

Assessing Digital Primary Source Usage: Value and Impact for Academic Libraries

Alex Butler, Adam Matthew Digital, July 2020

ABSTRACT

This paper aims to explore how digital primary sources are used in parallel with how academic librarians assess usage and value in these collections. By examining key metrics, case studies and literature around assessing impact of digital primary source collections, this white paper interrogates how patrons utilize digital primary sources and whether those case studies draw parallel with the ways in which librarians judge impact and value of such collections in the context of their own institutions.

INTRODUCTION

94% [of respondents] felt that digitized materials would be important, very important or essential to their future work.

In producing digital primary source collections for nearly 20 years, Adam Matthew Digital is aware of how valuable digitization has been to scholars since the 1990s. Digitized sources are fast becoming essential for Humanities research. Meyer et al. showed in their 2009 survey that “94% [of respondents] felt that digitized materials would be important, very important or essential to their future work (Meyer, Eccles, Thelwall, & Madsen, 2009, pp.151-152).” A recent EU Commission report argued that “there is probably no greater ambition than to perpetuate our rich cultural heritage” and that “digitization breathes new life into material from the past, and turns it into a formidable asset for the individual user and an important building block of the digital economy” (Comité des Sages, 2011, p.1&4). Digitized primary sources are a new resource that libraries must ensure access to for their patrons in the future.

As more archival materials become digitized, universities are weighing up the value of individual scholars taking research leave as opposed to enabling online access to materials for their institution’s entire research community. There are few disputing the difference between digital access and a physical

archive as Carol Steedman has shown¹. Many librarians and scholars argue for the “increasing options for users and undisputed accessibility” that digital archival collections beyond all else provide (Tanner & Deegan, 2011, p.34). This is exacerbated by the current and near-future restrictions to archive access due to COVID-19. Publishers are seeing the demand for digitized archival material growing across a range of disciplines and research areas. Wiley’s 2019 White Paper titled ‘Primary Source Digital Archives: Effectively Measuring a Digital Archive’s Value’ shows that primary source use was “almost equally dispersed among the health sciences, life sciences, physical sciences, and social sciences and humanities.” (p.2)

As digitized archival content becomes more of a mainstay in library holdings, it competes for funding with more traditional publisher offerings such as eBooks, databases and learning solution platforms. By looking at usage case studies and interviewing librarians about how they perceive digital primary sources in the arts and humanities and social sciences fields in particular, this paper aims to investigate how libraries assess value in digital primary sources in parallel with some high-impact user case studies.

¹ Steedman, 2002, p.7

METRICS AND ASSESSMENT TOOLS AVAILABLE

Publishers utilize a range of internal analytics tools that allow them to monitor engagement with a product. Audited, transparent usage statistics have become universal and independent thanks to COUNTER's continuous auditing and releases. They serve an important role in helping libraries understand their e-resource usage. Adam Matthew's AMalytics portal and COUNTER work in tandem to provide clear usage analytics; both are featured in the case study examples in this section.

Parallel to quantitative metrics, publishers record and track many qualitative metrics such as case studies, user feedback as well as courses naming resources on reading lists, Blackboard or Moodle. These are also essential in helping publishers and libraries understand how their digital primary source collections are being used. This white paper and the following case studies investigate how different metrics help measure usage and engagement with this type of specialist content.

NOTE ON METRIC AND ANALYTICS VARIANCE ACROSS GRAPHICS:

I have kept internal AMalytics and COUNTER data separated to ensure accurate reporting, and to emphasize the difference in reporting styles. AMalytics metrics are specifically designed to treat all the pages on the products equally. COUNTER's DB1 master reports focus primarily on searching and accessing documents, but omit contextual pages, academic essays, interactive tools and videos.

Adam Matthew Metrics: *Page views, sessions, Page views per session, session duration, downloads, unique users*

COUNTER 4 DB1 Metrics: *Regular Searches, Result Clicks, Record Views*

COUNTER 5 DB1 Metrics: *Total Item Investigations, Unique Item Investigations, Unique Title Investigations, Total Item Requests, Unique Item Requests, Unique Title Requests, Searches Regular*

'WHAT DOES GOOD IMPACT LOOK LIKE?' ASSESSING ADAM MATTHEW USAGE CASE STUDIES

There are challenges to assessing usage and the impact of digital primary source content. These challenges are not always present with homogeneous journal or eBook platforms that "all follow the same basic rules" and can be "measured using similar bibliographic techniques" (Meyer et al, 2009). The following case studies demonstrate that primary source collections differ widely in material type and research subject area and using a single metric to assess value across this content range would paint a misleading picture. 19th century digital newspapers require a different style of enquiry to 16th century commonplace books or manuscripts, for example. Trying to fit these types of content usage into the same expectations of usage can be problematic for any publisher or librarian attempting to make collection development or acquisitions decisions.

