NCAR researchers?

For some publishers, NCAR users retained access to the same titles as before, with no additional content added. We expected usage statistics to remain stable from year to year despite the new method of access requiring authentication via EZproxy. The American Chemical Society (ACS) is an example where no changes were made to content availability, and the only change was to mode of access. We found that PDF usage for ACS titles was relatively stable from 2016 to 2018 (see Figure 1), but HTML usage saw a 300 percent increase, mostly driven by a single title, but did see increases across most titles (see Figure 2). This suggests that researchers were

able to find and access the content through authentication via EZproxy for resources they were

already familiar with.

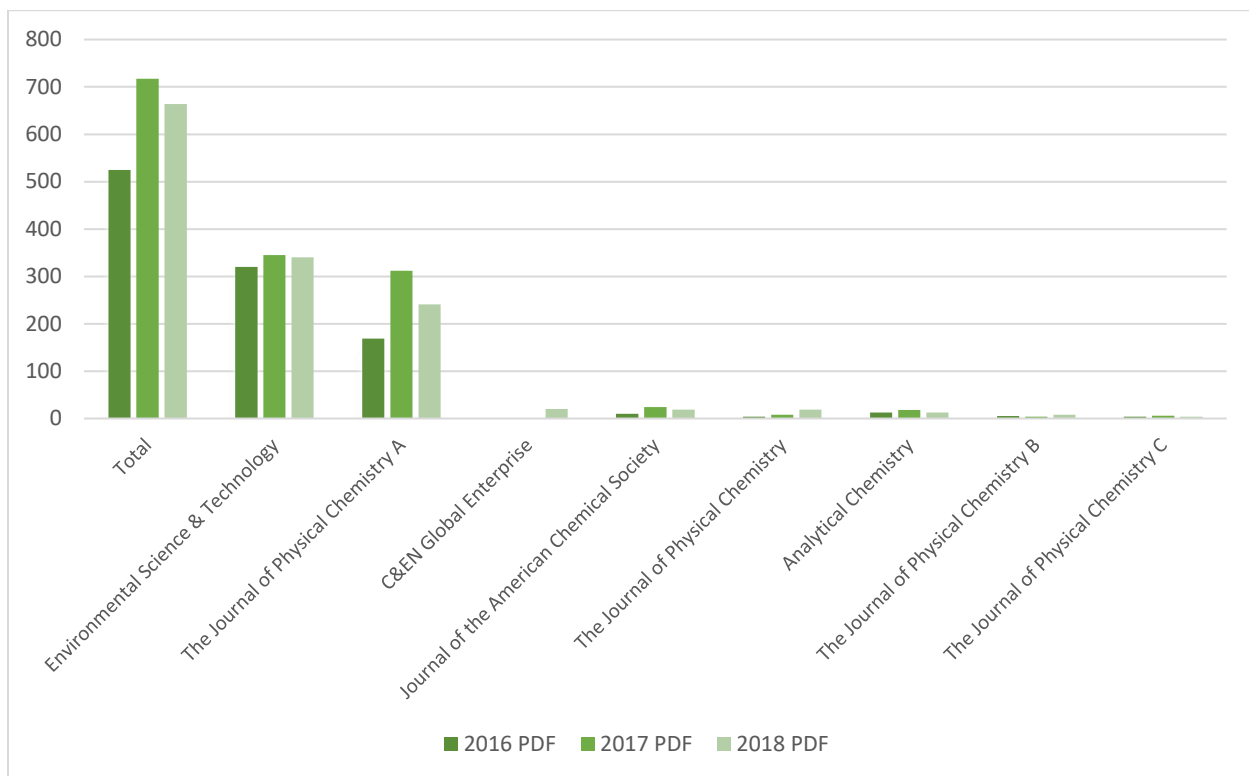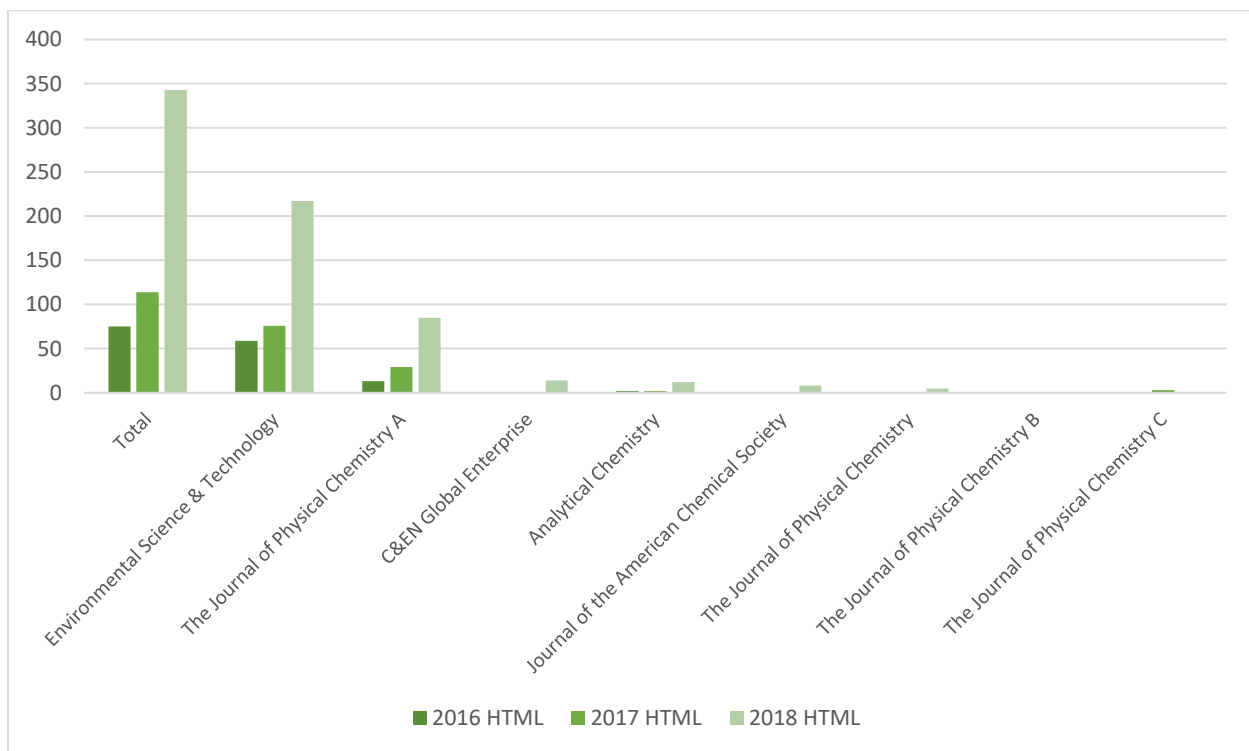**Figure 1** ACS PDF Downloads by Journal and Year, COUNTER4 JR1

