

2020 State of Open at the University of Colorado Boulder

An Update on Open Access Practices Based on Data from 2019

August 3, 2020

Prepared by the Data and Scholarly Communication Services Initiative, Center for Research Data and Digital Scholarship, University Libraries, University of Colorado Boulder

*Andrew Johnson, Associate Professor, Director, Data and Scholarly Communication Services Initiative
Melissa H. Cantrell, Assistant Professor, Scholarly Communication Librarian
Ryan Caillet, Institutional Repository Program Manager*



University Libraries

UNIVERSITY OF COLORADO **BOULDER**

TABLE OF CONTENTS

I. Executive Summary	2
II. Articles Published by CU Boulder Authors in Full Open Access Journals	3
III. CU Boulder Libraries Open Access Fund	14
IV. Open Access Content in CU Scholar	19
V. Open Data at CU Boulder	20

I. Executive Summary

Using data from 2019, this report is the second annual update to the “State of Open at the University of Colorado Boulder: A Baseline Analysis of Open Access Practices from 2012 to 2018”: <https://doi.org/10.25810/vprn-v113>. It includes analyses of open access (OA) journal publishing, OA repository usage, and open data practices by researchers at the University of Colorado Boulder (CU Boulder). Data used to produce this report can be found here: <https://doi.org/10.25810/p5fa-y621>.

Key findings from this report include:

- 9.95% of all articles published in 2019 by CU Boulder authors were published in full OA journals, which is up slightly from 9.63% in 2018;
- In 2019, the CU Boulder Libraries OA Fund funded author fees totaling \$100,588.42 for 65 journal articles published by CU Boulder authors in full OA journals (up from \$91,041.36 for 57 articles in 2018);
- As of January 2020, there were 11,426 OA items in the CU Scholar institutional repository (up from 10,638 at the time of the previous report), and these items were downloaded a total of 760,400 times in 2019 (up from 625,325 in 2018);
- In the annual Faculty Report of Professional Activities (FRPA), faculty reported 65 published data sets in 2019 (up from 56 in 2018) with 44 of these citations including Digital Object Identifiers (DOIs) (up from 34 in 2018) and 56 identifying a formal data repository (reported for the first time in this report);
- The Libraries and its partners registered 197 DataCite DOIs for published data sets in 2019 (up from 112 in 2018).

II. 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 2019 revealed a relatively consistent rate of OA publishing. In 2019, 3,797 unique articles were published by CU Boulder authors. 3,419 of those articles were published in journals not included in the Directory of Open Access Journals (DOAJ) while 378 articles were published in DOAJ journals. 303 of the DOAJ articles were added to the CU Scholar institutional repository through our DOAJ workflow. The remainder were either excluded due to authors not being affiliated with the university at the time of publication, conflicting rights statements from the publisher, or the articles already having been added to the repository through other workflows. 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	2018	2019
9.48%	9.35%	10.14%	11.11%	10.83%	10.20%	9.63%	9.95%

OA publishing remained consistent across campus during the years studied, but certain units were more heavily involved in OA publishing than others. Eight 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, Atmospheric and Oceanic Sciences, and Ecology and Evolutionary Biology. Chemistry and Biochemistry splitting into two independent departments contributed to their lower publication totals 2018-present, but both remain in the top ten.

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 2-9 show the top ten departments with OA contributions by year.

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

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/Unit

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/Unit

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/Unit

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/Unit

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/Unit

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

Table 8. 2018: Number of OA Article Contributions by Department/Unit

Department/Unit	Number of OA Contributions
Cooperative Institute for Research in Environmental Sciences (CIRES)	80
Mechanical Engineering	44
Ecology and Evolutionary Biology	37
Molecular, Cellular & Developmental Biology (MCDB)	29
Psychology and Neuroscience	22
Chemistry	20
Atmospheric and Oceanic Sciences (ATOC)	20
Physics	19
Integrative Physiology	18
Geological Sciences	18
Biochemistry	18
Institute of Arctic & Alpine Research (INSTAAR)	17
Civil, Environmental and Architectural Engineering	16

Table 9. 2019: Number of OA Article Contributions by Department/Unit

Department/Unit	Number of OA Contributions
Mechanical Engineering	35
Chemistry	34
Atmospheric and Oceanic Sciences (ATOC)	34
Civil, Environmental and Architectural Engineering	34
Cooperative Institute for Research in Environmental Sciences (CIRES)	32
Ecology and Evolutionary Biology	30
Molecular, Cellular & Developmental Biology (MCDB)	27
Physics	19
Integrative Physiology	18
Biochemistry	17
Electrical, Computer and Energy Engineering (ECEE)	16
Psychology and Neuroscience	15
Environmental Studies Program	15

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 2019. Tables 10-17 show the top ten OA journals by the number of articles published by CU Boulder authors by year.

