

State of Open At the University of Colorado Boulder

A Baseline Analysis of Open Access Practices from 2012 to 2018

July 16, 2018

Prepared by the Data and Scholarly Communication Services Unit, Department of Open and Digital Scholarship Services, University Libraries, University of Colorado Boulder

Andrew Johnson, *Assistant Professor and Head of Data and Scholarly Communication Services, University Libraries*

Katie Mika, *Assistant Professor and Data Services Librarian, University Libraries*

Melissa H. Cantrell, *Assistant Professor and Scholarly Communication Librarian, University Libraries*

Ryan Caillet, *Institutional Repository Program Manager, University Libraries*



University Libraries

UNIVERSITY OF COLORADO **BOULDER**

TABLE OF CONTENTS

I. Executive Summary	2
II. Introduction	3
1. Open Access to Articles and Data: A Brief History	3
2. Open Access at CU Boulder	4
III. Articles Published by CU Boulder Authors in Full Open Access Journals	6
IV. CU Boulder Libraries Open Access Fund	12
1. 2013-2017 Trends in the CU Boulder Open Access Fund	12
2. Trends in North American Open Access Funds	16
V. Open Access Content in CU Scholar	18
VI. Open Data at CU Boulder	19
VII. Conclusion	23
VIII. References	25

I. Executive Summary

Open Access (OA) to journal articles and research data is an increasingly important issue to researchers, institutions, funders, and journal publishers. University of Colorado Boulder (CU Boulder) researchers are engaging in OA practices at a growing rate, and the CU Boulder Libraries support a number of OA initiatives and services.

This report is an attempt to quantify, to the greatest extent possible given available data and methods, the current “State of Open” at CU Boulder. In addition to this initial effort to capture data on the OA activities underway at CU Boulder since 2012 when the Libraries first began coordinated efforts in this space, this report will serve as a baseline from which to measure future progress toward a more open research landscape on campus. Data used to produce this report can be found here:

<https://doi.org/10.25810/znb1-3k90>

Key findings from this report include:

- 10.20% of all articles published in 2017 by CU Boulder authors were published in full OA journals;
- From 2013 to 2017, the Libraries OA Fund funded author fees totalling \$163,522.40 for 116 journal articles published by CU Boulder authors in full OA journals;
- As of April 2018, there were 7,732 OA items in the CU Scholar institutional repository run by the Libraries, and these items were downloaded a total of 477,695 times in 2017 alone;
- Published data sets became a reportable scholarly activity in the annual Faculty Report of Professional Activities in 2014, and faculty reported 33 data sets in 2017 (up from 12 in 2014) with 16 of these citations including Digital Object Identifiers (DOIs).

II. Introduction

1. Open Access to Articles and Data: A Brief History

While there are several definitions of open access (OA), OA broadly refers to a world in which there are no costs to readers or users, beyond having an internet connection, to access and use scholarly information. In the case of journal articles, much OA work has arisen in response to increasingly unsustainable costs of traditional journal subscriptions that only allow access for researchers affiliated with individual subscribing institutions (Piwowar et al., 2018). With regard to research data, which in many fields is still only accessible to a small number of researchers, OA efforts are driven largely by the potential for reuse of data to advance the pace of discovery and enable new forms of inquiry while making research more transparent and reproducible.

Over the past two decades, a number of prominent efforts emerged to address the lack of openness in the way research is communicated. The Budapest Open Access Initiative (BOAI) was established in 2002 in order to find concrete solutions for making scholarly articles freely available (“Budapest Open Access Initiative,” n.d.). The BOAI provided one of the first definitions of OA, and it also established two strategies for achieving OA for journal articles that have been put into practice by numerous institutions, research funders, and publishers.

The first of these strategies involves self-archiving versions of journal articles (not always the final published version) in OA repositories like the National Institutes of Health (NIH) PubMed Central repository or the Harvard University DASH repository. OA repositories are often closely linked to OA policies that attempt to provide blanket copyright permissions for all researchers that are, for example, funded by an agency like NIH or employed by a specific institution like Harvard. Funder policies that require depositing some version of journal articles stemming from publicly and privately funded research in OA repositories are now widespread. Similarly, a large and growing number of academic institutions have taken it upon themselves to establish local policies governing primarily faculty researchers employed on their campuses. While these policies provide clarity on what researchers can legally do to make their articles openly available regardless of where they originally publish, OA policies can be quite difficult and labor-intensive to implement effectively.

The other common strategy for achieving OA involves establishing journals that utilize a number of funding models to cover the costs of OA publishing. Thousands of OA journals have been launched since the initial BOAI meeting in 2002, and many have become well-regarded in a number of disciplines. Some OA journals use a model that charges authors article processing charges (APCs) to cover the cost of publishing while making all published articles openly available to anyone with internet access. In order to

support researchers publishing in journals that use these models, many institutions have established OA funds to help cover or offset these costs. Some journal publishers have leveraged the momentum behind OA journals to create hybrid publishing models whereby researchers are charged APCs in order to make individual articles openly available to readers, yet the rest of the journal content (e.g., articles from researchers who do not elect to pay the APC) still requires an expensive subscription to access. While all OA journals increase the number of OA articles available to readers, there are still many questions surrounding the sustainability of various funding models.

OA to research data has largely been driven by research funders that require researchers to describe how they will make the data from funded research as openly and easily accessible as possible. Notably, the National Science Foundation (NSF) has required such a “Data Management Plan” for all grant proposals since 2011, and essentially all US federal funding agencies have now adopted such requirements (“Research Funder Data Sharing Policies,” n.d.; Nosek et al., 2015). Similarly, many journals have established data access policies that require authors to provide a means for publicly accessing the research data underlying journal articles at the time of publication. Typically, this is achieved by submitting data to one of a list of recommended research data repositories. Many disciplines and communities have established such repositories for various types of data, and a number of institutions support general repositories that are available to researchers on their campuses. While the importance of open data to funders and publishers continues to grow, journal articles are still by and large the primary currency of academic reward systems, including promotion and tenure. A great deal of work is still required to elevate research data to the level of a first class research output.

