5-6-2024

# Are We Relying on the Wrong Data? Analysis of E-journal Usage Data at Seven R1 and R2 Research Universities

Simon Robins
*University of Dayton*, srobins1@udayton.edu

# Are we relying on the wrong data? Analysis of e-journal usage data at seven R1 and R2 research universities

Simon Robins
University of Dayton Libraries, Dayton, Ohio
srobins1@udayton.edu
https://orcid.org/0000-0002-3264-5968

**ABSTRACT**
Using data from seven R1 and R2 universities in Ohio, this study seeks to determine the extent of usage of full-text e-journal articles stored across multiple platforms. Specifically, COUNTER 5 data is presented from five major academic journal publishers (Taylor & Francis, SAGE, Oxford, Wiley, and Springer), looking at usage within each of the publishers' respective platforms, as well as in EBSCO, JSTOR, and OhioLINK's Electronic Journal Center platform. The overarching goal of the analysis is to identify where exactly usage occurred for this e-journal content within the different plat- forms, especially in cases where an individual title has duplicate full-text access. Analysis of this data was conducted using Tableau and Excel. Findings raise significant questions about end-user behavior in terms of students' and faculty's ability to locate full text, regardless of whether it is discoverable from Google or other commercial search engines. The findings also suggest a strong connection between EBSCO usage and the academic calendar, with peak usage (as a percentage of overall usage) occurring in October, November, March, and April. Lastly, findings for Taylor & Francis e-journals within EBSCO call into question what these duplicate holdings mean for "big deal" subscription packages.

**KEYWORDS:** COUNTER 5; Wiley; SAGE; Taylor and Francis; Springer; Oxford University Press; EBSCO discovery service; electronic journals; data visualization

## Introduction

At many universities and colleges across the country, a customized version of EBSCO Discovery Service (EDS) is available alongside subscriptions to EBSCO's extensive collection of subject data- bases, providing users with varying degrees[1] of access to thousands of full-text e-journals (and millions of journal articles). These full-text holdings within EBSCO include content from several prominent publishers, most notably Taylor & Francis, SAGE, Oxford, Wiley, and Springer. The end-user does not have to leave EBSCO to view or download full-text e-journal articles from these and other publishers, and when users search within EDS, the full-text content from EBSCO subject databases feeds directly into the EDS interface. This is different from using the discovery layer to link out to external platforms (see **Figure 1**).[2]

When universities have large subscriptions to e-journal content with major publishers, in addition to these subscriptions to EBSCO databases and EDS, significant overlap can occur between the full-text content provided by EBSCO and the full-text content provided by publisher plat- forms such as Wiley Online Library or Taylor & Francis Online. Additional subscriptions in plat- forms like JSTOR and Project MUSE, which aggregate full-text content from multiple publishers, can also contribute to content overlap. Moreover, in certain states like Ohio, there can be overlap with any sort of state-run consortial database that stores the full-text, as is the case with OhioLINK's Electronic Journal Center (EJC).[3]

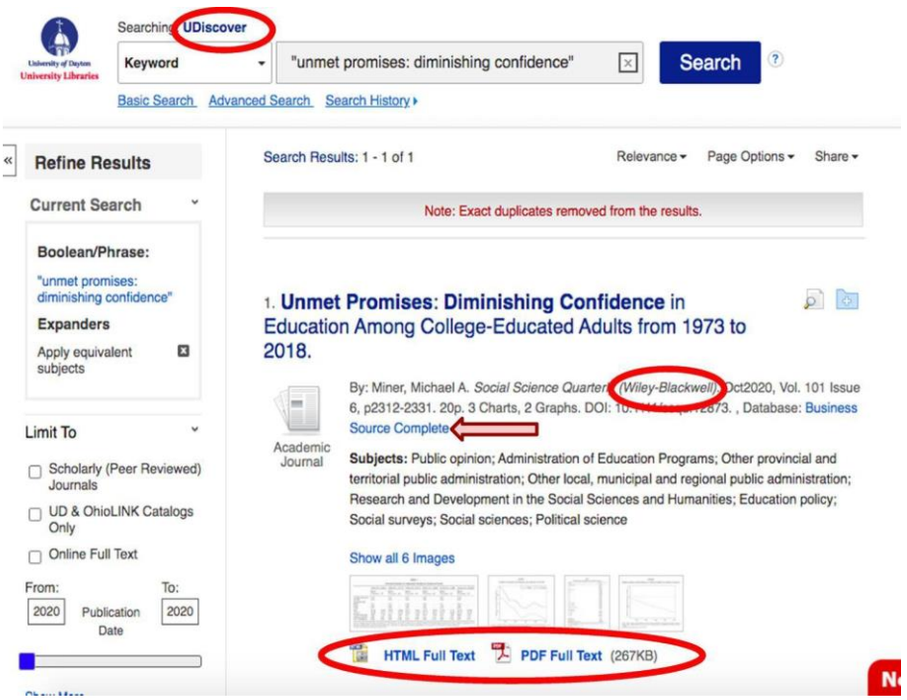



Figure 1. This search result is found within UD's EDS search (UDiscover). It is provided by Wiley, and the "PDF Full Text" link takes the user to the full text *without leaving EBSCO.* In addition to this full-text version, UD and all other universities featured in this study have access to the full text *via* Wiley Online Library, as well as a full-text copy through OhioLINK's EJC platform. In other words, there are three separate PDFs in three separate platforms for the exact same article.

This duplication of content has implications for both the end user and for the librarians and administrators who manage these subscriptions. For the student or faculty member, it creates a situation where a single journal article can be available for download across multiple platforms, all of which are active library subscriptions. For the librarians and administrators, in addition to having to pay twice for certain content, content duplication can skew usage and referral data. This in turn can have significant ramifications for library decision-making around subscription renewal, discovery layers, and subject databases.

No published studies have attempted to comprehensively measure, across multiple institutions, whether or not this duplication is negatively impacting referrals and download rates within different e-journal platforms. As a result, there is a gap in the existing research around attempts to understand the role of discovery layers and library subject databases in the work- flows of students and faculty. In fall 2021, a preliminary analysis of this type of usage data was conducted at the University of Dayton, and the findings strongly suggested that the collective EBSCO full-text holdings (and the EBSCO search interfaces) are serving as the primary platform for users to search for, locate, and retrieve relevant full-text articles.[4] This initial study was published in the Journal of Electronic Resources Librarianship in July 2022 (https://doi.org/10.1080/1941126X.2022.2064105). The logical next step was to explore the same type of usage data at other universities to see if the pattern is generalizable. Accordingly, the aim of this study is to answer the following research question: To what extent do end users at research universities in Ohio use EBSCO's search platforms to directly retrieve EBSCO's full-text holdings of e-journals from major publishers, even when the same content is retrievable from additional platforms external to EBSCO? To answer this question, usage across seven universities–Case Western Reserve University, Cleveland State University, Kent State University, Ohio University, University of Akron, University of Dayton, and Wright State University–were analyzed collectively.[5]

## Literature review

Over the past decade, numerous voices within the e-resources community have sought to better understand how discovery layers, subject databases, and publisher platforms fit into the workflows of students and faculty when they seek to access full-text e-journal articles. Several prominent voices in the field (Cummings, 2021; Dempsey, 2020; Evans & Schonfeld, 2020; Hayman, 2017; Lean Library, 2021; Nicholas et al., 2017; Robinson et al., 2013) have argued that discovery layers and library-managed subject databases are declining sharply in importance, constituting a smaller share of overall usage than Google Scholar and commercial search engines. Many of these studies rely on web referral data to prove their point, and this referral data (so far as this author can tell) is supplied by the publishers, often anonymously, with little transparency on which analytics plat- form was used (Cummings, 2021, pp. 72–73; Dempsey, 2020 p. 12; Evans & Schonfeld, 2020, p. 15; Hayman, 2017; Lean Library, 2021, p. 28; Robinson et al., 2013, p. 198).

Separately, several studies have reported on the impact of introducing a discovery layer on e- journal usage within publisher platforms (Calvert, 2015; Evelhoch, 2016; Ngo et al., 2019). These authors assert that discovery layers did have an impact when introduced at their university, although the results are mixed in terms of how the implementation affected usage. Central Washington University implemented Primo and saw decline in journal article requests (Evelhoch p. 198), Western Carolina University implemented EDS and saw a sharp increase in e-journal requests (Calvert p.89), and UC Berkeley also implemented EDS and found more publisher platforms with decreases in the number of article requests than those with increases (Ngo et al., 2019, p. 227).

As was the case with this author's previous study on usage at the University of Dayton, the goal of this study is to address these two trends found in the literature: the trend found in several publications reporting low referral rates from library-managed databases and the separate trend found in other publications reporting a noticeable impact of discovery layers on e-journal usage. Unlike the previous study solely focusing on UD, this study seeks to address this gap by assessing usage at multiple research universities across Ohio. As mentioned in the literature review of this earlier study, Ngo et al. (2019) concludes their study on the impact of EDS at UC Berkeley by stating "full-text availability in the EDS interface may in fact negatively correlate with usage reported by a publisher or platform" and calls for future studies (p. 236). While this study does not seek to show a statistical correlation, it nevertheless seeks to study this conclusion by comparing the extent of usage within EBSCO's full-text holdings to the extent of usage of these same holdings outside of EBSCO.

## Materials and methods

This study relies on secondary data to measure full-text retrievals of e-journals by faculty, staff, and students at public and private research universities in Ohio. To a far lesser extent, it uses secondary data from these same universities to measure the number of searches, through the "platform searches" metric, and nearly all forms of interaction with the e-journal content through the "item investigations" metric. The platform search metric measures all search clicks within a platform, while the item investigation metric includes any form of interaction with item records. This includes full-text views, link outs to resolvers, abstract views, and other forms of interaction. It does not include searches, but it does include requests (see **Figure 2** below).

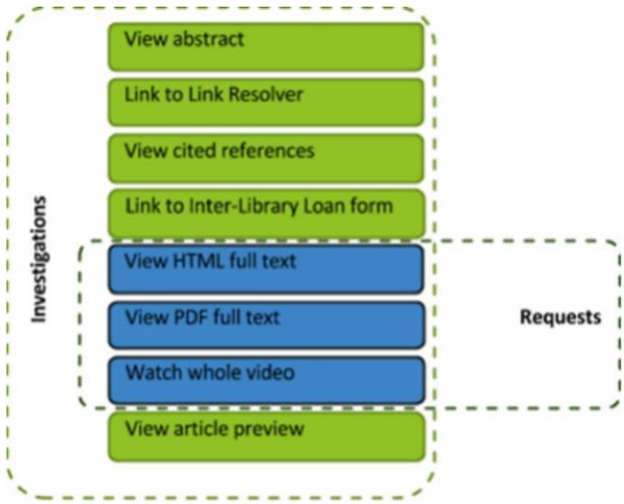



Figure 2. Breakdown of COUNTER 5 metrics "investigations" vs. "requests" (source: project COUNTER).

For the purpose of this study, "research university" is defined according to two Carnegie Classification of Institutions of Higher Education categories: "Doctoral Universities—Very High Research Activity (R1)" and "Doctoral Universities—High Research Activity (R2)" categories.[6] In Ohio there are five R1 institutions (Case Western Reserve University, Kent State University, Ohio State University, Ohio University, and the University of Cincinnati) and eight R2 institutions (Bowling Green State University, Cleveland State University, Miami University, University of Akron, University of Dayton, University of Toledo, and Wright State University). This study focuses on universities that use EBSCO Discovery Services (EDS) as a discovery layer, provide access to a package of EBSCOhost subject databases, and provide a common set of e-journal subscriptions provided through membership in the OhioLINK consortium (as provided by the OhioLINK EJC e-journal subscriptions negotiated by the consortium on behalf of members). As a result, to qualify for participation, any of the R1 or R2 universities listed earlier in this paragraph had to belong to the OhioLINK consortium, have access to the OhioLINK expanded pack of EBSCO subject databases, have access to OhioLINK's Electronic Journal Center (EJC), and use EBSCO Discovery Layer (EDS) as their primary discovery layer. Out of the thirteen R1 and R2 universities in Ohio, nine of them met the criteria listed in the preceding bullet point: Case Western Reserve University, Cleveland State University, Kent State University, Ohio University, Miami University, University of Akron, University of Dayton, University of Toledo, and Wright State University. In fall 2022, representatives from all nine university libraries were invited to have their university participate via email. In all cases, this involved contacting the e-resources or acquisitions librarian at each institution. Included in the email were two attachments, one detailing what participation would involve in terms of pulling the COUNTER 5 data and a time commitment, and a second document detailing the benefits of participating (see Appendix B for screenshots of these documents).[7] The collegial environment fostered by our joint participation in OhioLINK allowed for a degree of familiarity that might not have been present had this study looked beyond Ohio. Out of the eight universities contacted, six agreed to participate (Case Western Reserve University, Cleveland State University, Kent State University, Ohio University, University of Akron, and Wright State University). Along with the University of Dayton, this created a total of seven participating universities.