Table 10. 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 11. 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 12. 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 13. 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 14. 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 15. 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

Table 16. 2018: Number of OA Articles Published by CU Boulder Authors by Journal

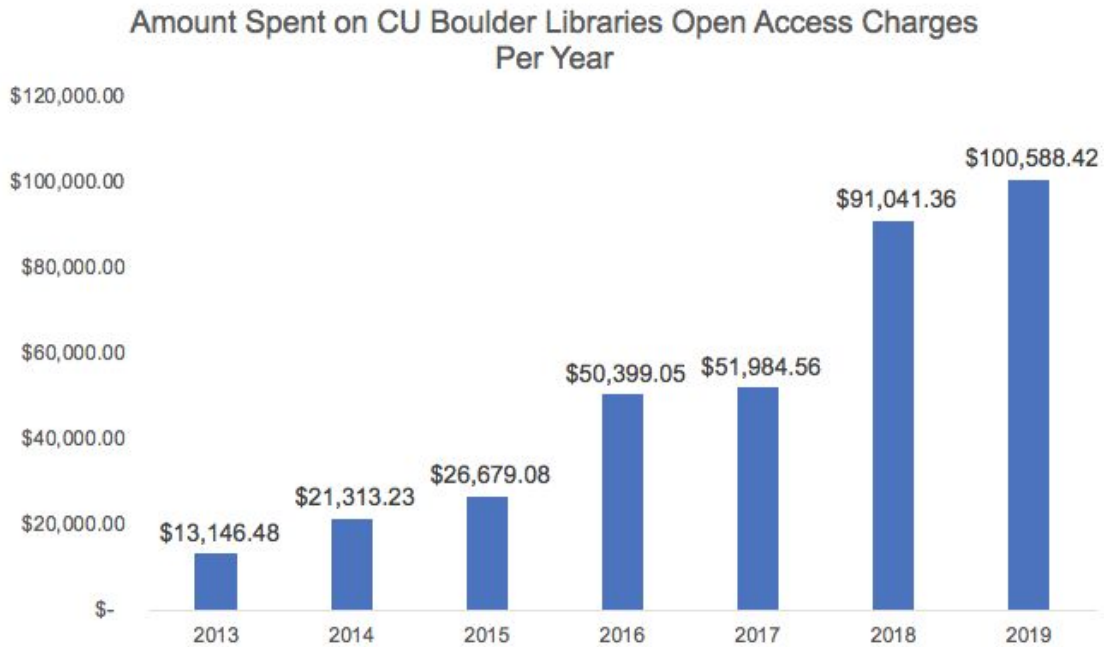
Journal	Number of OA Articles Published
Atmospheric Chemistry and Physics	46
Scientific Reports	28
PLoS ONE	27
Atmospheric Measurement Techniques	11
Optics Express	11
The Cryosphere	8
Environmental Research Letters	7
New Journal of Physics	7
eLife	6
Frontiers in Microbiology	6
Geoscientific Model Development	6
PLoS Genetics	6

Table 17. 2019: Number of OA Articles Published by CU Boulder Authors by Journal

Journal	Number of OA Articles Published
Atmospheric Chemistry and Physics	46
PLOS ONE	36
SCIENTIFIC REPORTS	25
Atmospheric Measurement Techniques	20
SUSTAINABILITY	17
JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS	15
CRYOSPHERE	14
OPTICS EXPRESS	13
ENVIRONMENTAL RESEARCH LETTERS	10