Academic libraries are involved with all of these efforts to promote OA to articles and research data. Libraries do this in a number of ways, including advocating for institutional OA policies, establishing OA funds, developing and promoting OA repositories for publications and data, and helping researchers navigate an increasingly complex landscape of funder and publisher requirements for OA.

2. Open Access at CU Boulder

In 2016-2017, CU Boulder received \$508 million in research funding with \$354 million coming from federal agencies (“Research and Creative Work 2016-17,” 2017). According to data from CU Boulder Elements (CUBE), CU Boulder researchers used or will use this funding in part to produce approximately 5,000 journal articles per year (“CU Boulder Elements (CUBE),” 2015). A very rough estimate of one data set per article would also yield around 5,000 research data sets produced at CU Boulder per year. All told, CU Boulder creates a significant amount of research articles and data,

much of which stems from federal funding, so efforts to help make this research openly available could be quite impactful.

The CU Boulder Libraries began coordinated efforts to promote OA in 2012. Much of the work related to OA for journal articles took place under the auspices of the Scholarly Communication Working Group. This group led the initial efforts to establish:

- an OA Fund to pay for APCs for CU Boulder researchers publishing in high quality OA journals (launched in 2012);
- an OA repository called CU Scholar (<https://scholar.colorado.edu/>) to provide a platform for openly disseminating CU Boulder research (launched in 2014);
- a campus-wide faculty OA policy in order to make it easier for CU Boulder researchers to make scholarly articles openly available (adopted in 2015).

Also in 2012, the Libraries began engaging in campus efforts to promote open research data, most notably via the establishment of Research Data Services, which was a collaboration involving the Libraries, Research Computing, and the Office of the Vice Chancellor for Research (OVCR). Research Data Services began as a result of the recommendations of the OVCR's Data Management Task Force (Vice Chancellor for Research's Data Management Task Force, 2012). Research Data Services helped launch:

- a consulting service to assist researchers with making data openly available, often in conjunction with funder or publisher policies, via disciplinary and institutionally-supported data repositories and platforms;
- an institutionally-supported tool for complying with research funder data management plan requirements (DMPTool);
- an institutionally-supported platform for managing and making research data openly available (Open Science Framework for Institutions);
- inclusion of published data sets as reportable scholarly activities in the annual report of Faculty Report of Professional Activities (FRPA).

In support of all of these activities, the Libraries engaged in many outreach and educational efforts aimed at increasing awareness and understanding of OA for both publications and data. This work and the management of the majority of the established OA services on campus now reside in the Libraries' Open and Digital Scholarship Services Department and the Center for Research Data and Digital Scholarship, a partnership between the Libraries and Research Computing. Many other units within the Libraries (e.g., the Scholarly Resource Development Department) and on campus play key roles in this collaborative work as well.

What follows in this report is an attempt to quantify, to the greatest extent possible given available data and methods, the current “State of Open” at CU Boulder. This report utilizes data from CU Scholar, CUBE, the OA Fund, and FRPA. In addition to this initial effort to capture data on the OA activities underway at CU Boulder since 2012 when the Libraries first began coordinated efforts in this space, this report will serve as a baseline from which to measure future campus progress toward a more open research landscape. All data used to produce this report can be found here: <https://doi.org/10.25810/znb1-3k90>

III. Articles Published by CU Boulder Authors in Full Open Access Journals

An analysis of all articles published in full OA journals by CU Boulder authors between 2012 and 2017 revealed a relatively consistent rate of OA publishing. All data in this section comes from CUBE and uses inclusion in the Directory of Open Access Journals as the criterion for a full OA journal. Table 1 shows the percentage of OA articles published by CU Boulder authors by year.

Table 1. Percentage of OA Articles Published by CU Boulder Authors by Year

2012	2013	2014	2015	2016	2017
9.48%	9.35%	10.14%	11.11%	10.83%	10.20%

While OA publishing increased slightly across campus during the years studied, certain units were more heavily involved in OA publishing than others. Seven departments or institutes were in the top five highest OA publishing units one or more years between 2012 and 2017. Those units are the Cooperative Institute for Research in Environmental Sciences (CIRES), Molecular, Cellular & Developmental Biology (MCDB), Physics, Chemistry and Biochemistry, Psychology and Neuroscience, Mechanical Engineering, and Ecology and Evolutionary Biology.

Due to high rates of intra- and inter-departmental collaboration on campus, quantifying the rate of participation in OA publishing for academic and research units required a metric beyond the total number of articles published by authors in a particular unit. Instead we totaled the number of OA contributions to account for multiple CU authors working on the same paper. A unit’s contributions were based on how many times its authors were credited as an author or co-author across all OA publications for that year. Tables 2-6 show the top ten departments with OA contributions by year.

Table 2. 2012: Number of OA Article Contributions by Department

Department/Unit	Number of OA Contributions
Cooperative Institute for Research in Environmental Sciences (CIRES)	56
Chemistry and Biochemistry	35
Molecular, Cellular & Developmental Biology (MCDB)	33
Physics	33
Psychology and Neuroscience	32
Ecology and Evolutionary Biology	21
Integrative Physiology	21
Institute of Arctic & Alpine Research (INSTAAR)	20
Electrical, Computer and Energy Engineering (ECEE)	16
Mechanical Engineering	15

Table 3. 2013: Number of OA Article Contributions by Department

Department/Unit	Number of OA Contributions
Cooperative Institute for Research in Environmental Sciences (CIRES)	56
Chemistry and Biochemistry	34
Ecology and Evolutionary Biology	32
Psychology and Neuroscience	28
Molecular, Cellular & Developmental Biology (MCDB)	24
Physics	24
Institute of Arctic & Alpine Research (INSTAAR)	16
Atmospheric and Oceanic Sciences (ATOC)	15
Integrative Physiology	13
Mechanical Engineering	13