Each participating university was required to pull COUNTER 5 data reports to measure usage across five specific publishers: Taylor & Francis, SAGE, Oxford, Wiley, and Springer. Other major publishers, such as Elsevier, were excluded from this study because they do not provide full-text of their content outside their proprietary platform.

### *Data collection instruments*

Participating institutions pulled the COUNTER 5 "Journal Requests, Excluding OA_Gold (TR_J1)" (JR1) report for the following platforms: EBSCO (to measure EDS and all EBSCO subject databases), Taylor & Francis Online, SAGE Journals, Oxford Academic, Wiley Online Library, SpringerLink, JSTOR, and OhioLINK's EJC. Participating institutions also provided COUNTER 5's "Platform Master Report" (PMR) for these same platforms. These additional reports were pro- vided to gain further insight into end user behavior. Platform Master Reports include metrics for "platform searches" and "item investigations," in addition to item request metrics, and they also include any open access usage.

To ensure that the data represented usage over time, participants pulled these reports for calendar years (CY) 2019, CY2020, CY2021, and CY2022. For EBSCO, COUNTER 5 data for CY2019 expired in EBSCOadmin toward the end of 2022, so the data for CYs 19–21 was requested in fall 2022. This initial request was followed by an additional data collection for CY22 in spring 2023. Participants received detailed instructions (a step-by-step PDF with screenshots - see **Appendix B**) showing how to retrieve each COUNTER report. There were also instructions on how to send the data as a set of CSV files. No further time commitment was required by participating institutions once this data was provided.

### *Data analysis techniques*

Once the data was received from all participating institutions, it was cleaned and prepared for analysis using Excel. This involved creating two separate groups of spreadsheets, one for the PMR data and one for the JR1 data. For the PMR data, all data from each of the seven

universities, across all years, were combined within a single spreadsheet. For the JR1 data, data for each year was combined within individual spreadsheets because there were simply too many rows of data to allow for all years to be combined. Each year averaged well over 250,000 rows of data, and in order to process this within Tableau, each year could only be processed with a single "Metric Type", either "total item requests" or "unique item requests."[8] As data was received from each university, it was copied and pasted to the relevant spreadsheet (this data constituted columns A - W) with labels added for the university (column X) and year for the PMR data (column J). A controlled vocabulary was used for each university to keep the institutions anonymous while analyzing and visualizing the data. For the JR1 spreadsheets, a controlled vocabulary was also used for the five publishers. Routledge was folded into Taylor & Francis because it is wholly owned by Taylor & Francis and all the content is provided on Taylor & Francis Online. The same is not true for Nature content in relation to Springer because Nature has its own platform separate from SpringerLink.

Crucially, the COUNTER data was left relatively unaltered from how it was received. This allowed for the inclusion of all available variables while the visualizations and analysis evolved within Tableau and Excel. For any visuals showing usage at the individual title level (**Figure 9**), the "Print_ISSN" from the COUNTER 5 data. Print ISSNs were far more consistent than the online ISSNs provided by the data. For rows of data without any Print ISSN available, an attempt was made to either include the relevant ISSN or, in cases where the publication did not have an ISSN, an alternate unique identifier was used (typically the "Publisher_ID" metric provided by COUNTER). Regional campuses of the universities were included when their usage data was included within the initial report for the main university. This occurred with two platforms, JSTOR and the EJC. Also, for EBSCO, JSTOR, and the EJC, the usage data was filtered down to include only the e-journal content from the five publishers. Same as the 2022 study on the University of Dayton, only titles from the five publishers were included within the analysis. This especially applied in the case of the EJC, EBSCO, and JSTOR data, but it also applied to titles found within several additional platforms. For each platform, the JR_1 data was formatted as a table so that the publisher column could be filtered down to only include the relevant publishers. To keep things organized, a controlled vocabulary was used for the five publishers. This allowed for combining multiple spreadsheets–for multiple years, for example. In addition, similar to the previous study, Routledge was included under the Taylor and Francis data because Routledge content is available within Taylor & Francis Online, and a Taylor & Francis journal has few discernable differences from a Routledge journal within the platform. Routledge is fully owned by Taylor & Francis. For Nature, however, even though it is owned by Springer, its content is only available from Nature.com, not from SpringerLink. As a result, Nature journals are not included in this study.

Conversely, unlike the 2022 UD study, this study seeks to be far more conclusive (as opposed to more of an exploratory study). One of the key limitations of the earlier study was not only that it explored only one institution, but it was also limited in terms of being inconclusive on whether the trends found at the University of Dayton applied elsewhere. This study seeks to bridge that gap. Moreover, to offer more definitive conclusions, this study looks at multiple years.[9] Through the iterative process of creating the visualizations in Tableau, the analysis evolved and informed the final analysis and manuscript drafts. Separate analyses for each participating institution were created with a unique set of visuals for each calendar year.

In addition to being a visualization tool, Tableau was also used as a method to transpose and extract data. This involved creating visualizations within Tableau, extracting the data from these visualizations back as a CSV file and then reformatting the data to create new visualizations back in Tableau. The reason behind taking this route has to do with the ability for Tableau to quickly calculate totals for different variables. For example, to determine how many full-text article down- loads of Taylor & Francis journals occurred within EBSCO for each calendar month in 2022, Tableau can expedite the process for Taylor & Francis (and all other publishers included in this study) by layering them all into a single visualization. This visualization essentially compiles the relevant data, which can then be exported to excel, and then, for a multi-year analysis, the extracted data can be combined with the extracted data for 2019, 2020, and 2021. Similarly, this extracted data from Tableau served as the basis to create percentage visualizations featured later in this article. Excel formulas for percentage were used to determine what percentage of total downloads for a publisher (by month and by year) occurred within a specific platform. For example, these formulas allowed the author to see what percentage of Oxford e-journal article downloads occurred in OhioLINK's EJC platform in October 2021.

## Profiles of participating universities

As mentioned in the previous section, the participating universities for this study were Case Western Reserve University (CWRU), Cleveland State University (CSU), Kent State University (KSU), Ohio University (OU), University of Akron (UA), the University of Dayton (UD), and Wright State University (WSU). Across these seven institutions, there are 153,000 individuals who are entitled to access the individual university libraries' collections. **Table 1** provides a full breakdown of each institution according to student, faculty, and staff populations. As can be seen, there are significant differences between the different institutions in terms of the number of faculty, students, and staff. The two private universities have some of the lowest enrollments, as is typically the norm, with state schools and land grant universities having far higher enrollments. The three institutions with the least number of combined faculty and staff are CSU, WSU, and UA. As will be discussed later in this study, training and education on how to use library resources might be a key driver of usage of e-resources. Lower library staffing levels and/or underfunded library instruction programs could have had a negative impact on usage for several of the platforms discussed in this study.

Table 1. Reflects 2022 numbers. Sourced from university websites.

| University | # of undergraduates | # of graduates | # of faculty | # of staff | Total # of users entitled to access library e-resources |
|---|---|---|---|---|---|
| CSU | 10,626 | 4,838 | 511 | 1,007 | 16,982 |
| CWRU | 5,792 | 6,277 | 3,657 | 3,144 | 18,870 |
| KSU | 27,768 | 5,441 | 2,427 | 3,488 | 39,124 |
| OU | 22,668 | 6,419 | 1,604 | 2,530 | 33,221 |
| UA | 10,383 | 4,608 | 1,092 | 1,559 | 17,642 |
| UD | 8,443 | 3,327 | 994 | 2,180 | 14,944 |
| WSU | 7,477 | 2,818 | 1,062 | 908 | 12,265 |
| Total across all: | | | | | 153,048 |

In addition to these variations between the different universities, there are several other characteristics worth identifying. The first has to do with authentication methods at each university. All universities except for CWRU use ezproxy as their mode of off-campus authentication for the plat- forms analyzed in this study. On campus, a majority of institutions allow for automatic authentication to e-resources *via* IP range authentication. This last point is important because a student or faculty member does not need to enter credentials to gain access to paywalled content. Case Western uses Open Athens access. Coincidentally, the one institution here that requires on-campus authentication, even though they have ezproxy, Ohio University, is also the university with the highest numbers of Federated Access configurations (apart from CWRU). In practical terms for the end user, this means that if a student or faculty member reaches a publisher platform such as SpringerLink through Google while on campus, they will be able to authenticate using their cam- pus-wide login by selecting an "Access *via* your institution" link. At all other institutions, apart from UA and OU, this was not configured at the time of the study, but because of IP range authentication, on-campus users had no need to authenticate at all, so there was less need to con- figure this Federated Access. For off-campus access, however, these five institutions likely lost traffic as a result because the end user might have clicked the "Access *via* your institution" link within one of the platforms and been unable to find their institution listed among the options. Note that Federated Access is a highly time-consuming process to configure for certain platforms and also requires the university to be part of the InCommon federation. It is of little surprise that the two institutions with the strongest research mission and largest endowments are also the universities with the most access options available to their end users. Moreover, despite OhioLINK being part of InCommon, so far as the authors of this study can tell, it has not turned on Shibboleth as a form of Federated Access for any of its platforms, despite the fact that it could. The only university from this study that does not belong to the InCommon federation is Wright State.

The second trend worth noting are the different graduate professional schools and colleges at each institution. CWRU, KSU, OU, and WSU have medical schools; CSU, CWRU, UA, and UD have law schools; CWRU has a school of Dental Medicine; CWRU, KSU, UA and WSU have nursing schools; and all institutions have business schools and teacher education programs. While it is not possible within the parameters of this study to determine exactly how these different graduate school portfolios impact usage of e-resources, it is important to keep in mind the variation between the different institutions. Each university is unique.

Lastly, there have been significant enrollment declines at some of these institutions, most not- ably the public universities. As several local news stories have noted, the COVID-19 pandemic greatly exacerbated an already existing issue for many of these institutions, and in Ohio this has particularly affected the University of Akron, Wright State University, and Cleveland State University. Most notably, the University of Akron saw its enrollment decline from more than 30,000 students in 2011, to fewer than 15,000 students in 2023 (Morris, 2023; Pignolet, 2022). For the two private universities included in this study, UD and CWRU, enrollment remained steady throughout the period. Within the scope of this study, it is not possible to show a direct correlation between these enrollment declines and e-resource usage trends, but it is important to keep these declines in mind when viewing the results from this study, identifying them as one of many variables that might be causing various trends.

By chance, there was a 100% participation rate among qualifying institutions in Northeast Ohio (four universities) and Southeast Ohio (one university). For the Northeast Ohio institutions, the data discussed in this study offers a fairly comprehensive picture of usage at universities within the Cleveland-Akron metropolitan area.

## Results

### *Overall usage trends at these universities (Platform Master Report data)*

While the primary focus of this study is on e-journal usage through the JR1 COUNTER 5 data, the Platform Master Report (PMR) COUNTER 5 data was also requested from all universities to provide a more comprehensive understanding of usage. This data provided an initial window through which to understand the broader trends at the different institutions. **Figure 3** visualizes this PMR data, and unlike subsequent visuals in this study using the JR1 data, it does include open access usage, some- thing which on average increases totals by 10%. It also includes all data "types", which includes (but is not limited to) e-journals, e-books, reports, newspapers, thesis/dissertations, and multimedia. **Figure 3**, showing four-year trends for item requests and item investigations, serves as a foundation for understanding and

thinking critically about the JR1 COUNTER 5 data visualizations, and two major trends are evident from this visual. First, nearly all platforms saw a decline in their usage over the timespan included in this study. The second has to do with the way EBSCO's investigations–and their relationship in terms of ratio to item requests–are different from nearly all other platforms. This visual indicates that a request in EBSCO subject databases and EDS means something very different than a request in SpringerLink or Wiley Online Library, something which other researchers have recently started identifying (Getsay & Chen-Gaffey, 2021). For all of the institutions, there are "big deal" subscriptions to Springer and Wiley, and many of the campuses have automatic on-cam- pus IP range authentication. The presence of these two variables–big deal subscriptions and blanket IP range access–results in the end users clicking on a result to simply see the abstract of an article and triggering one item request and one item investigation. Crucially, this occurs regardless of whether the intent of the end user (they could simply wish to view the abstract) because the full text automatically displays. For many results within EBSCO databases or EDS, this automatic item request is not triggered because the HTML full text is not typically available. Therefore, it is worth considering the implications of this trend within Wiley Online Library, SpringerLink and other plat- forms, to think about the impact it likely has on higher download counts. Conversely, in EBSCO it has the opposite effect, hence the wide difference between EBSCO's item request and item investigation totals. This consideration should be taken into account when reviewing item request metrics for any of the visualizations in this publication, as has been pointed out by researchers at Slippery Rock University of Pennsylvania (Getsay & Chen-Gaffey, 2021, p. 214).