III. CU Boulder Libraries Open Access Fund

Figure 1

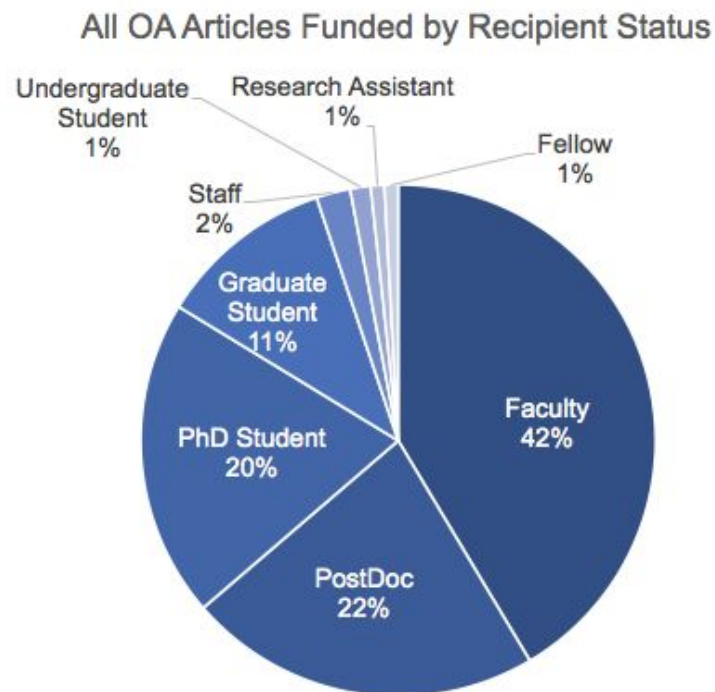


The CU Boulder Libraries spending on article processing charges (APCs) in fully open access journals for faculty, staff and students continued to rise in 2019, with just over \$100,000 (\$100,588.42) spent for the year. Sixty-five total articles were funded in 2019, for an average APC cost of \$1,547.51. Similar to 2018, most of the APCs were paid in the latter half of 2019 after the fiscal year renewal, with only fourteen being paid from January to July 2019¹.

Cumulatively, the CU Boulder OA Fund has helped authors publish fully open access articles in 100 unique journal titles. Of the 65 journal articles funded in 2019, 22 of them were published in journal titles not previously funded through the CU Boulder OA Fund. It's possible this could be indicative of an acceleration in the diversity of fully OA journals in which authors are interested in publishing.

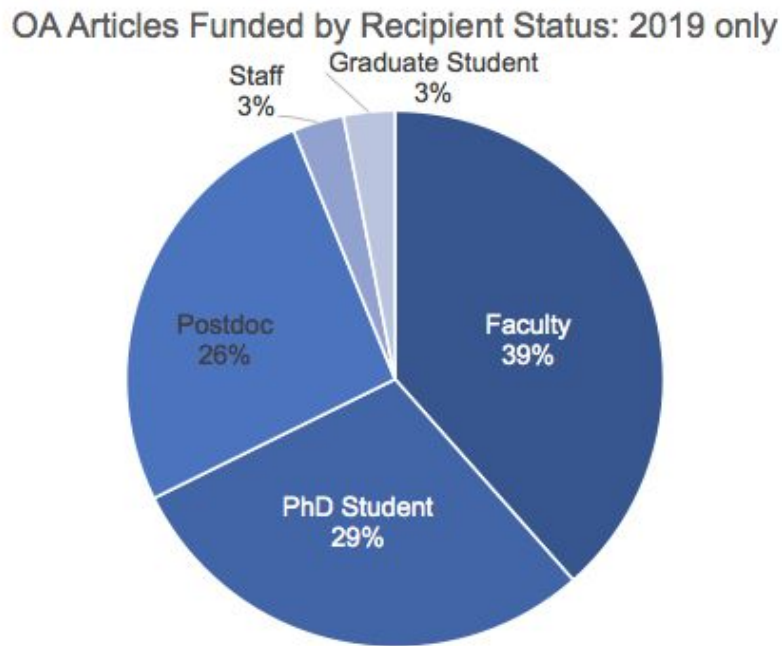
¹ There were requests made in Spring 2019 that could not be fulfilled due to funding being depleted.

Figure 2



The proportion of funded authors by university status over the lifetime of the OA Fund remains largely unchanged, as faculty still represent just over two-fifths of funded articles (42%) and graduate students and postdoctoral researchers represented a little over half (53%) of the articles funded. Research Assistants, Staff, Fellows, and undergraduate students collectively represent 5% of the total articles funded since the inception of the OA Fund in 2013.