Table 4. 2014: Number of OA Article Contributions by Department

Department/Unit	Number of OA Contributions
Cooperative Institute for Research in Environmental Sciences (CIRES)	73
Chemistry and Biochemistry	42
Ecology and Evolutionary Biology	32
Molecular, Cellular & Developmental Biology (MCDB)	31
Psychology and Neuroscience	25
Mechanical Engineering	18
Physics	17
Integrative Physiology	17
Atmospheric and Oceanic Sciences (ATOC)	17
Institute of Arctic & Alpine Research (INSTAAR)	14

Table 5. 2015: Number of OA Article Contributions by Department

Department/Unit	Number of OA Contributions
Cooperative Institute for Research in Environmental Sciences (CIRES)	141
Molecular, Cellular & Developmental Biology (MCDB)	48
Chemistry and Biochemistry	40
Mechanical Engineering	34
Physics	29
Institute for Behavioral Genetics (IBG)	28
Ecology and Evolutionary Biology	24
Psychology and Neuroscience	23
Integrative Physiology	15
Atmospheric and Oceanic Sciences (ATOC)	15

Table 6. 2016: Number of OA Article Contributions by Department

Department/Unit	Number of OA Contributions
Chemistry and Biochemistry	58
Molecular, Cellular & Developmental Biology (MCDB)	44
Cooperative Institute for Research in Environmental Sciences (CIRES)	42
Ecology and Evolutionary Biology	34
Physics	27
Mechanical Engineering	24
Psychology and Neuroscience	24
Atmospheric and Oceanic Sciences (ATOC)	23
Integrative Physiology	19
Aerospace Engineering Sciences	15

Table 7. 2017: Number of OA Article Contributions by Department

Department/Unit	Number of OA Contributions
Cooperative Institute for Research in Environmental Sciences (CIRES)	74
Chemistry and Biochemistry	58
Molecular, Cellular & Developmental Biology (MCDB)	38
Physics	36
Atmospheric and Oceanic Sciences (ATOC)	35
Ecology and Evolutionary Biology	32
Mechanical Engineering	28
Integrative Physiology	19
Electrical, Computer and Energy Engineering (ECEE)	17
Geography	16

The analysis of OA articles published at CU Boulder also revealed strong trends in journal selection. Certain journals consistently ranked highly in the number of articles published every year studied. Atmospheric Chemistry and Physics, PLoS One, Optics

Express, Atmospheric Measurement Techniques, and Scientific Reports all ranked in the top ten journals every year from 2012 to 2017. The mean for number of articles published per journal grew from 2.85 in 2012 to 3.29 in 2015 even though the majority of journals only published one OA article from CU authors in a given year. Tables 8-13 show the top ten OA journals by the number of articles published by CU Boulder authors by year.

Table 8. 2012: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
PLoS One	49
Atmospheric Chemistry and Physics	45
Optics Express	22
Atmospheric Measurement Techniques	13
Scientific Reports	8
Cryosphere	7
PLOS Genetics	6
Atmospheric Chemistry and Physics Discussions	5
International Journal of Antennas and Propagation	5
Nucleic Acids Research	5

Table 9. 2013: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
PLoS One	48
Atmospheric Chemistry and Physics	42
Optics Express	18
Atmospheric Measurement Techniques	11
Scientific Reports	10
BIOGEOSCIENCES	9
Frontiers in Psychology	8
Cryosphere	7
Environmental Research Letters	7
Nucleic Acids Research	6

Table 10. 2014: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
Atmospheric Chemistry and Physics	60
PLoS One	55
Atmospheric Measurement Techniques	19
Cryosphere	13
Atmospheric Chemistry and Physics Discussions	11
Scientific Reports	11
Optics Express	10
Frontiers in Psychology	8
Nucleic Acids Research	7
ELIFE	6

Table 11. 2015: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
Atmospheric Chemistry and Physics	64
PLoS One	44
Scientific Reports	21
Atmospheric Measurement Techniques	20
Optics Express	13
Environmental Research Letters	10
BIOGEOSCIENCES	8
ELIFE	8
New Journal of Physics	8
PHYSICAL REVIEW SPECIAL TOPICS-PHYSICS EDUCATION RESEARCH	8

Table 12. 2016: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
Atmospheric Chemistry and Physics	52
PLoS One	38
Scientific Reports	28
Atmospheric Measurement Techniques	18
Optics Express	13
New Journal of Physics	11
Cell Reports	9
Environmental Research Letters	9
Cryosphere	8
ELIFE	8

Table 13. 2017: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
Atmospheric Chemistry and Physics	49
PLoS One	27
Atmospheric Measurement Techniques	25
Scientific Reports	25
Optics Express	17
Journal of Advances in Modeling Earth Systems	11
Geoscientific Model Development	10
ELIFE	9
Environmental Research Letters	9
Cell Reports	6

IV. CU Boulder Libraries Open Access Fund

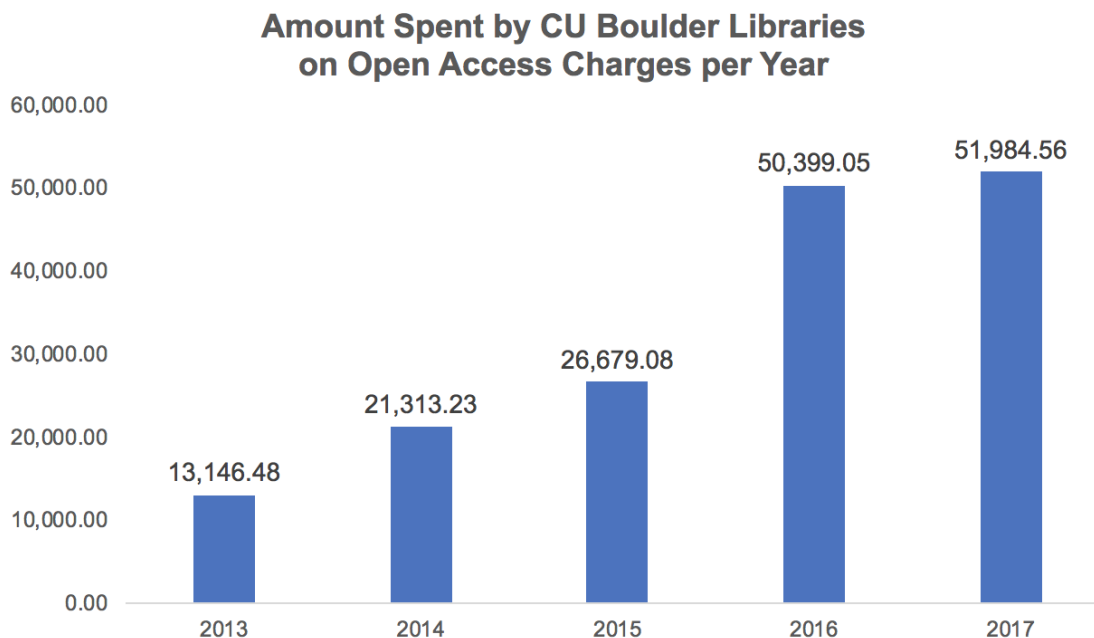
1. 2013-2017 Trends in the CU Boulder Open Access Fund