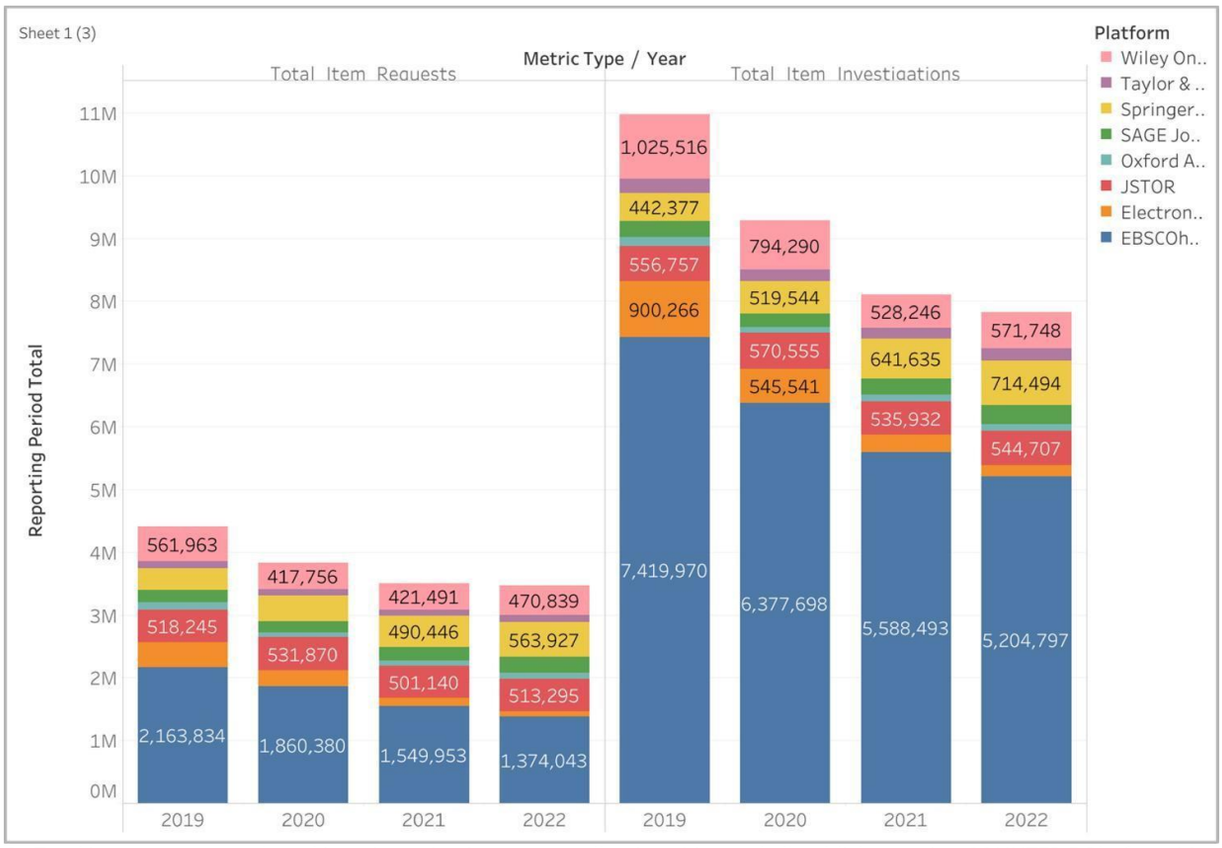



Figure 3. Total item requests and total item investigations across all universities and all publishers using COUNTER 5 Platform Master Reports. In the case of EBSCO there is a vast difference between the number of investigations and requests, whereas for other platforms, the ratio difference is far less significant.

If we take into account how investigations result in automatic full-text views for many major publisher platforms, while at the same time, these same forms of investigations/interactions do *not* result in automatic full-text view for a high percentage of EBSCO records, it's worth noting that investigations would allow for a more accurate understanding of the extent of usage within the different platforms. This study relies more heavily on item requests than item investigations, but investigations remain a highly-important future direction for this field of research.

### *Popular e-journal publishers within EBSCO's full-text holdings*

Above all, the JR1 reports allowed the author of this study to identify how e-journal item requests within EBSCO, JSTOR, and the EJC compared to requests of e-journal articles within the publisher platforms of the five publishers. Crucially, unlike the PMR data, the JR1 data allows for analysis at the individual title level, allowing the author to parse out individual titles across all platforms.

To begin the analysis of JR1 data, EBSCO's 2022 data was analyzed across all universities to determine the most popular publishers. In the 2022 study on UD's usage, the same five publishers were used as this 2024 study, and a first step for this study was to determine whether those same publishers were popular within EBSCO across all seven universities. **Figure 4** shows a breakdown of the most publishers across all seven universities in 2022. Taylor & Francis was by far the most popular publisher within EBSCO with 18.5% of overall usage. This finding was unsurprising based on the results of the initial study at UD.

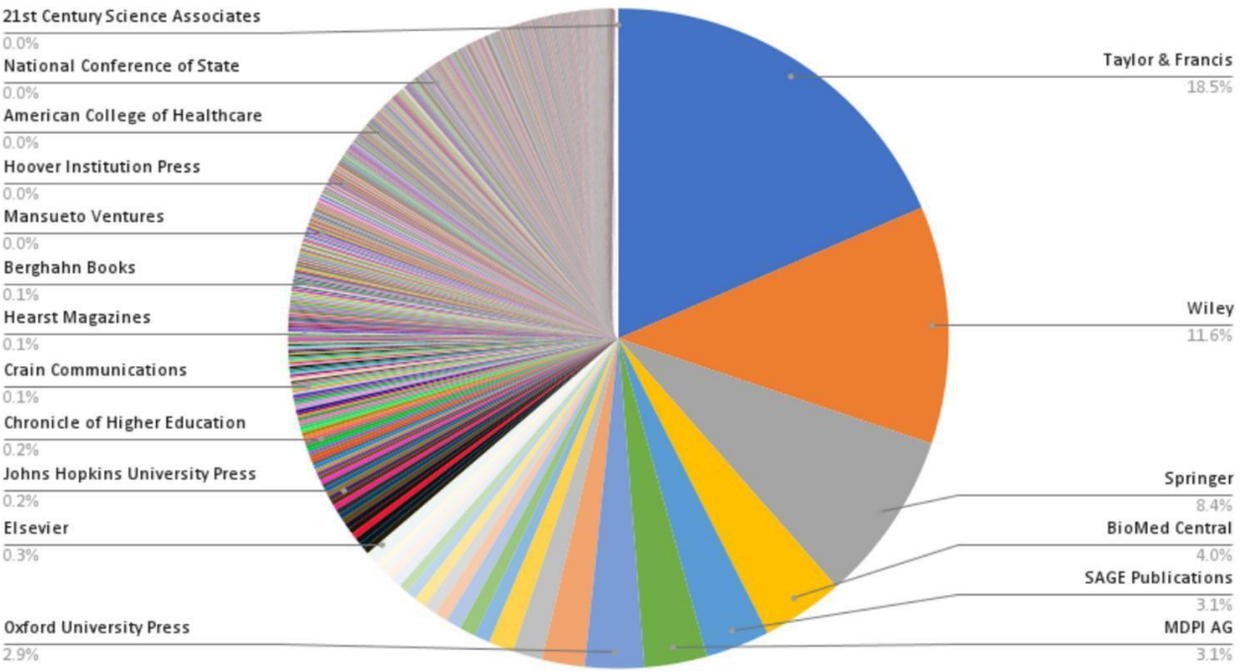
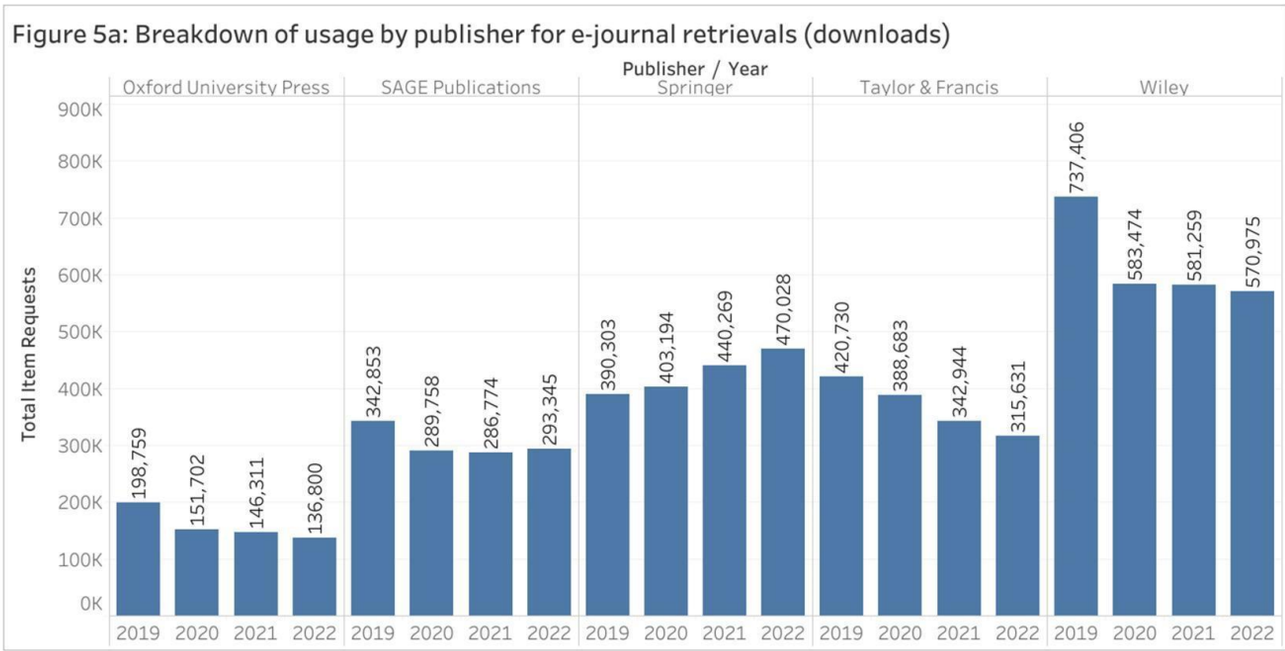



Figure 4. Total item requests within EBSCO (EDS and EBSCOhost subject databases) by different publishers in 2022, across all seven universities. Note, in addition to including the five publishers, the visual includes all other publishers with full-text hold- ings available within EBSCO databases.

### *Breakdown of usage by platform for the different publishers*

Next, broadening the analysis of JR1 data across all platforms, it was important to determine which of the eight platforms were the most popular. In 2022, across all platforms and universities Wiley and Springer were the two most popular e-journal publishers out of the five (**Figure 5a**). As can be seen in **Figure 5a**, this trend has remained fairly consistent over the past four years, although in 2019 Taylor & Francis was more popular than Springer. The breakdown within each publisher of where these downloads are occurring, in terms of platform, present a far more complicated picture (**Figure 5b**). In the most recent calendar year of this study, 2022, the percentage of overall downloads for these five publishers occurring within EBSCO varied: 24% for Wiley, 69% for Taylor & Francis, 21% for Springer, 13% for SAGE, and 25% for Oxford. When JSTOR and the EJC are combined with EBSCO, 32% of Wiley e-journal downloads occurred outside of Wiley Online Library; 75% of Taylor & Francis occurred outside of Taylor & Francis Online; 26% of Springer occurred outside of SpringerLink; 26% of SAGE occurred outside of SAGE Journals; and 47% of Oxford occurred outside of Oxford Academic. In other words, in terms of the platforms used (as seen in **Figure 5b**), a majority of the usage occurred within the publisher platforms for every publisher except Taylor & Francis.[10]

Therefore, in 2022, the combined percentage of downloads occurring in platforms unreachable *via* Google or other commercial search engines, namely the EJC and EBSCO, ranged between 17% and 69% (Wiley 26%; Taylor & Francis 69%; Springer 22%; SAGE 17%; Oxford 26%).[11] Worth noting, the percentage occurring within JSTOR was really low for every publisher except Oxford University Press: Taylor and Francis 6%; Wiley 5%; Oxford 21%; SAGE 8%; and Springer 1%. For the EJC, usage was also really low, and on top of this it decreased dramatically every year. The impact of the pandemic shutdowns on changes in usage between 2019 and 2020 was likely significant, contributing to sharp declines. We also see a decline in usage between 2019 and 2022 for four out of the five publishers, with only Springer seeing an increase every year. EBSCO usage declined every year too, although not nearly as much as the EJC.