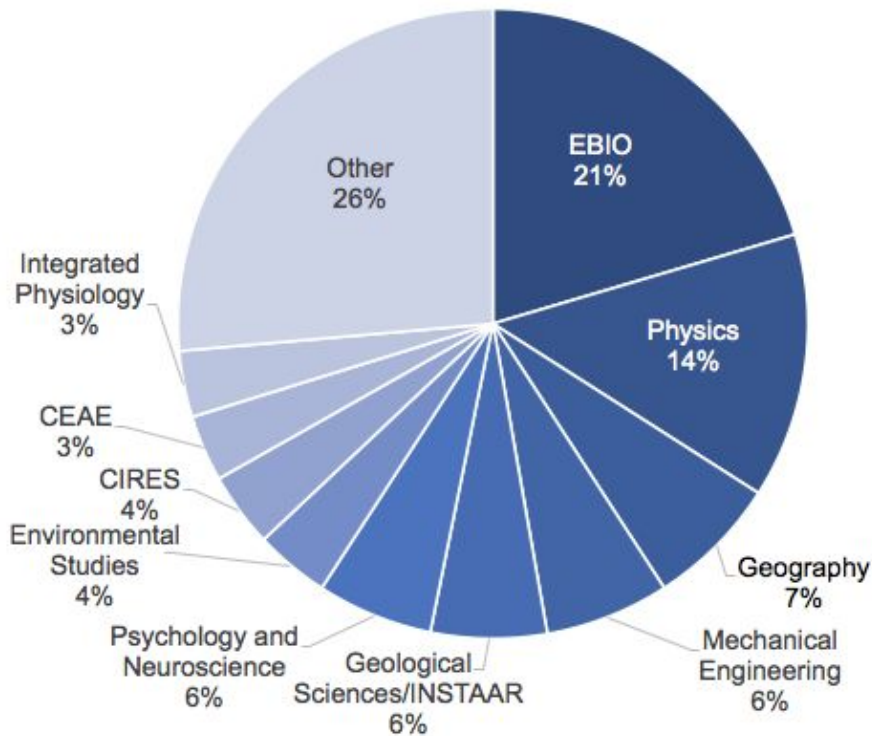
Figure 3



Articles funded in 2019 only by recipient status also look similar to previous years, with some small changes. While faculty represented just over two-fifths of funded articles in 2018, that proportion decreased slightly in 2019, with just under two-fifths (39%) of articles funded going to faculty. This shift is accompanied by a slight increase in funding recipients indicating their status as a graduate student, PhD student, or postdoctoral researcher - from 53% of funded articles in 2018 to 58% in 2019. The percentage of recipients identifying specifically as a “graduate student” has remained well below the cumulative portion of those identifying as such for the past couple of years.

Figure 4

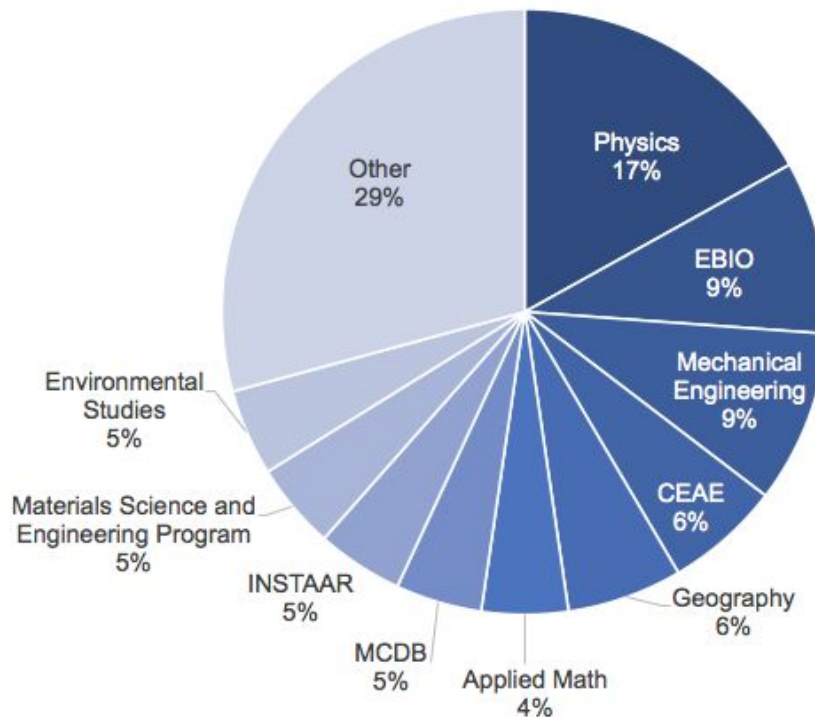
OA Funded Articles by Department - Cumulative