In 2012, the CU Boulder Libraries initiated an OA Fund to serve as a promotional tool for OA journal publishing as well as to offset the article processing costs (APCs) being charged to CU Boulder faculty who were engaging in OA publishing. The OA

Fund provides up to \$2,000 per year to each student, staff, or faculty member publishing in a fully OA and peer reviewed journal.

Between 2013 and 2017, the CU Boulder Libraries have funded a portion or total amount of author fees for 116 journal articles. Over five years the Libraries have made a total investment of \$163,522.40, and the amount of funding that the Libraries have invested in the OA Fund has increased every year since 2013:

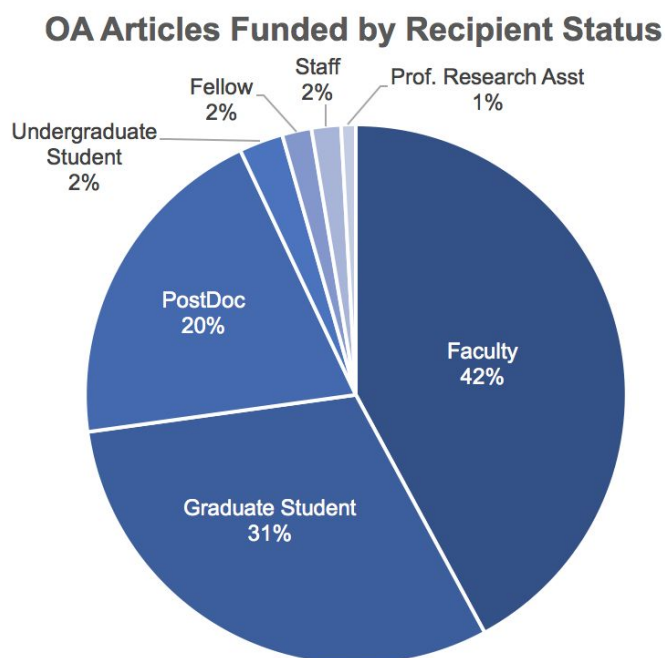
Figure 1



The average APC amount that was funded through the program was \$1,421.93. The smallest amount approved for funding was \$200 and the largest \$2,000 (this being the limit of the fund). Some APCs do exceed \$2,000, in which case authors pay the remainder through grant funds, departmental funds, or other sources.

Unsurprisingly, faculty (as a single category) by far apply for and receive the most funding assistance through the program. Over two-fifths (42%) of funds granted go to faculty, while roughly half of the funds go to graduate students and postdocs (31% and 20%, respectively). The remaining 8% of funds go to undergraduate students, staff, fellows, and one professional research assistant:

Figure 2

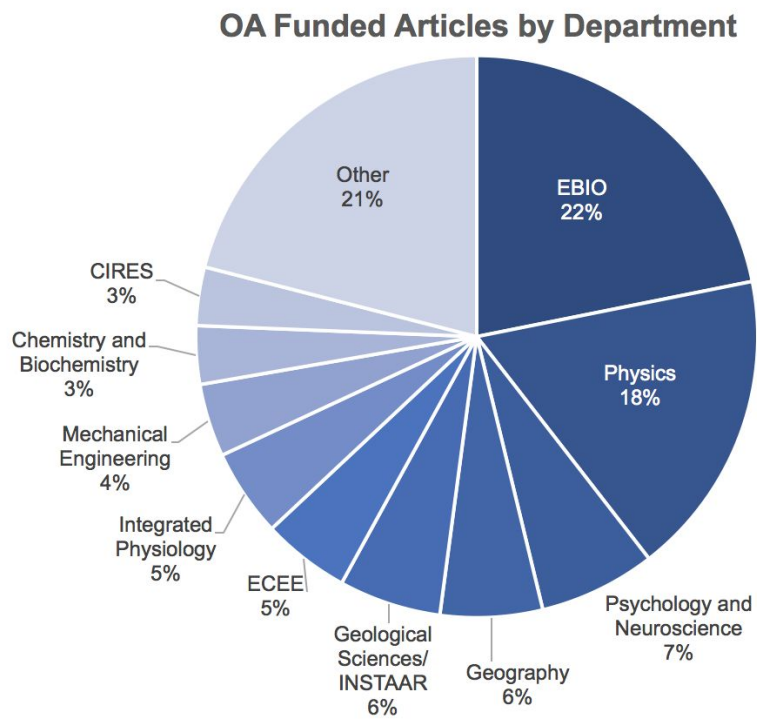


Although “PhD student” and “Master’s student” are both selections on the application, 41% of all graduate level applicants used the blank field to self-select as a “Graduate Student,” so it is not possible to draw meaningful comparisons across graduate students at different levels.

By department, Ecology and Evolutionary Biology (EBIO) and Physics receive by far the largest proportion of the funding. A total of 24 different departments across the university have had recipients of the OA Fund, yet these two departments represent two-fifths (40%) of the funded articles with a combination of 47 approved applications. This equals \$62,665.81 worth of funding between these two departments, which is more than a third (38%) of all of the funds that have been distributed since 2013. In fact, the Physics department itself has received more OA funding than 18 other funded departments combined.