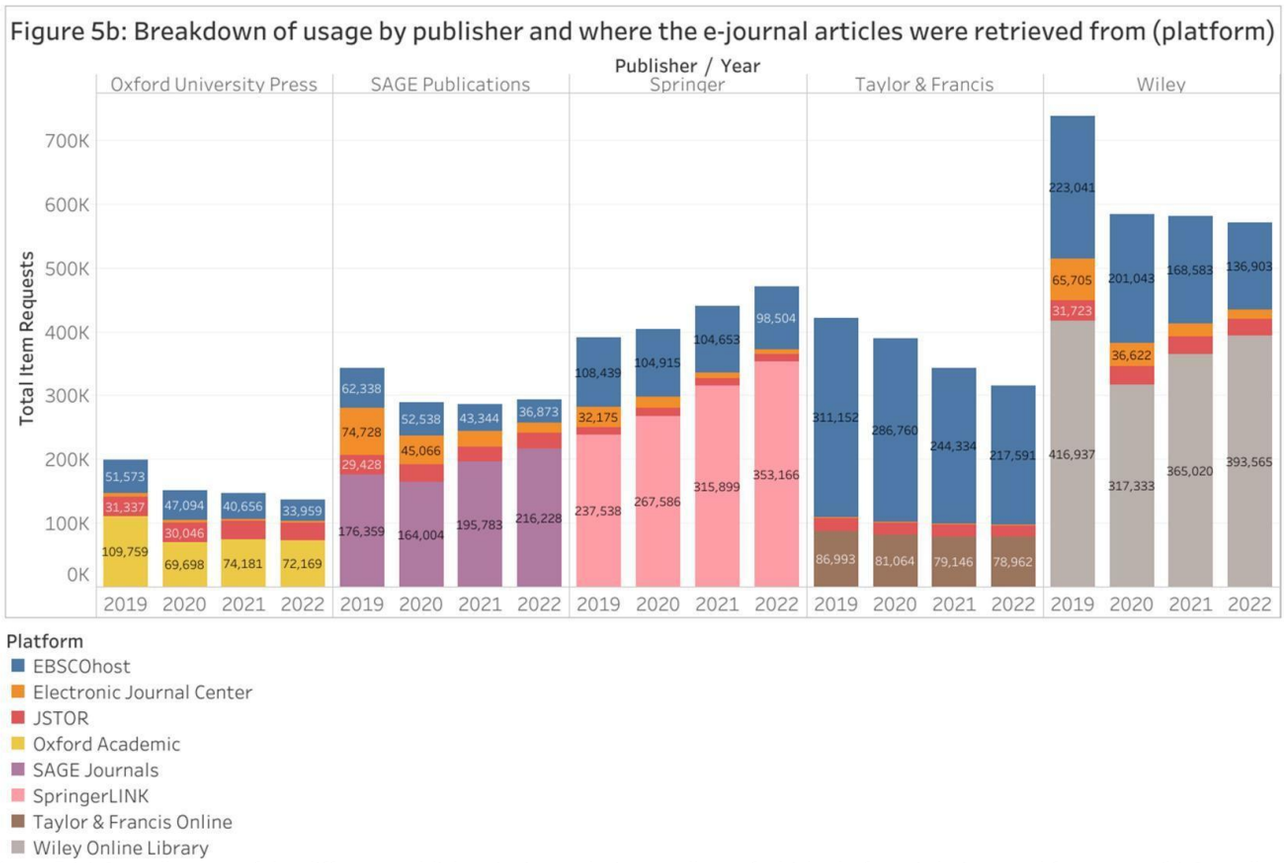
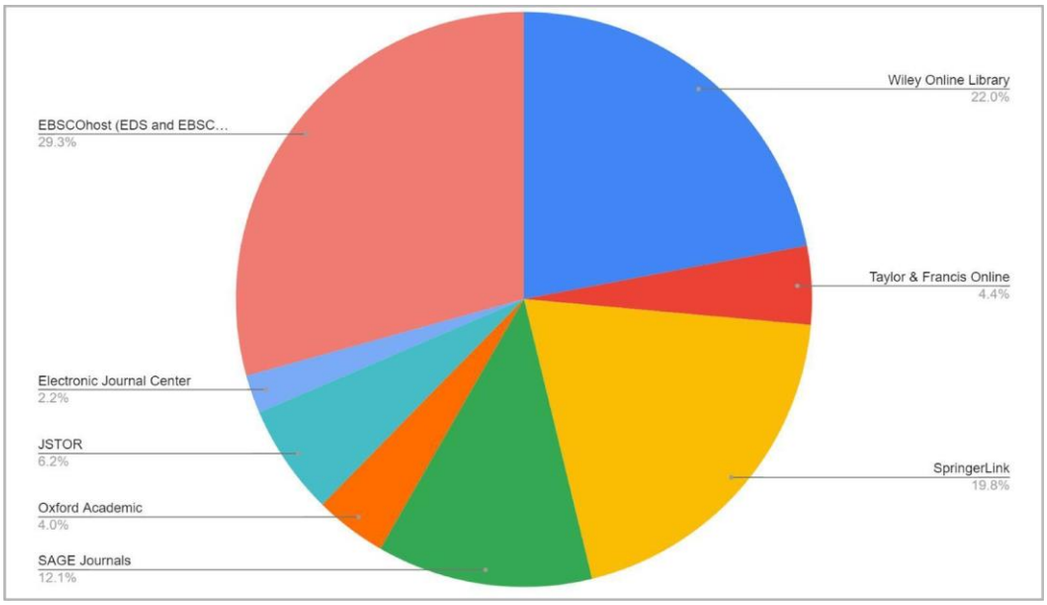




Figure 5 (a) Overall usage of the different publishers' e-journals, irregardless of which platform is being used for access. (b) Overall usage of the different publishers' e-journals, with a breakdown by different platforms being used.

### *Percentage of retrievals potentially reached via google*

In 2022, 1,787,347 downloads occurred for these five publishers across all participating universities. As can be seen by totaling the different percentages in **Figure 6**, sixty two percent of these downloads occurred across the respective five publisher platforms. Thirty two percent occurred within EBSCO and the EJC, and 6% occurred within JSTOR. If we assume that 10% of the 1,225,017 downloads that occurred within platforms reachable from Google and other public-facing search engines (JSTOR and the five publisher platforms) were referrals from EBSCO linkouts, then we would claim 122,501 of these downloads occurred as a result of EBSCO and the EJC (alongside the 562,330 downloads occurring within EBSCO and the EJC). This totals 684,831 downloads (or 38% of the 1,787,347 downloads that occurred in 2022).

Figure 6. Breakdown of usage by different platform for 2022.




In terms of how this usage breaks down by institution, six out of the seven institutions saw a similar extent of usage occur within EBSCO and the EJC. The major outlier to this trend was Case Western Reserve University, which relied on EBSCO far less than the other institutions, as can be seen from **Figure 7**.

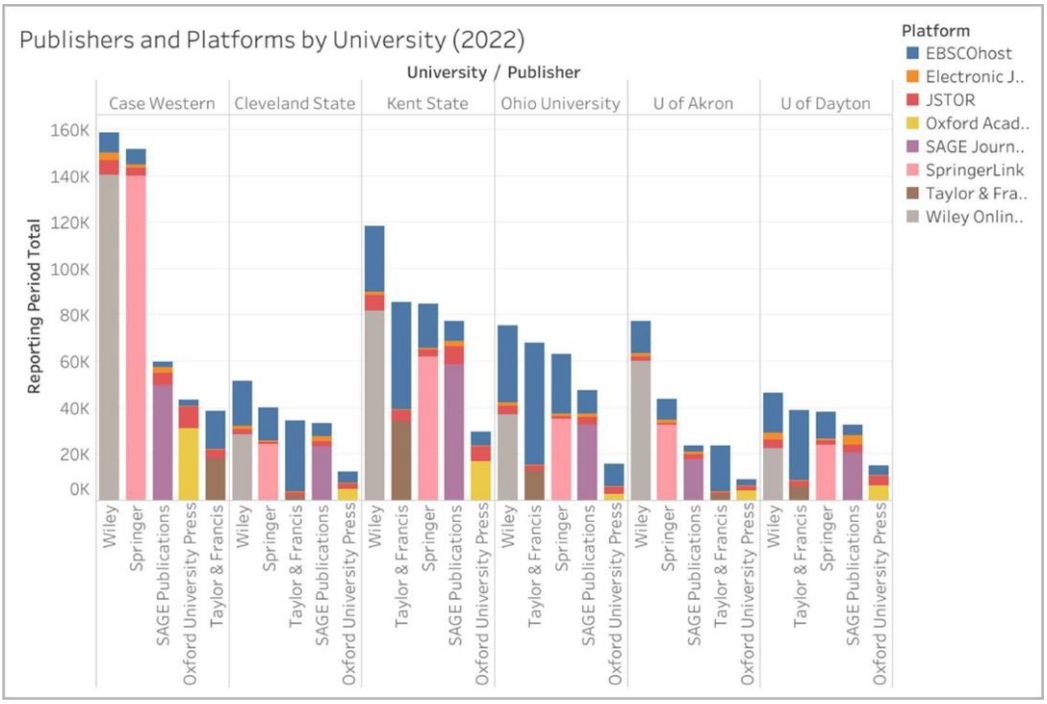



Figure 7. Total item requests of e-journals by university. Bars for each of the five publishers includes a breakdown by color to depict where the articles were retrieved from (platform). Note: Wright State University is excluded from this visual because the institution did not grant permission for their data to be de-anonymized at the institution level.

## *Usage by month*

Relying on the monthly data from the JR1 reports, it was possible to see how usage by platform (as a percentage of overall usage for different publishers) fluctuated within a twelve-month period. As can be seen from **Figure 8**, the spikes for the publisher platforms occurred consistently in August of each year, and the spikes for EBSCO platforms occurred in March, April, October, and November. Moreover, several publisher platforms saw steep drops in usage in March and April and 2020, while the same did not occur for EBSCO. This trend can be clearly seen in the case of Wiley, Oxford, and Taylor & Francis (**Figure 8**). An additional pattern evident from these visuals is how these patterns are highly consistent year over year.

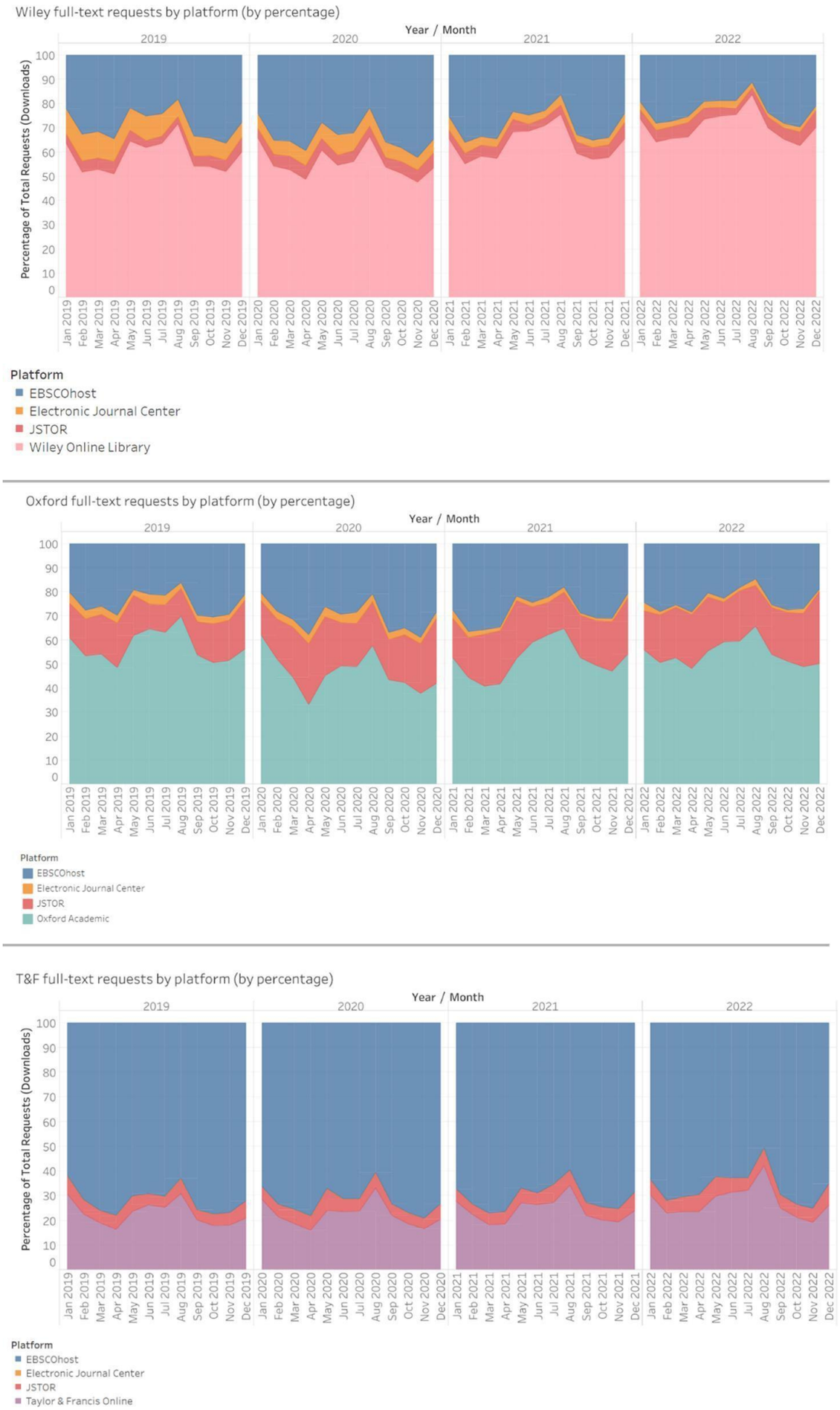



Figure 8. Wiley, Oxford, and Taylor & Francis (T&F) e-journal full-text requests (total item requests) showing different the percentages by platform used. Note the really high EBSCO usage for T&F, as well as the spikes in EBSCO usage during peak curriculum months.