Cumulative trends for the departments whose articles were funded also follow previous patterns, with some slight changes. EBIO fell from 24% of all articles funded to 21%, while Physics held a steady share of funded articles at 14%. Additionally, CEAE rose into the top ten departments receiving funding, while ECEE fell off of that list. The proportion of “Other” departments having articles funded also rose from 21% in 2018 to 26% in 2019. Departments and programs with individuals receiving funding for the first time in 2019 include Journalism, the Program in Environmental Design, Wardenberg Medical Services, and three articles originating from the Materials Science and Engineering Program.

Figure 5

OA Funded Articles by Department: 2019 Only

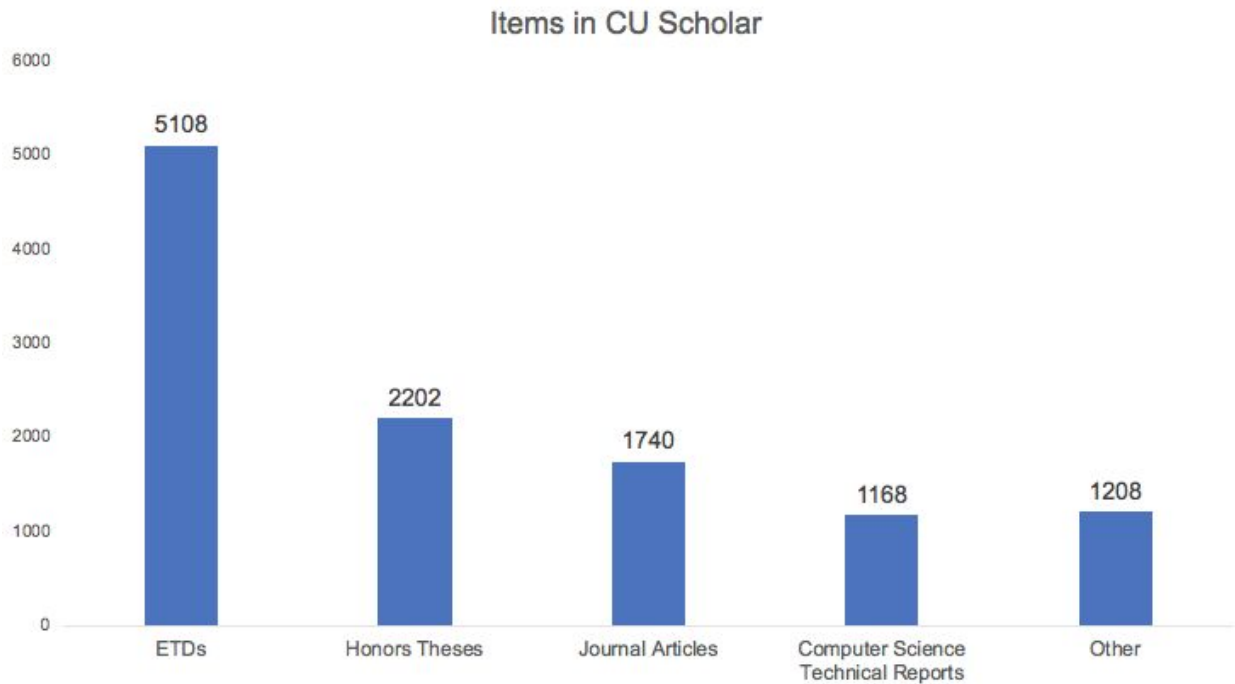


When isolating the 2019 data of articles funded by the department of the author recipient, there are some stark differences compared to 2018. While in 2018 EBIO represented a quarter of funding recipients and Physics only 5% of recipients, in 2019 authors from the Physics department were awarded the highest proportion of article funds (17%) and the share of funds awarded to EBIO was only 9%. CEAE, which in 2018 was not in the top ten funded departments, tied with Geography as the fourth most funded department in 2019 (each receiving 6% of the total fund awards). Chemical and Biological Engineering, Psychology and Neuroscience, and CIRES, which were in the top ten funded departments in 2018, were replaced in 2019 by Applied Math, CEAE, and the Materials Science and Engineering Program.