The OA Fund is also much more heavily used by STEM departments than any other broad disciplinary category. Only five departments or centers outside of STEM have had a recipient of the fund: The Latin American Studies Center, Media Research and Practice, the School of Education, Silicon Flatirons Center at the Law School, and Sociology. Out of these five, Sociology has utilized the fund the most with three successful applicants. No humanities department has had an author either apply for or receive OA funding through this program.

Figure 3



The OA Fund has been used to fund more than one article in just 20 of the 55 journals in which OA Fund recipients have published, as illustrated by Table 14.

Table 14. OA Journals Receiving OA Funds More Than Once

Journal	Number of times paid via CU Boulder OA Fund
PlosOne	18
Nature Communications	8
Optics Express	8
EcoSphere	6
New Journal of Physics	5
Nucleic Acids Research	4
Physical Review X	4
Environmental Research Letters	3
Zookeys	3
Atmospheric Chemistry and Physics	2
Biomedical Optics Express	2
BMC Public Health	2
Ecology & Evolution	2
Frontiers in Human Neuroscience	2
Frontiers in Microbiology	2
Frontiers in Psychology	2
International Journal of Environmental Research and Public Health	2
Neurobiology of Sleep and Circadian Rhythms	2
Science Advances	2
Scientific Reports	2

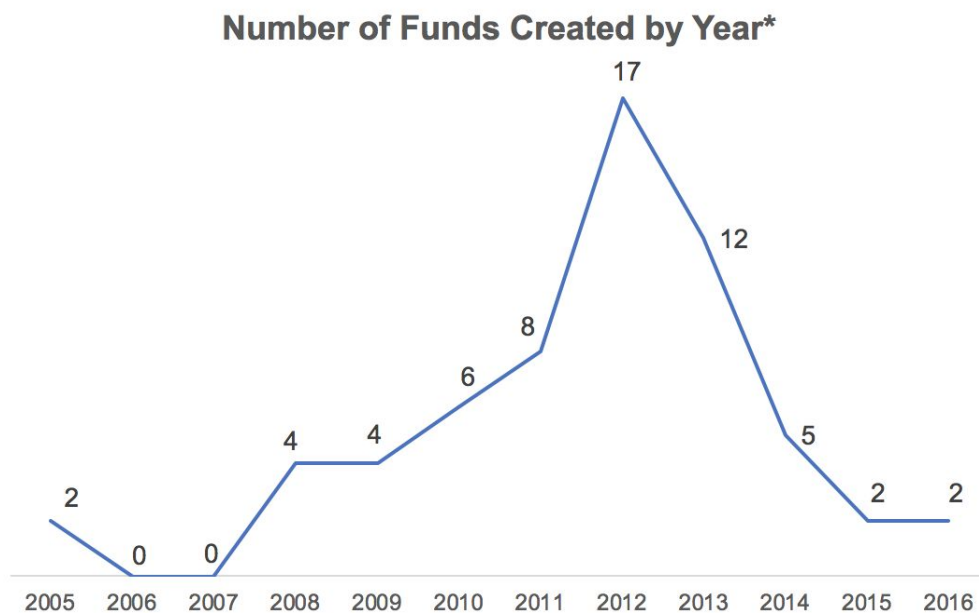
PlosOne, Optics Express, and Nature Communications have received the most funding from the Libraries OA Fund, with a combined \$51,279.25 paid to these three journals since 2013. That means that just these three journals have received almost a third (31%) of all of the funds distributed by the OA Fund.

2. Trends in North American Open Access Funds

According to the SPARC Open Access Funds in Action 2017 report, 63 universities reported data about OA funds, although approximately 19 explicitly state that their funds are no longer active (“Open Access Funds in Action,” 2017). The oldest OA funds were created by University of Wisconsin - Madison and University of North Carolina - Chapel Hill in 2005, but neither of these funds are still active according to the

latest report. The longest running programs for OA funding come from University of Calgary, University of California, Berkeley, University of Tennessee - Knoxville, and Wake Forest University, which all began operating funds in 2008. The newest funds, both launched in 2016, come from Wayne State University and the University of Pennsylvania. CU Boulder created its fund in the peak year of 2012 when 17 universities reported creating funds for OA publishing:

Figure 4



*One university did not report the date of fund establishment

Since 2005, these 63 universities report 7,654 articles approved for funding. The average total number of articles funded per school is 120, and CU Boulder sits slightly under that average with 99 total articles funded as of 2016 (the most recent year covered in the report). There is a large, spread, however, between the universities that have accepted the most and the least articles for funding. University of California, Riverside funded the least number of articles (5), but the program only ran for a single year. The University of Calgary has funded the most (1,105), but it has been continuously running since 2008 and boasts a budget of \$100,000 per year.