### *Visualization and analysis at the journal title level*

**Figure 9**, shows the breakdown of most popular e-journals across the five publishers and all participating universities in 2022. Different colors correspond to the different platforms used to access the full-text. The blue sections represent EBSCO platforms, while red represents JSTOR. Barely visible is the orange which represents the OhioLINK EJC platform. As can be seen in the visual, multiple titles saw 50% or more of their usage occur *outside* of their publisher platforms.

Not visualized here are the most heavily used journals within EBSCO outside of these five publishers. While these titles technically fall outside the scope of this study, they are worth mentioning as a tangent. Across all these eight platforms, the analysis found the most popular title across all universities and publishers was PLoS One, an open access title, with 22,881 full text retrievals in CY 2022. Similarly, the third most popular is an MDPI OA title, the International Journal of Environmental Research and Public Health, with 14,099 full-text retrievals. Both of these publications are entirely open access and indexed by commercial search engines and Google Scholar, but this form of access is not represented in these totals. Instead, this usage of OA represents usage of *EBSCO's full text holdings*, meaning that users still had to authenticate to gain access. In other words, this occurs despite the fact that the EBSCO full-text access to these titles is not the fastest way for the end user to retrieve these articles: the user has to first access the EBSCO database or EDS, and if they are off campus they would not need to authenticate. Both titles also saw more downloads than the most popular titles visualized above. This trend occurred across all four years of this study.

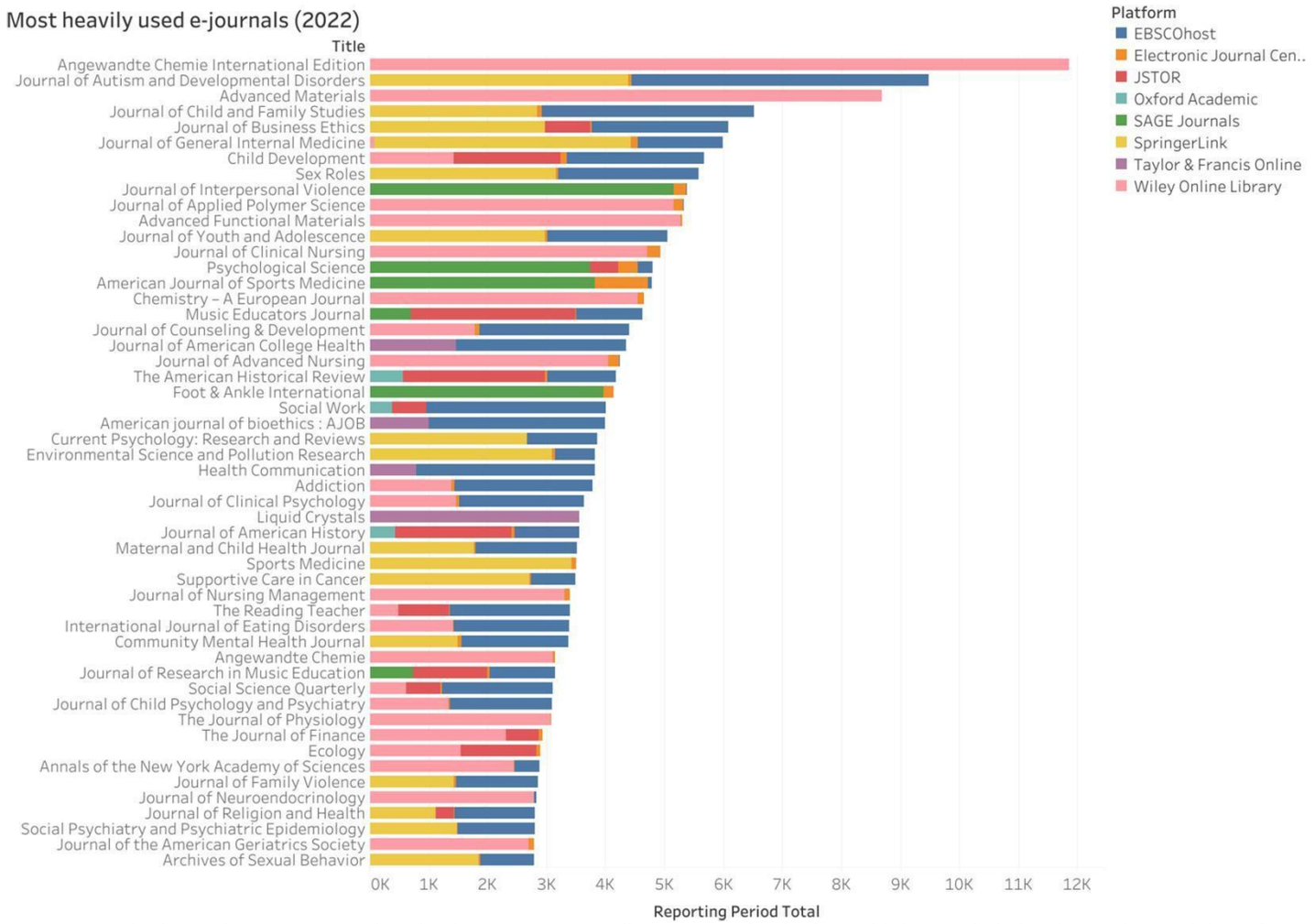



Figure 9. Top titles by total item requests (downloads) in 2022 across all universities for the five publishers included in this study. Breakdown for each title shows different colors for the different platforms.

## Discussion & conclusions

This study was conducted to determine the extent of usage of EBSCO's full-text holdings of major publishers, as a follow up on a previous study relying solely on University of Dayton data. The earlier study found a majority of e-journal usage occurring within EBSCO (Robins, 2022). While the results of this study do not show this same degree of EBSCO usage across all institutions, the data still reflects a significant amount of usage occurring outside of the publisher plat- forms, with many publishers seeing over 40% of their downloads occurring within EBSCO. This finding in particular, combined with the clear connection between the busy academic calendar months and increased usage of EBSCO, emphasizes a different understanding of what leading voices in the field have claimed about the role of the library website, the discovery layer, and library-provided subject databases within the workflows of faculty, staff, and students.

### *Going beyond the publisher platforms to determine usage*

These findings demonstrate the need to look at usage across multiple platforms when trying to understand e-resource usage at any academic library, and especially at research universities. It seems entirely reasonable to assume the vendor representatives from the five publishers included in this study would have an inaccurate response to the question, "what is the most popular [Taylor & Francis, Oxford, Springer, etc.] journal being used at the University of [blank]?" In all likelihood, the vendor rep would respond based solely on their respective platform's COUNTER 5 data–defining the most popular journal within their respective *publisher platform* and, therefore, showing only one part of overall usage. There is no way for these publishers to obtain EJC data to answer these questions, but considering that they license their content to platforms like EBSCO and JSTOR, there probably should be a way for publishers to disclose the usage data when working with client universities and colleges. It might also be the case that they already receive this data from EBSCO and JSTOR, in which case more transparency would be beneficial to libraries. Related to this, it is perhaps a little misleading for Wiley and other publishers to not acknowledge the likelihood that the duplicate presence of their full-text holdings within EBSCO might have an impact on referral data when they send it to authors (Cummings, 2021, pp. 72–73; Evans & Schonfeld, 2020, p. 15). As can be seen from multiple findings in this study, individual titles are rarely accessed from a single platform when they are available in full text from multiple platforms. E-resources, unlike print resources, present a maze of multiple platforms to assess usage, and the LIS field should work with publishers to make it easier to understand usage at the publisher or title level across platforms.

LibInsight from Springshare is one platform which does offer this form of dynamic analysis in a fairly straightforward manner. And one advantage this platform provides is in allowing for analysis at the journal title level for item investigations, in addition to item requests. This allows for further insight on where the usage is occurring in terms of platform. This study did not pull this type of data for all participants because of the limited availability of item investigation data at the title level. However, because the University of Dayton uses LibInsight for its internal assessments of e-resource usage, investigation data was on hand for an analysis at that institution. As can be seen in **Figure 10**, this allowed for further insight on how external platforms relate to e-journals and conference proceedings from publishers whose full text is almost exclusively available within their publisher platforms, namely IEEE and Elsevier. As can be seen in the figure, a significant amount of usage for Elsevier occurs outside of its publisher platform ScienceDirect, within EBSCO. For IEEE, conversely, nearly all usage occurs within their publisher platform IEEE Xplore. Further studies on what this fully means is needed, but for the purposes of this study, the Elsevier data does strongly suggest that the end users at UD rely on EDS and subject data- bases to view Elsevier article abstracts, link out to ScienceDirect, and search for Elsevier journals. For 2022, the year of the visual, roughly 27% of investigations of Elsevier e-journals occurred within EBSCO.[12] Again, this is at odds with the 10% reported by various studies. For the additional publishers included in the visual, the percentages are far higher and reinforce this point further.

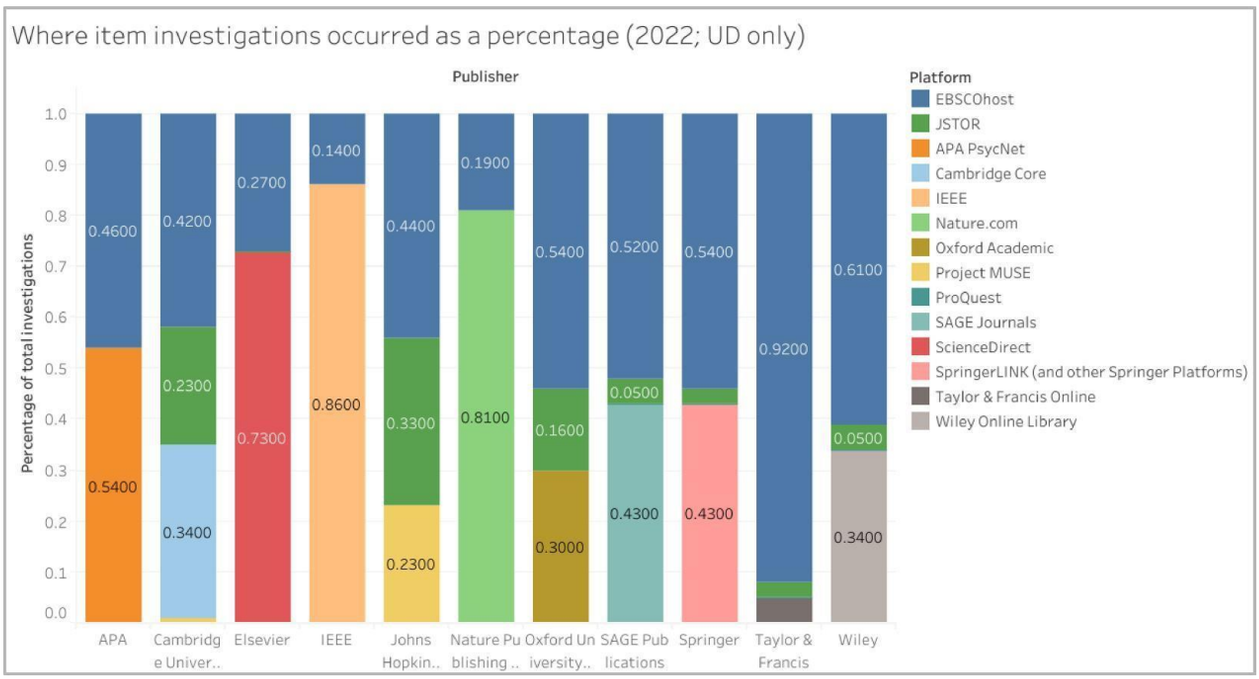



Figure 10. Breakdown of percentage of item investigations by different platform (UD only). Includes additional publishers and platforms, most notably Elsevier/ScienceDirect and IEEE.