Figure 2 ACS HTML Usage by Journal and Year, COUNTER4 JR1

While ACS usage looked promising, we also examined usage of *Science Magazine* from 2016 through 2018 (see Figure 3), prior to and after the implementation of the project, to understand usage via the former method versus via EZproxy authentication. The subscription to *Science* expired on the old IP-based account on January 31, 2018, so there was a month of overlap between the old and new methods of access. Usage data from 2018 shows a transition for users in the way they access, but there was still significant usage outside of EZproxy throughout the year. A more troubling trend was found when comparing total usage from 2016 through 2018. Use of *Science* in 2018 through the new access and discovery environment did not reach levels of usage in the previous environment, differing significantly from what was found with ACS. Is there a barrier to access via authentication for some content and not others? If so, what

are the discovery and access pathways that researchers are accustomed to with *Science*, and how might EZproxy be integrated into this pathway? Furthermore, does authenticating via a username and password have an impact on researchers' ability to access content?

Figure 3 AAAS Science Total Usage by Year, COUNTER4 JR1
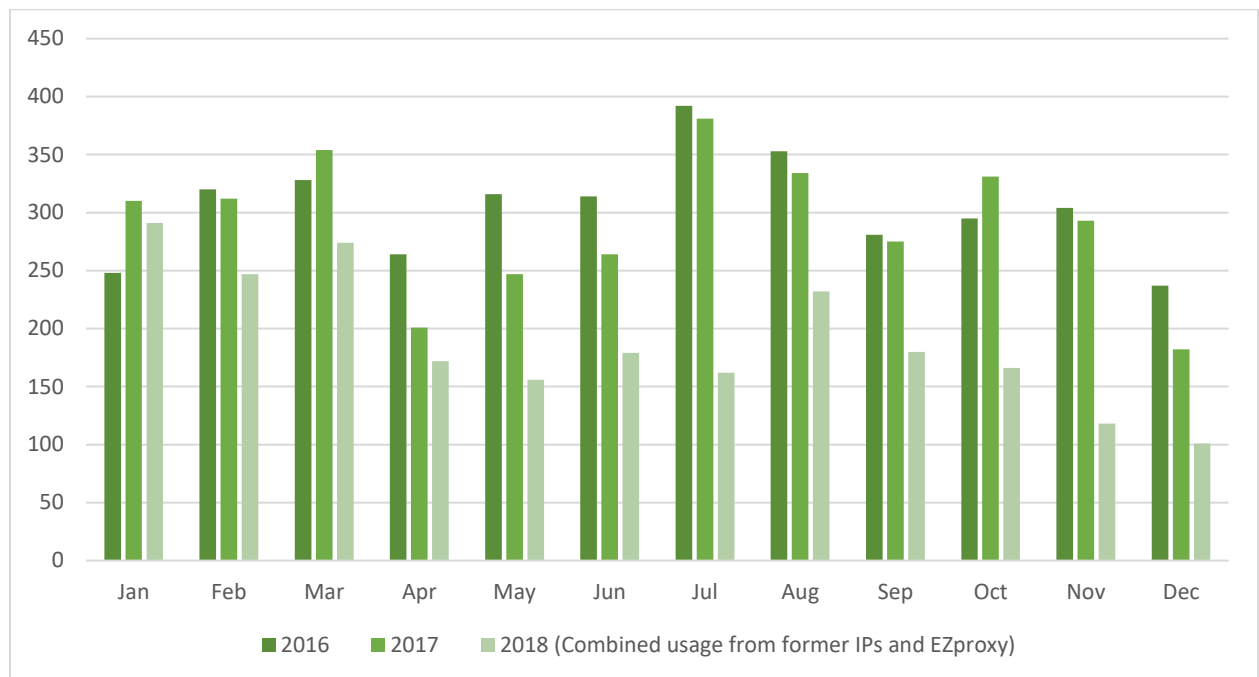
Finally, assessment of usage in most cases was limited, due to the way NCAR researchers were indistinguishable from CU Boulder faculty and students in COUNTER JR1 reports. We attempted various methods to extrapolate usage without COUNTER reports by analyzing link resolver statistics from SFX for NCAR and 360 Link for CU Boulder, but unfortunately, this method did not result in any meaningful conclusions, with the data showing inconclusive variation. Tools built for EZproxy log file analysis may be a potential solution for gaining better usage data in the future.

Conclusion

While this paper discusses one specific example of a collaborative research environment and the challenges libraries and researchers may face when attempting to provide access to licensed resources in a paywalled environment, it highlights questions and issues applicable beyond this specific case. Usage data thus far raise questions on whether or not more content is better when data suggest more content does not equate to higher usage. Changes in discovery and access, such as routing NCAR users through authenticated access via EZproxy, raise questions on whether these changes also create unintended barriers to access. Finally, a review of definitions of authorized users raises questions about whether librarians are doing enough to meet researchers' needs in a collaborative research environment, and what we can do to reduce barriers to content. This case study highlights the complexity of the collaborative research environment and how the paywalled model of access, combined with the way identity management and license agreements define who is considered a member of an institution, inhibit and challenge effective collaboration. Even when libraries put forth effort to support collaborative research, are they ultimately perpetuating a confusing, complex, and restrictive research environment because of their entanglement with paywalled content and authorization protocols? Is an open future of access and discovery a potential solution to eliminating barriers for collaborative research?

Metadata & Serials Librarian, for support in collaborating on this presentation and providing data.

NOTES

[1] Office of Information Technology University of Colorado Boulder, "Person of Interest (POI) HCM Records Practices," January 15, 2019, https://oit.colorado.edu/identikey-accounts/person-interest-poi-hcm-records-practices.

[2] "American Chemical Society Publications Division Online Products Institutional Access Agreement," accessed July 18, 2019, https://pubs.acs.org/pb-assets/documents/infocentral/ACS_Institutional_Access_Agreement_Academic-1559222638637.pdf.

[3] National Center for Atmospheric Research, "NCAR Author Publications Fact Sheet: 2005-2015," Spring 2016.

[4] Jennifer Phillips, email message to authors, June 3, 2019.

CONTRIBUTOR NOTES

Juleah Swanson is Head of Acquisitions Services, University of Colorado Boulder, Boulder, Colorado.

Steven W. Brown is E-Resources Access Manager, University of Colorado Boulder, Boulder, Colorado.