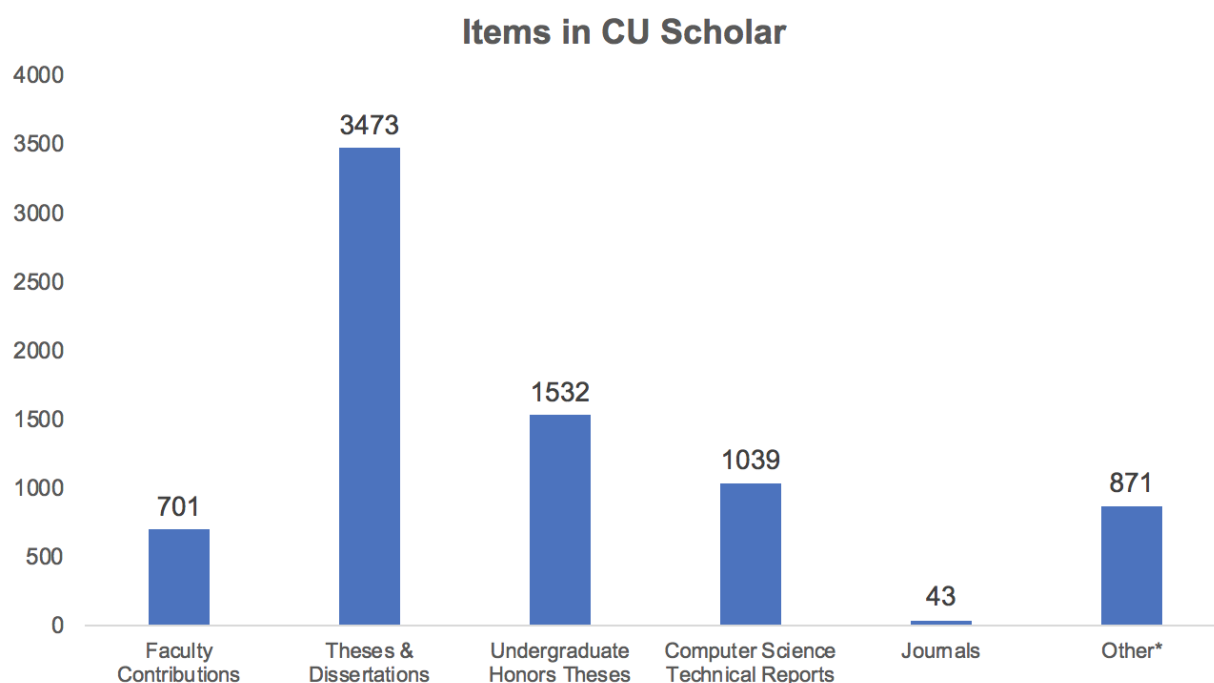
Based on the data there does not appear to be a “normal” or “average” method of running an OA fund. Most funds are very dependent on local context and available resources, and there has been a series of funds closing since 2013 coupled with a clear drop-off in the creation of new funds (see Figure 4). Although closings may be due to the depletion of a one-time fund for these kinds of programs, this drop-off may suggest that some universities find that there is not a clear or advantageous return on

investment for continuing to support or for launching a fund for OA publishing, although the reasons for cancellation are not specifically stated.

V. Open Access Content in CU Scholar

As of April 2018, CU Scholar contained 7,732 items including faculty publications, graduate theses and dissertations, undergraduate honors theses, journals, event series, conferences, and graduate and undergraduate contributions (other than theses and dissertations). Graduate theses and dissertations make up 44.92% of all content in the repository and account for approximately 24% of all downloads from CU Scholar. Faculty and staff contributions make up 9.07% of the repository and account for 11.06% of all downloads from CU Scholar. Undergraduate honors theses account for 19.81% of items in the repository yet drive 64.90% of all downloads.

Figure 5



A series of computer science technical reports makes up 13.44% of the repository content, but it is not actively growing beyond a few items per year. Hosted journals, conferences, event series, and graduate and undergraduate contributions (other than theses and dissertations) combined represent less than 1% of the repository's contents and downloads.

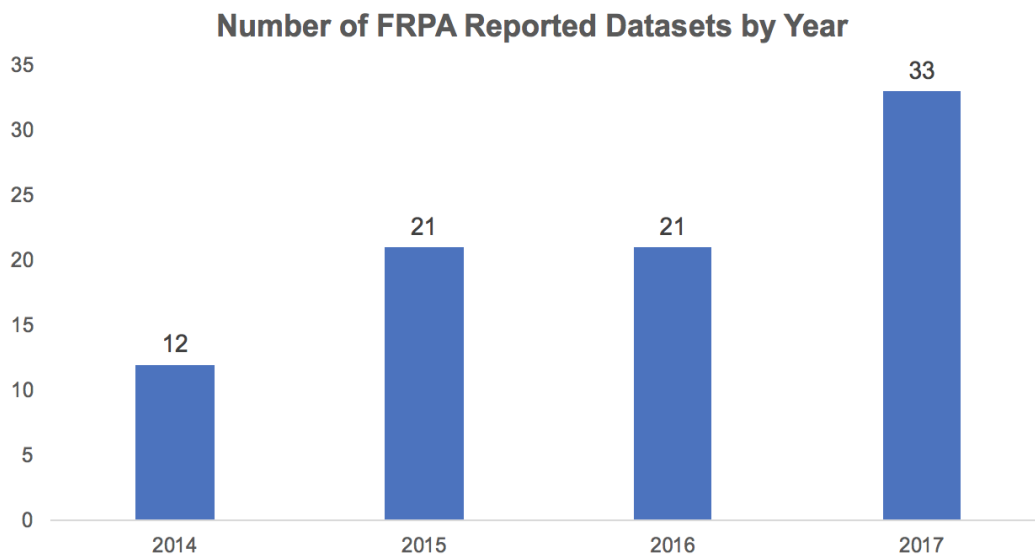
Use of the repository varies by department and academic unit on campus. The units most active with regard to uploading their scholarship are the University Libraries with 106 submissions, Physics with 96 submissions, Chemistry & Biochemistry with 42 submissions, and MCDB and Astrophysical and Planetary Sciences each with 33 submissions to CU Scholar. Around 70 departments on campus have yet to deposit any faculty research into CU Scholar on their own accord. The Libraries are addressing these gaps and the low rate of self-deposit across academic disciplines through mediated deposit projects, especially with regard to implementing the campus OA policy for faculty articles.

CU Scholar has seen a steady increase in usage since its inception. CU Scholar saw 107,266 full-text downloads in 2015 which was the first full year the repository was operational. In 2016 that number grew to 227,711, an increase of 112.29%. Similar growth (109.78%) was seen in 2017 with 477,695 downloads.