On top of this, it is important to keep in mind how the platform can influence the download rate. For example, in Wiley, ScienceDirect, and other prominent publisher platforms, the presence of two variables–campus-wide IP range access and big deal subscriptions–guarantees that an abstract view by an end user will almost always guarantee a full text retrieval is also triggered, even if the user only intended to view the abstract. The same is not the case with EBSCO, as has been identified by other researchers at Slippery Rock University of Pennsylvania (Getsay & Chen-Gaffey, 2021, p. 214).

### *End-user behavior*

The results suggest end users are far savvier than the literature seems to suggest. EBSCO and OhioLINK EJC records cannot be reached *via* google or other search engines; therefore, for any period where there is a significant percentage of overall usage occurring within those

platforms, the implications are that the end user needs to learn how to use the library's website to reach the content. To varying degrees, EBSCO and the EJC were used significantly at every university except for Case Western Reserve University (CWRU) over the past four years. This likely has to do with CWRU being the only university featured in this study to use OpenAthens as their authentication method. OpenAthens, unlike EZproxy, significantly removes the need for the off- campus end user to start with a proxied link, typically provided from the library website or A-Z database list. Accordingly, the findings seem to suggest a trend: if you give users the option to reach everything *via* Google for on-campus and off-campus access, as is the case with CWRU (and their additional configuration of Open Athens), then they do follow the pattern reported by the leading voices in the field (of only using EBSCO databases and EDS 10% of time). However, if you have the slightly different situation where users do have automatic IP authentication to e-resources while on-campus, but are required to authenticate using library-provided EZproxy links when off-campus, then EBSCO and EDS are used far more often–most likely because they provide a reliable space to access full text. For these institutions where EZproxied links are the dominant mode of access (at least for off-campus users), this in turn means the library requires users to learn that they must go to the library website and authenticate to gain access to library resources. The data from this study for these universities seems to suggest that end users do seem to figure out that the discovery layer and EBSCO subject databases are a reliable way to locate full text. This, of course, occurs as a result of direct training *via* librarian-led instruction, library service desk interactions, word of mouth, patrons learning on their own through online tutorials such as LibGuides, or a combination of all these factors. No better example of this exists than OU where in 2022 four out of five of the publishers saw over 40% of e-journal downloads occurring outside of the publisher platforms and JSTOR, in spaces that can- not be reached *via* commercial search engines such as Google (**Figure 7**).

Additionally, the findings from this study present ample reason for the LIS and publisher com- munities to shift their focus away from a preoccupation with where users "start" their research to one concerning the extent of overall usage. Reframing the conversation in this way–at least for the research university user groups featured in this study–shows a far different picture of end-user behavior among faculty, staff, and students in their usage of discovery layers and subject databases. For example, between 2019–2021 4 out of every 10 Wiley journal article downloads were occurring outside of Wiley Online Library, and 3 out of every 10 were occurring within EBSCO. For Taylor & Francis the trends are even more staggering and seem to highly suggest that library-managed search interfaces (which cannot be reached *via* commercial search engines) constitute the main place where end users access this content considering that for all four calendar years 7 out of every 10 Taylor & Francis journal article downloads occurred outside of Taylor & Francis Online, and during the peak academic months (March, April, October, and November), this percentage would creep even higher, in some years reaching nearly 80%. This stands at odds with what many leading voices in the field of e-resources have stated over the past decade. If we add the claim by many studies that 10% of referrals to these publisher platforms comes from the discovery layer, then it is not hard to see how the referral data studies are at odds with the usage patterns shown here.

As identified in this author's previous study, the higher percentage of downloads within EBSCO of these five publishers' e-journals likely decreases the number of linkouts from EBSCO to the publisher platforms (Robins, 2022). This pattern of end user behavior likely misleadingly influences researchers who have noted a 10% referral rate average from EBSCO in their studies to conclude that EBSCO does not play a significant role (Cummings, 2021, pp. 72–73). Again, the data from this study shows a far higher percentage of *downloads* (not referrals) occurring within EBSCO, and this does not account for the number of linkouts (or even abstract views) also occurring from EBSCO to these publisher platforms. Speculating on the combined number of linkouts and downloads occurring within the broader EBSCO ecosystem of EDS and the subject databases points to a significant percentage of usage relying on (if not starting with) EBSCO. The magnitude of not accounting for these other forms of end user interaction (i.e., only considering the referral data) could be significant in the long term for the LIS profession. For example, OhioLINK's 2020 White Paper claims–based on referral data–that a rather insignificant amount of Wiley Online Library sessions across all OhioLINK institutions began with EBSCO (p. 15). This visualization by OhioLINK was then referenced by Sage's Librarian Futures Report in 2021, where they claimed that "only 6% of patron discovery began at the library" across "discovery workflows for patrons in the OhioLINK network of 117 libraries:" (p. 28). While this reference by Sage seems completely off-base to begin with, considering that the vagueness of "at discovery workflows for patrons" (Sage makes no mention of Wiley Online Library and imply that this % is across all platforms), it is equally damaging to the library community in terms of the false impression it provides to readers, and it is worth speculating whether OhioLINK and Sage would have changed their comments if they had seen the data presented in this study when they were authoring their con- tent. This article encourages librarians and administrators to look beyond referral data.

### *What this could mean for big deal subscriptions*

The results, especially for Taylor and Francis, raise significant questions about end-user behavior in the absence of a big deal package. Out of these publishers, Taylor and Francis is the only publisher that falls outside of OhioLINK's major subscriptions. The EJC does not provide access to any current Taylor and Francis journals. Moreover, nearly all of the participating universities in this study do not have extensive local subscriptions to Taylor and Francis titles. This absence seems to have led users to access Taylor and Francis content (often embargoed) through EBSCO's full-text holdings instead.[13] Once again, this suggests that, when necessary, the end user seems fully capable of finding a way to the full-text content, which in this case is being provided by EBSCO.[14]

One can argue that the journal articles being downloaded from this publisher are only being paid for once, and this is typically through the EBSCO subscription. Despite this sole location for access, usage rates are still strong relative to the other publishers from this study. In an era of declining budgets and exorbitant year-over-year increases from major publishers, it might be worthwhile for academic libraries to see these EBSCO full-text holdings as an alternative to big deal e-journal subscriptions, despite the fact that many of these Taylor & Francis titles are embargoed within EBSCO, thereby restricting the ability of the researchers and students to access the latest content. There are, of course, downsides to this approach. EBSCO only leases these Taylor & Francis titles, and the holdings could change at any time, creating a lack of stability for the end- user. Still, this prompts the question: if OhioLINK dropped its big deal subscriptions to Wiley or Springer, would users

turn to EBSCO's holdings for these publishers instead? Could this in turn save OhioLINK members a significant amount of money while still meeting their users' needs?

### *Connection to curriculum*

As indicated by **Figure 8**, findings show a strong surge of usage for EBSCO and the EJC during peak academic months. Conversely, the summer months are when the publisher platforms see the greatest proportion of overall usage. This strongly suggests that students rely heavily on the discovery layer and subject databases. Several articles from the past decade focus on "early career researchers" and their preference for Google and Google Scholar (Hayman, 2017 and Nicholas, D., et. al. 2017). Considering the slant toward undergraduates among the universities featured in this study, it could still be possible for both things to be true: undergraduates appreciate having the guardrails, so to speak, provided by the discovery layer and subject databases, while early career researchers are at a stage where they've moved beyond this phase of their trajectory. This logic supposes that on their way to becoming early- career researchers (or simply on their way to graduation), the preference for and reliance on EBSCO seems strong, while those who devote their careers toward research no longer rely on EBSCO (in the same way they used to) once they are employed researchers. Further study in this area is needed.

In addition, these findings reinforce the likely impact library instruction sessions (and likely LibGuides, handouts, and video tutorials) have on usage of these platforms. This also deserves further research to determine, but it does not seem far-fetched to suggest that more library instruction on databases that are not reachable *via* Google results in more usage of these plat- forms. In other words, these findings provide ample reason to not discount the value of instruction programs to train users, especially students, on how to use these platforms. Likewise, administrators and leadership within libraries should not discount the importance of EBSCO's full-text holdings, especially in the workflows of students.

### *Limitations and future directions - Results at the consortial-level vs. a subset of universities*

A key paradox within the LIS field is arguably our aspirational goals to test things at the local level for UX design and other user needs, whether it's for the layout of the catalog, discovery layer, or library website (Adams, 2023; Sonsteby & DeJonghe, 2013), while at the same time, when it comes to usage trends, we often default to data reported at a statewide or national level to understand how usage is universal. Key examples include the Sage Librarian Futures Report and the OhioLINK white paper (Evans & Schonfeld, 2020, p. 15; Lean Library, 2021, p. 28). This study attempts to chart more of a middle ground, looking at a specific type of university, R1 and R2 universities. But even within this subset of universities within OhioLINK, there is really wide variation between the different institutions when it comes to how e-journals are accessed. CWRU fits the narratives put forward by leading voices in the field, while all other universities, apart from maybe the University of Akron, do not fit this narrative. Excluded from this study are the two largest universities in the state–Ohio State University and the University of Cincinnati. Could it be that their usage trends would tilt the narrative in a different direction?

Another key set of questions to ask here concern other subsets of OhioLINK institutions. How are the needs of small liberal arts universities different? Or, are there wide variations between private and public universities? Does the difference between public and private have to do with students' social and economic backgrounds? Or does the variation hinge on perceptions around the "legitimacy" or "quality" of a more elite private university library versus the all-too-often under- funded state school? Does this reflect tremendous inequities between different institutions within OhioLINK in terms of staffing and resources for collections maintenance and discovery layer con- figuration? And lastly, how do these patterns compare to different consortiums or groups of R1 and R2 universities?

What's apparent in asking all of these questions is the need for *more* research on usage pat- terns and more *localized* research, moving away from an over-reliance on surveys and referral data supplied by publishers. If we fail to conduct additional studies, we risk a profound lack of understanding when it comes to the needs of our users. The author of this study encourages all institutions to comprehensively analyze their institutions' usage data at the local level, while also acknowledging the significant amount of time required for such an analysis. There is a strategic disadvantage many libraries face when it comes to the staffing and time required to reach conclusions on e-resources, while at the same time there is a strategic disadvantage when libraries fail to comprehensively analyze how their e-resources are being used. This study seeks to provide a comprehensive approach for any library to follow to uncover usage patterns at their institution.

## Notes

1. E-journal titles within EBSCO frequently have full-text embargoes for 1-2 years.
2. Mentions of "full-text downloads" or "retrievals" occurring within EBSCO's full text holdings implies a direct download of PDFs or viewing of full text within a web browser. Crucially it implies that the user does leave EBSCO's interface to view the full text (i.e. they are not going to an external publisher platform via a link out to view the full text).
3. OhioLINK's Electronic Journal Center (EJC) is a database available to most OhioLINK consortium members and provides over 32 million full-text articles across 10,000 journals.
4. See the "Discussion & Conclusions" section of this article for a visualization of 2022 data from U of Dayton.
5. These seven universities total roughly 150,000 students, faculty, and staff.
6. Learn more about the classifications: https://carnegieclassifications.acenet.edu/classification_descriptions/basic.php
7. See Appendix B for these attachments and a sample email.
8. "total item requests" was chosen over "unique item requests" because the focus of this study is on the extent of usage. Moreover, any comparison with the data between the two metrics–during the analysis– did not show any difference between the two metrics, apart from unique item requests being lower than total item request; therefore, had unique item requests been used, the totals would only have been lower for each platform.
9. At the same time, the author of this study is aware of just how difficult it can be to measure all usage occurring at these institutions, not only because of usage occurring off-campus, undetected by IP addresses or proxies. The author of this study is aware of the myriad of ways a complex web of siloed administrative accounts and IP ranges can evolve within a university campus and thwart any attempt by the library to comprehensively track usage. Accordingly, this study–and the previous study–represents the usage associated with the primary administrative accounts for each of the university libraries

and for each of the included platforms. At the same time, for relevant platforms, regional campus data was included in this study.

10. In 2022, sixty eight percent of Wiley e-journal requests occurred within Wiley Online Library; 25% of Taylor & Francis occurred within Taylor & Francis Online; 74% of Springer occurred within SpringerLink; 74% of SAGE occurred within SAGE Journals; 53% of Oxford occurred within Oxford Academic.
11. See Table A1 for a full breakdown of these percentages.
12. See Table A2 for a full table detailing each platform for each publisher.
13. In addition, the nature of the Taylor & Francis content, leaning more towards the social sciences, could also be a factor.
14. There is no connection between the results from this study and deciding to publish in a Taylor & Francis journal. If anything, the results found here do not reflect well on Taylor and Francis business strategy up until now, considering that unlike the other four publishers, they have been far less successful in licensing big deal packages to OhioLINK and charging these OhioLINK institutions to manage multiple subscriptions to duplicate content.