IV. Open Access Content in CU Scholar

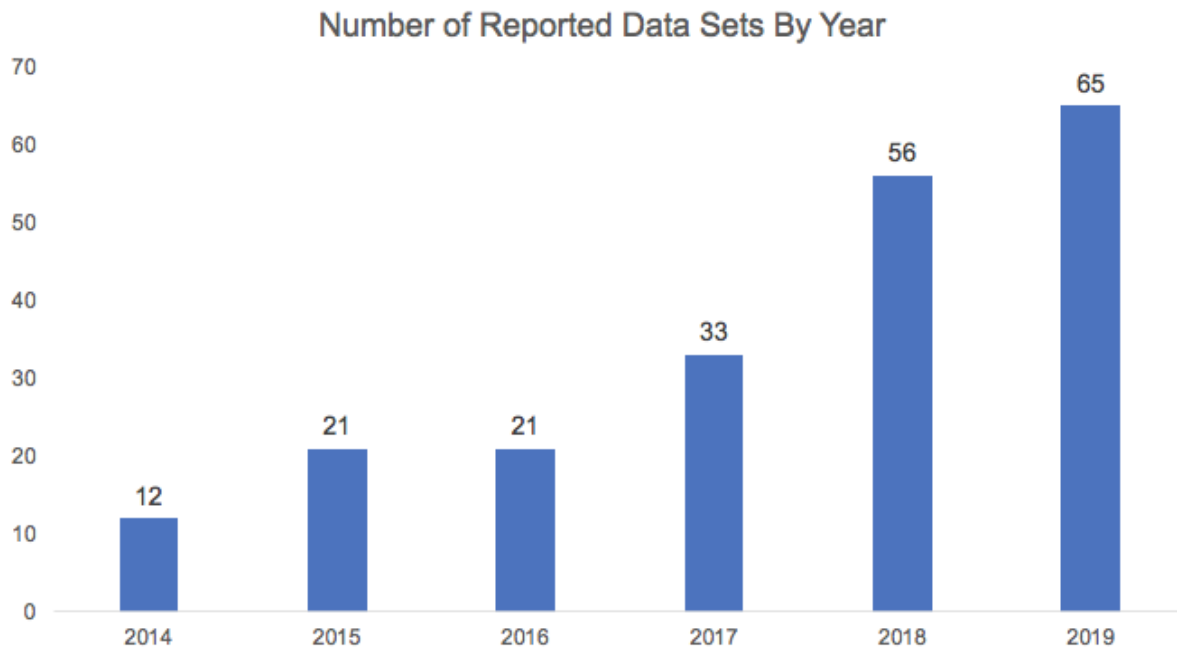
At the beginning of 2020, CU Scholar migrated from the bepress Digital Commons platform to the open source Samvera repository software. As a result, we will be changing the way we report open access content in CU Scholar going forward. Data will now be reported on the calendar year beginning with the updated report next year, which will contain data for 2020 (the first full year on the new Samvera platform). For this 2019 update, we will report the final data at the time of the migration off of the bepress Digital Commons platform at the end of 2019. At that time, CU Scholar contained 11,426 items including journal articles, data sets, graduate theses and dissertations, undergraduate honors theses, conference materials, books, and book chapters (up from 10,638 as of July 2019). CU Scholar items were downloaded a total of 760,400 times in 2019, and the total download count for all items in CU Scholar for all years prior to migration is 2,198,397. At the time of migration, graduate theses and dissertations made up 44.70% of all content in the repository. Journal articles accounted for 15.23% while undergraduate honors theses comprised 19.27% of repository items. Technical reports represented 10.22% of the repository with most reports belonging to a computer science technical reports collection that is not actively growing.