VI. Open Data at CU Boulder

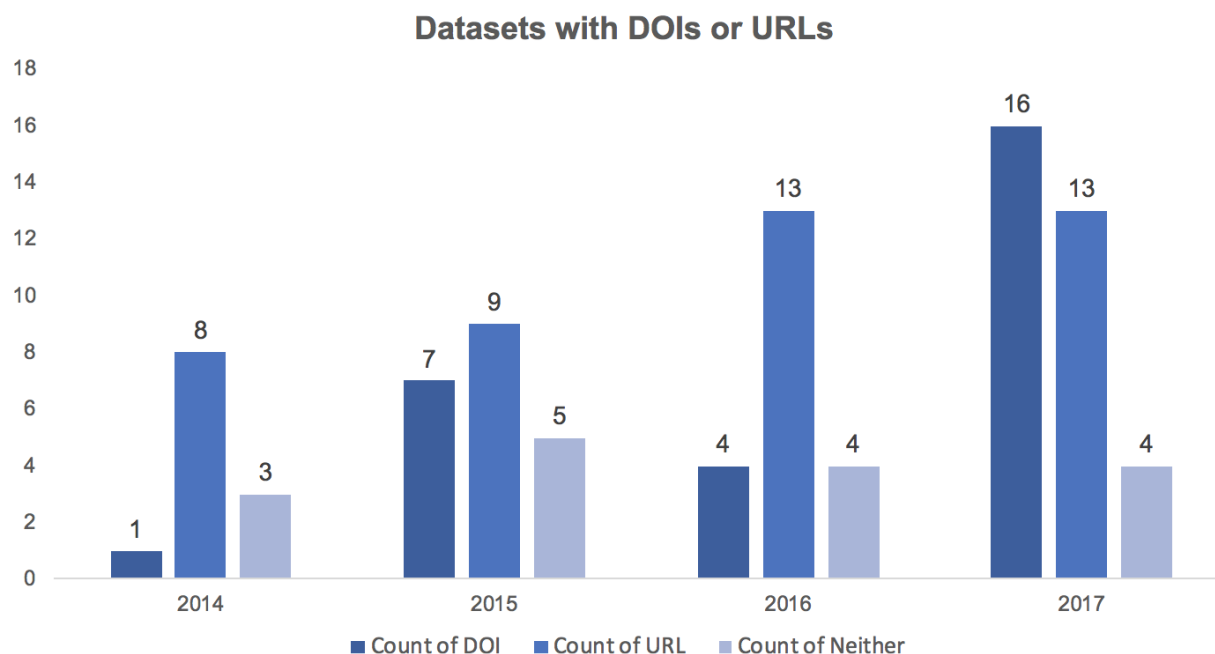
In addition to funder and journal publisher policies that promote open data, the Open Science movement has also contributed to the rise of research data as a first-class scholarly output. Driven in large part by the success of the OA movement, Open Science specifically promotes the open publication of research data in order to improve scientific inquiry, research, and communication (Molloy, 2011). Beginning in 2014, faculty at CU Boulder were extended the option to include published data in their Faculty Reporting of Published Activities (FRPA). A faculty member's FRPA collects evidence of teaching activity; scholarly work; grants, commissions, and fellowships; creative work; service activity; honors and awards; and works in progress, and is considered the official record of scholarly outputs for CU Boulder faculty ("What is FRPA?," 2016). All tenure-track faculty and full time instructors are required to submit the FRPA as part of the annual merit evaluation process. Published data is included as a category of scholarly work and is entered as a citation in a free form text box with instructions to "indicate if a dataset is refereed, and include information for identification and access, such as a DOI or URL" ("FRPA Activity Codes," 2016). According to reporting data from 2014-2017, there is a clear increase in faculty including data citations in annual FRPA reports.

Figure 6



The Digital Object Identifier (DOI) is an emerging standard for unique and persistent identification of published data. DOIs support accurate citation of data and resolve to URLs in order to provide increased discovery, access, and reuse of data. Furthermore, published data with a DOI, if best practices are followed, may have a high standard of curation and rich metadata. Due to the widely varying types of research data that CU considers as scholarly outputs for FRPA reporting, DOIs, while encouraged, are not required for publication. URLs, different types of unique identifiers, or other access details are also recommended. The quality of citations for published data in FRPA varies greatly. This may be due to a lack of communication or guidance regarding standards and best practices for publishing and citing data. Publishing data is a relatively recent phenomenon in many disciplines, and assigning DOIs for data is an even newer practice. Although there was a decrease in the number of citations containing DOIs from 2015 to 2016, the overall number of links to data increased. In 2017, nearly half (48.5%) of the data sets reported included DOIs.

Figure 7



Beginning in 2018, published data reported in FRPA will be included in CU Experts profile pages populated with data from CU Boulder’s Symplectic Elements instance (CUBE). The opportunity to enter citation metadata directly into Symplectic Elements should improve the quality of citations by requiring fields for DOIs, URLs, publishers, etc. As open data practices become more widespread, domain and data type standards for this information are likely to become more robust as well.

Departments across campus are adopting open data practices at different rates. According to the FRPA reports, CU Boulder departments and institutes publishing the most data are INSTAAR, CIRES, and EBIO; however, the absence of data reported in FRPA does not indicate an absence of published data in general. It is important to note that reporting data in FRPA is at the discretion of the faculty member, and faculty in many departments are publishing data that is not being reported in FRPA, including data in repositories like the Open Science Framework, Figshare, Dryad, and PANGAEA as well as in locally managed data centers. Due to the novelty and recency of the practice of publishing research data, many faculty may not yet consider data sets to be significant scholarly outputs. Table 15 shows the total number of datasets reported and authors reporting data sets by department in the FRPA system.