## Acknowledgements

Special thanks to all librarians at participating universities who generously agreed to gather and share the data for this study. This includes: Melanie McGurr and Gregory Harris at the University of Akron; Holly Talbott at Kent State University; Seth Sisler at Ohio University; Hannah Pearson at Cleveland State University; Deberah England and Karen Wilhoit at Wright State University; and Anne Kumer and Anne Trenholme at Case Western Reserve University.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Data availability statement

All data discussed/analyzed in this study for University of Dayton and Case Western Reserve University is available in figshare at: https://doi.org/10.6084/m9.figshare.25571643. Questions about data for all other universities featured in this study can be directed to Simon Robins srobins1@udayton.edu and will be evaluated on a case-by-case basis.

## References

Adams, A. L. (2023). UX resources. Public Services Quarterly, 19(1), 38–45. https://doi.org/10.1080/15228959.2022.2155756

Bates, J., Best, P., McQuilkin, J., & Taylor, B. (2017). Will Web Search Engines Replace Bibliographic Databases in the Systematic Identification of Research? The Journal of Academic Librarianship, 43(1), 8–17. https://doi.org/10.1016/j.acalib.2016.11.003

Breeding, M. (2015). Discovery product functionality. Library technology reports, 50(1), 5-32. https://doi.org/10.5860/ltr.50n1

Cummings, J. (2021). Online navigation to journal articles: How are journal articles retrieved by researchers and students at an academic institution; a quantitative examination of HTTP referer [sic] data. Journal of Electronic Resources Librarianship, 33(2), 63–74. https://doi.org/10.1080/1941126X.2021.1911446

Calvert, K. (2015). Maximizing Academic Library Collections: Measuring Changes in Use Patterns Owing to EBSCO Discovery Service. College & Research Libraries, 76(1), 81–99. https://doi.org/10.5860/crl.76.1.81

Dempsey, Lorcan 2020. Foreword to Resource Discovery for the Twenty-First Century Library: Case Studies and Perspectives on the Role of IT in User Engagement and Empowerment, xxi–xxxii. Edited by Simon McLeish. London: Facet Publishing. OCLC Research Foreword Preprint. https://doi.org/ 10.25333/p8j4-1411

Evans, G. & Schonfeld, R.C. (2020). It's Not What Libraries Hold; It's Who Libraries Serve: Seeking a User- Centered Future for Academic Libraries. Ithaka S+R. 23 January 2020.

Evelhoch, Z. (2016). Web-Scale Discovery: Impact on Library Database Web Page Views and Usage. Journal of Web Librarianship, 10(3), 197–209. https://doi.org/10.1080/19322909.2016.1191048

Getsay, H., & Chen-Gaffey, A. (2021). Understanding COUNTER 5 metrics in context: A case study on vendor database usage reports. Journal of Electronic Resources Librarianship, 33(2), 92–104. https://doi.org/10.1080/1941126X.2021.1911450

Hayman, R. (2017). Early Career Researchers Demand Full-text and Rely on Google to Find Scholarly Sources. Evidence Based Library and Information Practice, 12(4), 256. https://doi.org/10.18438/B8WD4G

Inger, S., & Gardner, T. (2016). How readers discover content in scholarly publications. Information Services & Use, 36(1–2), 81–97. https://doi.org/10.3233/ISU-160800

Lean Library. (2021). Librarian Futures. SAGE Publishing. https://doi.org/10.4135/wp.20211103

Levine-Clark, M., Mcdonald, J., & Price, J. (2014). The effect of discovery systems on online journal usage: a longitudinal study. Insights: The UKSG Journal, 27(3), 249–256. https://doi.org/10.1629/2048-7754.153

Morris, C. (2023). "Northeast Ohio universities still grappling with enrollment declines." Ideastream. https://www.ideastream.org/education/2023-02-27/northeast-ohio-universities-still-grappling-with-enrollment-declines

Ngo, L., Hennesy, C., & Knabe, I. (2019). The Impact of Web-Scale Discovery on the Use of Electronic Resources. Serials Review, 45(4), 227–238. https://doi.org/10.1080/00987913.2019.1695343

Nicholas, D., Boukacem-Zeghmouri, C., Rodríguez-Bravo, B., Xu, J., Watkinson, A., Abrizah, A., Herman, E., & Świgoń, M. (2017). Where and how early career researchers find scholarly information: How ECRs find scholarly information. Learned Publishing, 30(1), 19–29. https://doi.org/10.1002/leap.1087

Nicholas, D., Rowlands, I., Huntington, P., Jamali, H. R., & Hernandez Salazar, P. (2010). Diversity in the e-journal use and information-seeking behaviour of UK researchers. Journal of Documentation, 66 (3), 409–433. https:// doi.org/10.1108/00220410110384760

Pignolet, J. (2022). "University of Akron enrollment drops again, but new student numbers show promise." Akron Beacon Journal. https://www.beaconjournal.com/story/news/education/2022/09/23/university-of-akron-enrollment-drops-but-new-student-numbers-promising/69511428007/

Robins, S. (2022) Full-text retrievals and EBSCO Discovery Service: Assessing usage of e-journals across multiple platforms, Journal of Electronic Resources Librarianship, 34:2, 89-107, DOI: 10.1080/1941126X.2022.2064105

Robinson, L., Snyder, R., & Burke, D. (2013). Discovery and Analysis of the World's Research Collections: JSTOR and Summon Under the Hood. The Serials Librarian, 64(1–4), 196–199. https://doi.org/10.1080/0361526X.2013.760317

Sonsteby, A., & DeJonghe, J. (2013). Usability Testing, User-Centered Design, and LibGuides Subject Guides: A Case Study. Journal of Web Librarianship, 7(1), 83–94. https://doi.org/10.1080/19322909.2013.747366

Wang, Y. (2020). Web-scale discovery and Google Scholar: A study in use patterns. Journal of Electronic Resources Librarianship, 32(1), 1–10.

https://doi.org/10.1080/1941126X.2019.1709722
Wang, Y., & Howard, P. (2012). Google Scholar Usage: An Academic Library's Experience. Journal of Web Librarianship, 6(2), 94–108. https://doi.org/10.1080/19322909.2012.672067

# Appendix A

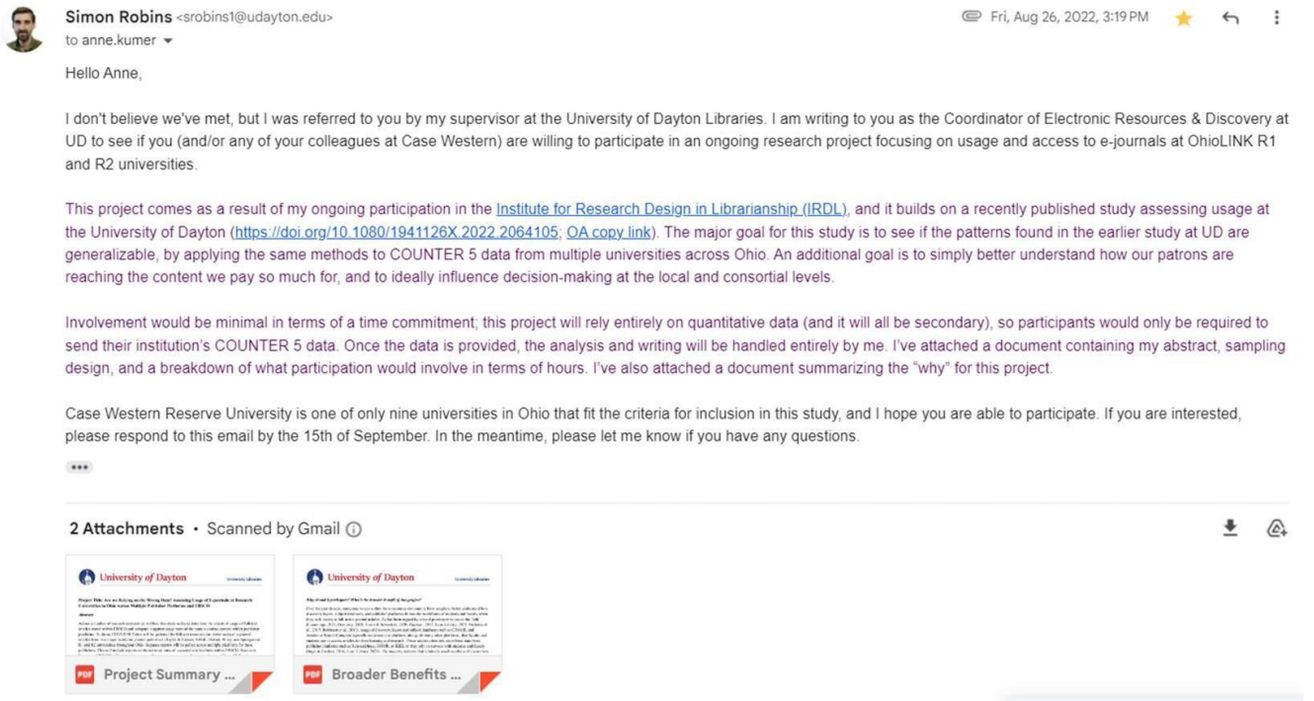
Table A1. Percentage of downloads by publisher and platform (across all universities).

| Publisher | % of downloads by platform | 2019 | 2020 | 2021 | 2022 |
|---|---|---|---|---|---|
| Oxford University Press | EBSCOhost | 26% | 31% | 28% | 25% |
| | Electronic Journal Center | 3% | 3% | 2% | 1% |
| | JSTOR | 16% | 20% | 20% | 21% |
| | Oxford Academic | 55% | 46% | 51% | 53% |
| SAGE Publications | EBSCOhost | 18% | 18% | 15% | 13% |
| | Electronic Journal Center | 22% | 16% | 8% | 5% |
| | JSTOR | 9% | 10% | 8% | 8% |
| | SAGE Journals | 51% | 57% | 68% | 74% |
| Springer | EBSCOhost | 28% | 26% | 24% | 21% |
| | Electronic Journal Center | 8% | 5% | 2% | 1% |
| | JSTOR | 3% | 3% | 2% | 2% |
| | SpringerLink | 61% | 66% | 72% | 75% |
| Taylor & Francis | EBSCOhost | 74% | 74% | 71% | 69% |
| | Electronic Journal Center | 0% | 0% | 0% | 0% |
| | JSTOR | 5% | 5% | 6% | 6% |
| | Taylor & Francis Online | 21% | 21% | 23% | 25% |
| Wiley | EBSCOhost | 30% | 34% | 29% | 24% |
| | Electronic Journal Center | 9% | 6% | 3% | 2% |
| | JSTOR | 4% | 5% | 5% | 5% |
| | Wiley Online Library | 57% | 54% | 63% | 69% |

Table A2. Percentage of item investigations by publisher and platform (at university of dayton only).

| Publisher | Platform | % of Total Item Investigations in 2022 |
|---|---|---|
| APA | APA PsycNet | 54% |
| | EBSCOhost | 46% |
| | ProQuest | 0% |
| Cambridge University Press | Cambridge Core | 34% |
| | EBSCOhost | 42% |
| | JSTOR | 23% |
| | Project MUSE | 1% |
| | ProQuest | 0% |
| Elsevier | EBSCOhost | 27% |
| | JSTOR | 0% |
| | ProQuest | 0% |
| | ScienceDirect | 73% |
| IEEE | EBSCOhost | 14% |
| | IEEE | 86% |
| Johns Hopkins University Press | EBSCOhost | 44% |
| | JSTOR | 33% |
| | Project MUSE | 23% |
| | ProQuest | 0% |
| Nature Publishing Group | EBSCOhost | 19% |
| | Nature.com | 81% |
| | ProQuest | 0% |
| Oxford University Press | EBSCOhost | 54% |
| | JSTOR | 16% |
| | Oxford Academic | 30% |
| | Project MUSE | 0% |
| | ProQuest | 0% |
| SAGE Publications | EBSCOhost | 52% |
| | JSTOR | 5% |
| | ProQuest | 0% |
| | SAGE Journals | 43% |
| Springer | EBSCOhost | 54% |
| | JSTOR | 3% |
| | ProQuest | 0% |
| | SpringerLink (and other Springer Platforms) | 43% |
| Taylor & Francis | EBSCOhost | 92% |
| | JSTOR | 3% |
| | ProQuest | 0% |
| | Taylor & Francis Online | 5% |
| Wiley | EBSCOhost | 61% |
| | JSTOR | 5% |
| | ProQuest | 0% |
| | Wiley Online Library | 34% |

## Appendix B. Sample email invitation and attachments (sent to possible participants)



**Simon Robins** <srobins1@udayton.edu>
to anne.kumer

Fri, Aug 26, 2022, 3:19 PM

Hello Anne,

I don't believe we've met, but I was referred to you by my supervisor at the University of Dayton Libraries. I am writing to you as the Coordinator of Electronic Resources & Discovery at UD to see if you (and/or any of your colleagues at Case Western) are willing to participate in an ongoing research project focusing on usage and access to e-journals at OhioLINK R1 and R2 universities.