Figure 6: Items in CU Scholar



V. Open Data at CU Boulder

Figure 7

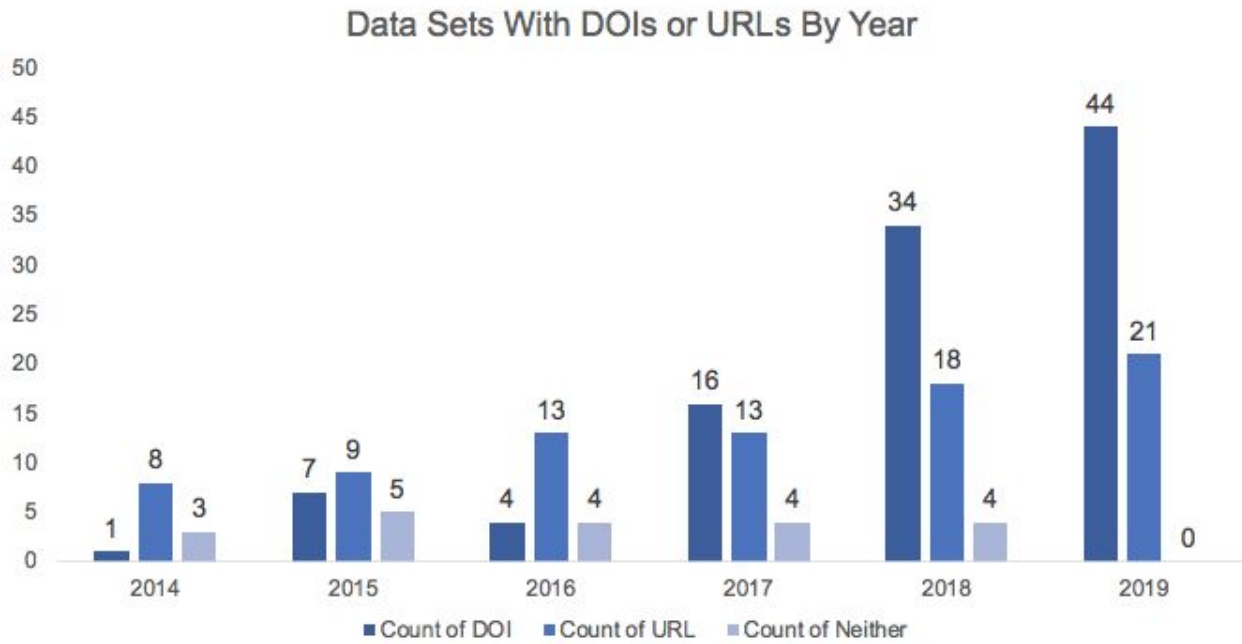


Reporting of published data sets on the annual Faculty Report of Professional Activities (FRPA) continued to increase steadily from 56 data sets in 2018 to 65 in 2019 (see Figure 7). This suggests a modest yet growing acceptance of data sets as research outputs worth reporting for the purposes of annual faculty evaluations. As might be expected, the departments/units with the largest number of reported data sets ($n > 10$) are primarily in disciplines where funding largely comes from funders with data sharing requirements (e.g., the National Science Foundation) and where journal publisher policies for data availability are increasingly common (e.g., American Geophysical Union). One new department reported published data sets in 2019 (Speech, Language, and Hearing Sciences) while all other reported data sets came from departments with at least one reported data set since 2014. Table 16 provides the total number of published data sets reported in FRPA for all departments with more than one data set since 2014.

Table 16. Reported Datasets by Department/Unit, 2014-2019 (n>1)

Department/Unit	Number of Reported Datasets
Cooperative Institute for Research in Environmental Sciences	53
Institute of Arctic and Alpine Research	24
Ecology and Evolutionary Biology	22
Atmospheric and Oceanic Science	12
Linguistics	10
Astrophysical and Planetary Sciences	9
Computer Science	9
Geography	7
Laboratory for Atmospheric and Space Physics	7
Geology	6
Leeds School of Business	6
Libraries	5
Chemical and Biological Engineering	3
Civil, Environmental, and Architectural Engineering	3
Education	3
Environmental Studies	3
History	3
Information Science	3
Media Studies	3
Aerospace Engineering	2
Asian Languages and Civilizations	2
Classics	2
Journalism	2
Natural History Museum	2
Physics	2
Speech, Language, and Hearing Sciences	2