Table 15. Reported Datasets by Department, 2014-2017

Department	Number of Reported Datasets	Number of Authors Reporting Data
INSTAAR	12	6
CIRES	18	5
EBIO	14	4
Linguistics	8	2
LASP	6	2
Computer Science	4	3
Leeds	2	2
Classics	2	2
University Libraries	2	2
Journalism	2	1
Media Studies	2	1
Physics	2	1
CASA	1	1
Environmental Studies	1	1
Natural History Museum	1	1
Atmospheric and Oceanic Science	1	1
Aerospace Engineering	1	1
CEAE	1	1
History	1	1
Music	1	1
Information Science	1	1
Education	1	1
Asian Languages and Civilizations	1	1
Geography	1	1
Geology	1	1

One way to increase data publication and reporting is to develop an incentive structure that rewards such practices. Currently, research published in journal articles, conference proceedings, and monographs earn premium credit while research data (and similar outputs) are often considered less valuable. As Open Science practices take greater hold in the research landscape, faculty reporting of published data will likely continue to increase. Introducing a minimal level of metadata curation may improve data citations as well. The CU Boulder Libraries have recently launched DOI minting and data curation services in order to impact the quality of published data on campus. In

addition, incentivising researchers to report data publishing activities in external repositories will also help to track a larger number of data sets produced by CU Boulder researchers. To that end, academic departments may consider offering credit for publishing data as a first class scholarly output. Efforts to identify, report, and make available openly published research data are a first step toward enabling such a shift in incentive structures.

VII. Conclusion

The findings in this report indicate that all OA activities at CU Boulder for which there is available data have steadily increased in use since the Libraries began formal initiatives for promoting and supporting OA to journal articles and research data in 2012. Due to a lack of standardization in reporting of this type of data nationally, it is difficult to know where CU Boulder stands relative to other institutions, but the numbers in this report seem consistent with published and anecdotal evidence of OA practices at other universities. With regard to the use of CU Scholar for archiving OA copies of journal articles, it should be noted that efforts to implement the campus-level OA policy that was adopted in 2015 only began in late 2017 due to vacancies in key positions within the Libraries. The findings in future reports should help to evaluate the effectiveness of efforts to utilize the OA policy to systematically increase the amount of OA journal articles in CU Scholar.

OA activities do seem to be concentrated in a few key departments/institutes on campus, including Physics, EBIO, CIRES, Chemistry and Biochemistry, Psychology and Neuroscience, Mechanical Engineering, and Integrative Physiology. It will be interesting to track if/how these disciplinary trends change going forward as outreach across campus expands. The findings from this report indicate that particular attention should be paid to many humanities and social science departments.

Even though there is much overlap in kinds of OA engagement (Physics, Chemistry and Biochemistry, and MCDB all self-archive and publish in full OA journals at some of the highest rates), there are other inconsistencies that cannot yet be accounted for with available data. If MCDB is one of the most active departments in self-archiving in CU Scholar and in full OA journal publishing, why has that department rarely utilized the OA Fund (with only 2 submissions over 5 years)? If CIRES shows the highest engagement with full OA journal publishing almost every year, why are they not in the top 5 in self-archiving and represent only 3% of OA Fund submissions? From the open data perspective, while some of the highest reporters of published data also engage in OA publishing, none of the three top engagers in OA journals and

self-archiving mentioned above (Physics, Chemistry and Biochemistry, and MCDB) have significant reported involvement with published data in the past two years.

Each of these areas of inconsistent engagement across departments raises a larger problem: it is very difficult to demonstrate the existence of or an increase in awareness or advocacy of open practices. Engagement, even among the most active participants, is too scattered to demonstrate an awareness or desire to employ the full range of open practices. A major question that these data raise, then, is how to move from engagement to advocacy? Moving closer to an answer will probably involve addressing some of the limitations of the available data. While the ability to track OA publishing has improved and is now quite reliable, tracking open data is far more difficult. While the Libraries are working to improve local procedures, the capacity to accurately quantify open research data at our institution is largely dependent on the systems and platforms through which that data is made available. Another major limitation, and one that prevents building a bridge from use to awareness and then to advocacy, is a lack of understanding of the perception of OA and open data among researchers at CU Boulder.

Despite these limitations, this report contributes to a better understanding of OA activities at CU Boulder and possibly serves as a model for other institutions to adopt in the interest of making data like this more widely available. Because reports like this, if implemented more widely and consistently, could provide a fuller picture of the progress towards OA and open research data at the institutional level, the data sets used to generate this report are openly available as well (Johnson et al., 2018). While the findings in this report are useful for capturing a snapshot of the “State of Open” on campus, there are still important questions that these numbers are not able to answer. Further exploration will be required in order to better understand the level of awareness and understanding of OA-related issues on campus as well as the benefits CU Boulder researchers have received or perceived as a result of engaging in these OA activities.

VIII. References

- Budapest Open Access Initiative. (n.d.). Retrieved June 12, 2018, from <http://www.budapestopenaccessinitiative.org/>
- CU Boulder Elements (CUBE). (2015, September 28). Retrieved June 13, 2018, from <https://www.colorado.edu/fis/CUBE>
- FRPA Activity Codes. (2016, July 19). Retrieved June 13, 2018, from <https://www.colorado.edu/fis/frpa-help-activity-codes>
- Johnson, A., Caillet, R., Cantrell, M. H., & Mika, K. (2018). Data from 'State of Open at the University of Colorado Boulder' report. University Libraries, University of Colorado Boulder. <https://doi.org/10.25810/ZNB1-3K90>
- Molloy, J. C. (2011). The Open Knowledge Foundation: Open data means better science. *PLOS Biology*, 9(12), e1001195. <https://doi.org/10.1371/journal.pbio.1001195>.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- Open Access Funds in Action (2017, November 10). Retrieved June 13, 2018, from https://sparcopen.org/wp-content/uploads/2015/11/OA-Funds-in-Action_Nov-11-17.pdf
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., ... Haustein, S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>

Research and Creative Work 2016-17. (2017, October 27). Retrieved June 13, 2018, from <https://www.colorado.edu/research/report/2016-17>

Research Funder Data Sharing Policies. (n.d.). Retrieved June 12, 2018, from <https://sparcopen.org/our-work/research-data-sharing-policy-initiative/funder-policies/>

Vice Chancellor for Research's Data Management Task Force (2012). Research data management at the University of Colorado Boulder: Recommendations in support of fostering 21st century research excellence. Office of the Vice Chancellor for Research, University of Colorado Boulder. <https://scholar.colorado.edu/ovcr/1>

What is FRPA? (2016, October 28). Retrieved June 13, 2018, from <https://www.colorado.edu/fis/frpa>