This project comes as a result of my ongoing participation in the Institute for Research Design in Librarianship (IRDL), and it builds on a recently published study assessing usage at the University of Dayton (https://doi.org/10.1080/1941126X.2022.2064105; OA copy link). The major goal for this study is to see if the patterns found in the earlier study at UD are generalizable, by applying the same methods to COUNTER 5 data from multiple universities across Ohio. An additional goal is to simply better understand how our patrons are reaching the content we pay so much for, and to ideally influence decision-making at the local and consortial levels.

Involvement would be minimal in terms of a time commitment; this project will rely entirely on quantitative data (and it will all be secondary), so participants would only be required to send their institution's COUNTER 5 data. Once the data is provided, the analysis and writing will be handled entirely by me. I've attached a document containing my abstract, sampling design, and a breakdown of what participation would involve in terms of hours. I've also attached a document summarizing the "why" for this project.

Case Western Reserve University is one of only nine universities in Ohio that fit the criteria for inclusion in this study, and I hope you are able to participate. If you are interested, please respond to this email by the 15th of September. In the meantime, please let me know if you have any questions.

**2 Attachments** • Scanned by Gmail

University of Dayton — Project Summary ...

University of Dayton — Broader Benefits ...

University of Dayton University Libraries

**Project Title: Are we Relying on the Wrong Data? Assessing Usage of E-journals at Research Universities in Ohio Across Multiple Publisher Platforms and EBSCO**

*Abstract:*

Across a number of research universities in Ohio, this study seeks to determine the extent of usage of full-text articles stored within EBSCO and compare it against usage rates of the same e-journal content within publisher platforms. To do so, COUNTER 5 data will be gathered for full-text retrievals (or downloads) of e-journal articles from five major academic journal publishers (Taylor & Francis, SAGE, Oxford, Wiley, and Springer) at R1 and R2 universities throughout Ohio. Separate reports will be pulled across multiple platforms for these publishers. This will include reports on the retrieval rates of e-journal articles from within EBSCO Discovery Service and EBSCO subject databases, as well as reports measuring the retrieval rates of the publishers' e-journals within their respective platforms (such as SpringerLINK or Wiley Online Library). An analysis of this data will be conducted using Tableau and other visualization software. If the findings indicate that the full-text article holdings of these e-journal titles are accessed more often within EBSCO than they are within the publisher platforms, then it could suggest that students and faculty rely heavily on the discovery layer and subject databases for their learning and research (and likely prefer to select EBSCO's full-text holdings over results that link out to external platforms). More broadly, the results could have wide-ranging implications for how librarians and administrators manage these systems and subscriptions, as well as how they incorporate them into their instruction and reference sessions.

*Sampling design*:

- This study will use secondary data to measure full-text retrievals of e-journals by faculty, staff, and students at public and private research universities in Ohio. No IRB will be required.
- For the purpose of this study, "research university" is defined according to two Carnegie Classification of Institutions of Higher Education categories: "Doctoral Universities – Very High Research Activity (R1)" and "Doctoral Universities – High Research Activity (R2)" categories. In Ohio there are five R1 universities (Case Western Reserve University, Kent State University, Ohio State University, Ohio University, and the University of Cincinnati) and seven R2 universities (Bowling Green State University, Cleveland State University, Miami University, University of Akron, University of Dayton, University of Toledo, and Wright State University).
- To quality for participation, any of the universities listed above must belong to the OhioLINK consortium, have access to the OhioLINK expanded pack of EBSCO subject databases, have access to OhioLINK's Electronic Journal Center (EJC), and use EBSCO Discovery Layer (EDS) as their primary discovery layer.

University Libraries • 300 College Park, Dayton, OH 45469-1360 • 937-123-4567 (phone) • 937-123-4567 (fax)
emailaddress@udayton.edu • *www.udayton.edu/libraries*

- Out of the twelve R1 and R2 universities in Ohio, nine of them meet the criteria listed in the preceding bullet point: Case Western Reserve University, Cleveland State University, Kent State University, Ohio University, Miami University, University of Akron, University of Dayton, University of Toledo, and Wright State University.
- All nine universities are being invited to participate. For those that agree, representatives from the university library (typically the e-resources or acquisitions librarian) will be required to provide COUNTER 5 data sets on their university's usage.

***What would agreeing to participate involve? What would be the time commitment?***

- Each university that agrees to participate will be required to pull COUNTER 5 data reports for five publishers: Taylor & Francis, SAGE, Oxford, Wiley, and Springer. If an institution does not subscribe to Taylor & Francis titles, then data from this publisher will not be required. Specifically, participating institutions will need to pull the COUNTER 5 "Journal Requests, Excluding OA_Gold (TR_J1)" report for the following platforms: EBSCO (to measure EDS and all EBSCO subject databases), Taylor & Francis Online, SAGE Journals, Oxford Academic, Wiley Online Library, SpringerLINK, OhioLINK's EJC, and JSTOR (only if their library has a subscription with JSTOR).
- To ensure that the data represents usage over time, participants will need to pull this report for calendar year (CY) 2019, CY2020, CY2021, and CY2022. For EBSCO, COUNTER 5 data for CY2019 will expire in EBSCOadmin towards the end of this year, so I will request data for CYs 19-21 in fall 2022. This initial request will be followed by an additional data collection for CY22 in spring 2023.
- Participants will receive detailed instructions (a step-by-step PDF with screenshots) showing how to retrieve each COUNTER report. There will also be instructions on how to send the data as a set of CSV files. No further time commitment will be required by participating institutions once this data has been provided. **In total, participation should not take more than 1-2 hours in fall 2022, followed by an additional 1-2 hours in spring 2023.**

***Data analysis techniques will be handled entirely by the researcher at the University of Dayton:***

- Once the data is received from all participating institutions, I will use Excel and OpenRefine for data cleaning, and Tableau and Excel will be used to analyze and visualize the data. Separate analyses will be conducted for each institution, with a unique set of visuals for each calendar year. I also expect the visualizations and analysis to evolve as the study progresses.
- For EBSCO, JSTOR, and the EJC, the usage data will be filtered down to include only the e-journal content from the five publishers. This will be done on my end; no adjustments will be required once the COUNTER data has been pulled from EBSCOadmin, JSTOR, and the OhioLINK members portal.
- In late summer/early fall 2023, the primary contacts from all participating institutions will be invited to meet virtually to attend a brief presentation on the findings of the study, as well as to discuss the findings

as a group. This cohort will likely consist primarily of e-resource and acquisition librarians. OhioLINK staff will be invited to attend this meeting too. Attendance is optional and should not exceed 90 minutes.

- A draft of the manuscript for submission (to be sent to peer-reviewed journals in fall 2023) will be sent to participating institutions in advance so that all participants can review how their institution is represented and offer any feedback. Any insight on factors that may influence e-resource usage (for example, instruction) will be encouraged at this stage and will be considered for inclusion in the discussion portion of the manuscript.
- All individuals who pull data and/or participate in the discussion session will be formally acknowledged (and thanked!) in any articles or presentations resulting from this research.

***Why should I participate? What's the broader benefit of this project?***

Over the past decade, numerous voices within the e-resources community have sought to better understand how discovery layers, subject databases, and publisher platforms fit into the workflows of students and faculty when they seek access to full-text e-journal articles. As has been argued by several prominent voices in the field (Cummings, 2021; Dempsey, 2020; Evans & Schonfeld, 2020; Hayman, 2017; Lean Library, 2021; Nicholas et al., 2017; Robinson et al., 2013), usage of discovery layers and subject databases such as CINAHL and Academic Search Complete typically constitute one platform, alongside many other platforms, that faculty and students use to access articles for their learning and research. These studies often rely on referral data from publisher platforms such as ScienceDirect, JSTOR, or IEEE, or they rely on surveys with students and faculty (Inger & Gardner, 2016; Lean Library, 2021). The majority indicate that relatively small numbers of researchers are using discovery layers or other library-controlled subject databases as their starting point to reach the full-text articles held within these publisher platforms. The resulting argument, based in some cases on referral data to support this assertion, is that students and faculty are more likely reaching the full-text held within the publisher platforms via other search interfaces, most often Google and Google Scholar, or through direct referrals (Cummings, 2021, pp. 72-73; Dempsey, 2020 p. 12; Evans & Schonfeld, 2020, p. 15; Hayman, 2017; Lean Library, 2021, p. 28; Robinson et al., 2013, p. 198).

The gap that these studies fail to address is the relationship between the full-text e-journal holdings provided by EBSCO and the holdings provided by the publisher platforms. For example, EBSCO–through EBSCO Discovery Service and through EBSCO subject databases–provides access to vast e-journal holdings which are often duplicates of e-journals found in prominent publisher platforms such as SpringerLINK and Wiley Online Library. The same is true for OhioLINK's EJC and JSTOR. This prompts the question: If a duplicate full-text PDF from a Wiley[1] journal is immediately available from EBSCO's servers, then why would the end user go to a publisher platform like Wiley Online Library to retrieve it? And wouldn't this trend cause the referrals from EBSCO to platforms like Wiley Online Library to be low? This research seeks to answer these questions.

In fall 2021, a preliminary analysis using COUNTER 5 e-journal usage data was conducted at the University of Dayton, and the findings strongly suggest that the collective EBSCO subject databases and discovery layer are serving as the primary platform for users to search for, locate, and retrieve relevant full-text articles, with users often downloading the EBSCO holdings for major publishers more often than the same holdings from publisher platforms. This initial study was published in the Journal of Electronic Resources Librarianship in July 2022 (https://doi.org/10.1080/1941126X.2022.2064105). The logical next step is to explore the same type of usage data at other universities to see if the pattern is generalizable and to answer the

[1] Wiley is being used as an example – it could be any other major publisher that has full-text holdings in EBSCO

following research question: To what extent do end users at research universities in Ohio use EBSCO's search platforms to directly retrieve EBSCO's full-text holdings of e-journals from major publishers, even when the same content is retrievable from additional platforms external to EBSCO?

Once completed, this study will provide new information about e-resources usage that could affect the decision-making of librarians who handle acquisitions, manage library databases, or handle other aspects of the electronic resources lifecycle. If the data shows that the majority of downloads of e-journal articles for these major publishers are occurring within the discovery layer and subject databases, then it would provide librarians and administrators with a more robust (and clearer) understanding of how these systems are being used. Moreover, if this is the case, then the results would strongly suggest that students and faculty likely start with and rely heavily on the discovery layer and subject databases for their learning and research, which would contradict what many prominent voices in the field are saying (Dempsey, 2020, p. 12; Evans & Schonfeld, 2020, p. 15; Lean Library, 2021, p. 28).[2] These findings could in turn impact instruction librarians, influencing them to focus more on demoing EBSCO when teaching. On the other hand, if the findings show that there is relatively little usage of these full-text holdings within discovery layers and databases compared to the usage of publisher platforms, then it would pose several important questions that would equally benefit this population of librarians and have implications for the future of subject databases and discovery layers.

Lastly, this study could impact the way libraries assess their subscriptions. By raising awareness of the duplication in holdings across different platforms and showing how to analyze this usage collectively, more libraries might be able to more fully understand how their e-journal titles are being used, and they might be encouraged by this study to use COUNTER 5 data and visualization software like Tableau. This can support more thorough data-driven decision-making when it comes to renewing subscriptions at the local-level and through consortiums like OhioLINK.

**References**

Cummings, J. (2021). Online navigation to journal articles: How are journal articles retrieved by researchers and students at an academic institution; a quantitative examination of HTTP referer [sic] data. *Journal of Electronic Resources Librarianship*, *33*(2), 63–74. https://doi.org/10.1080/1941126X.2021.1911446

Calvert, K. (2015). Maximizing Academic Library Collections: Measuring Changes in Use Patterns Owing to EBSCO Discovery Service. *College & Research Libraries*, *76*(1), 81–99. https://doi.org/10.5860/crl.76.1.81

[2] If there is a negative correlation, this assertion of users starting with the library databases and discovery layer would be relatively easy to determine because of the fact that the end user can only reach these systems from a proxy link. The same is not true for publisher platforms when users are within a university's on-campus IP range.