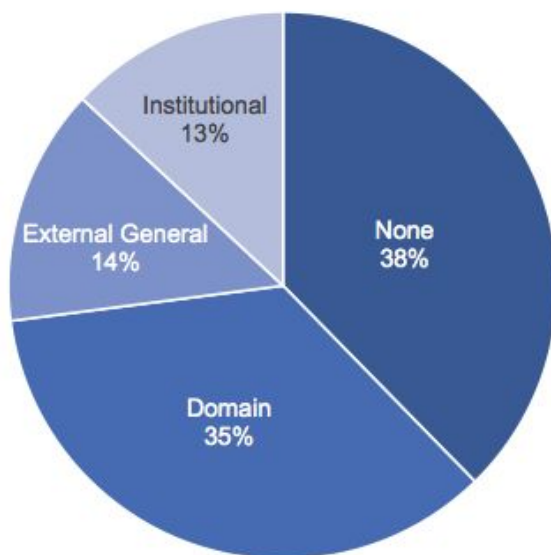
Figure 8



Use of DOIs for reported data sets increased from 34 in 2018 to 44 in 2019. The percentage of data sets with DOIs also increased to a new annual high of 67.7%. For the first year since 2014, all data sets included either a DOI or a URL for accessing the data. This trend suggests that data set citations, including a means for accessing data, are becoming a more common and consistent practice for faculty.

Figure 9

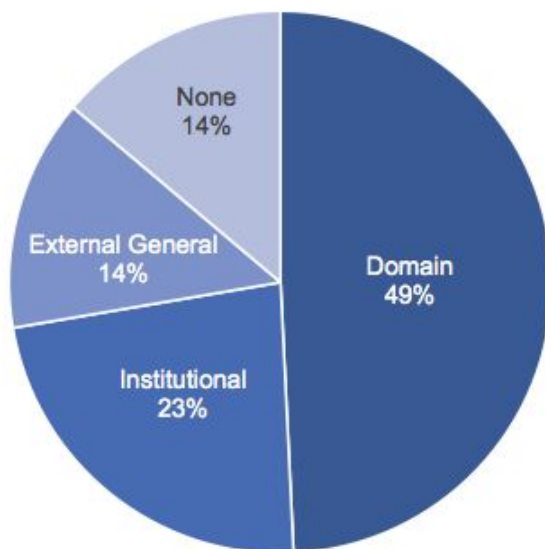
Repositories Used By Type, 2014-2019



With the use of DOIs steadily climbing, which suggests an increase in the use of formal repositories for publishing data sets, we decided to include data on repository types for the first time in this report update. Overall, a plurality (38%) of the reported published data sets dating back to 2014 have not used a formal repository. Instead, these data sets are commonly posted on individual faculty, research group, or project websites in addition to using other methods of publishing data sets. Of the data sets in formal repositories, 35% are housed in domain repositories that provide access to data from particular disciplines and/or to specific types of data (e.g., ICPSR, NSF Arctic Data Center, Protein Data Bank, etc.). General repositories that are external to CU Boulder and cover a wide range of disciplines and data types (e.g., figshare, Dryad, Zenodo, etc.) provide access to 14% of the reported published data sets in FRPA since 2014. Institutional repositories account for 13% of the reported published data sets with all but two of these data sets appearing in CU Scholar (the other two are housed in the Stanford University institutional repository).

Figure 10

Repositories Used By Type, 2019



In contrast to the overall picture of repository types since 2014, the data sets reported in 2019 were overwhelmingly published in formal repositories (only 14% do not appear in some type of repository). Nearly half (49%) of these data sets are housed in domain repositories while 23% appear in an institutional repository (all but one in CU Scholar). External general repositories account for the remaining 14% of data sets reported in 2019 that are published in repositories. It will be interesting to see if these trends toward formal repository use and domain/institutional repository use continue in future years.

Finally, in 2018, the Libraries began actively curating data sets in the CU Scholar institutional repository, including registering DataCite DOIs for every published data set. In addition, the Libraries provide DataCite DOI registration capabilities to a small number of campus partners through formal agreements. In 2019, the Libraries and its partners registered 195 DOIs for published data sets (up from 112 in 2018). The disconnect between this number of DOIs and published data sets reported by faculty could be due to a number of factors. Many of the data sets published by the Libraries are recurring data sets that receive a new DOI for every update but are only reported as a single data set for the purposes of annual faculty reports. Also, some data sets published by the Libraries and its partners were created by individuals other than faculty (e.g., graduate students, staff). It is also possible that some data sets might not be considered appropriate for faculty annual reports for a number of reasons. For example, data sets supporting journal articles might be seen as duplicative when the journal

article is already reported. Among other reasons, this could also account for some of the discrepancy between the presumably much larger number of data sets produced by CU Boulder faculty in domain and external general repositories than are reported by faculty annually.