The following case studies look at researchers and university courses at three different North American institutions, who have all made use of Adam Matthew collections. These case studies represent the three typical research and learning environments that benefit from online primary source collections. These environments include an instructor-led classroom activity, independent homework/student research task and scholarly research by faculty and graduate students.

PhD Researcher Large, Research-Intensive University, Canada

Postgraduate and faculty research represents a significant portion of Adam Matthew usage statistics. This usage is often categorized by small bursts of 'deep research' over the summer months, when faculty and researchers have undisturbed research time. This is demonstrable in this case study. A PhD student in the English department of a top-tier Canadian university was conducting research "going through every promptbook that was owned by an actor page by page, searching for any notes made in the actor's own hand...I have to go through them one by one from cover to cover". These 'short burst' spikes of usage do not always reflect the significant numbers that librarians would generally see when undergraduate classes utilize a resource. High average session duration, 'page views per session' and low session numbers would indicate

an individual researcher using the collection. This is an essential and positive method of access to individual researchers. According to the PhD student, "the database enables us to conduct far more exhaustive research than I could possibly have done without it".

The difference between the COUNTER 4 report and AMalytics is worth noting when considering this successful usage case study. COUNTER 4 focuses on the number of searches and record views but does not give any indication to the 'depth' of the research. It is common for researchers to run simple searches within an archival collection to spread the net wide and interrogate all results. COUNTER 4 'Record Views' also miss how long researchers spend in a document and how many pages they access.

AMalytics (internal analytics platform)

METRIC	2019-07-01 to 2019-08-31
SESSIONS	57
PAGE VIEWS	1,185
AVERAGE PAGE VIEWS PER SESSION	20.79
AVERAGE SESSION DURATION	14:15
PDF DOCUMENTS DOWNLOADED	36
PDF PAGES DOWNLOADED	196

COUNTER 4

USER ACTIVITY	2019-07-01 to 2019-08-31
REGULAR SEARCHES	38
SEARCHES-FEDERATED AND AUTOMATED	0
RESULT CLICKS	71
RECORD VIEWS	742

This case study emphasizes to both publishers and librarians seeking to measure impact the importance of qualitative feedback from individual researchers. Relying on simple usage metrics does not give a full picture of how online primary sources contribute to research needs, so being able to collate qualitative feedback will provide a fuller picture on positive and impactful usage.

Independent Student Research Assignment, Medium Regional University, USA

The second case study focuses on a second-year undergraduate class of 18 using *Race Relations in America* to study the history of the American civil rights movement. Each week students engage with a “seminal secondary work in the field” and then explore *Race Relations in America* “to locate documents that explain and create a context to their secondary reading on different eras or geographies”. The students were asked to explore the site and bring a source to class. The students would then write a blog post with references and analysis of the source they had curated.

Assigning independent retrieval of a primary source is another common type of Adam Matthew product engagement, and this case study is a good example of this type of assignment. Given the time restraints of a single week and student workload, these usage numbers reflect the quick retrieval of a document, and how retrieving material from a

largely full-text print collection is an easier process for undergraduates than engaging with handwritten manuscript content.

The difference in reporting styles between COUNTER 5 and AMalytics is again evident. One advantage of COUNTER 5 DB1 reports is its ability to indicate how many individual documents have been accessed on a site. It is also interesting to note the small number of PDF downloads. It is uncommon to see undergraduates taking primary sources offline and out of the context of the database. This case study represents a shorter-term engagement with an Adam Matthew product, but because the instructor shared qualitative feedback on the course, it is possible to frame how the database was used and the impact it had not only on usage numbers, but also on students achieving the learning outcomes set for the course.

AMalytics (internal analytics platform)

METRIC	February - April 2020
SESSIONS	115
PAGE VIEWS	1,372
AVERAGE PAGE VIEWS PER SESSION	11.93
AVERAGE SESSION DURATION	11:33
PDF DOCUMENTS DOWNLOADED	9
PDF PAGES DOWNLOADED	195

COUNTER 5

METRIC	February - April 2020
TOTAL ITEM INVESTIGATIONS	568
UNIQUE ITEM INVESTIGATIONS	196
UNIQUE TITLE INVESTIGATIONS	196
TOTAL ITEM REQUESTS	337
UNIQUE ITEM REQUESTS	147
UNIQUE TITLE REQUESTS	147
SEARCHES REGULAR	114

3

Instructor-led Classroom Activity, Medium State University, USA

This case study focusses on an upper-level History class 'American Colonial Civilizations' at a medium size institution in the Western United States. Students were asked to build a group presentation on a topic of their choice using primary sources they had discovered during the in-class "research week". The instructor allocated a "research week in which each [student] team located a question that arose during our discussion of the monograph, and spent the week doing research in the Colonial America collection to answer that question."

Students were then asked to provide their group presentation three weeks later in the Week 9 class.

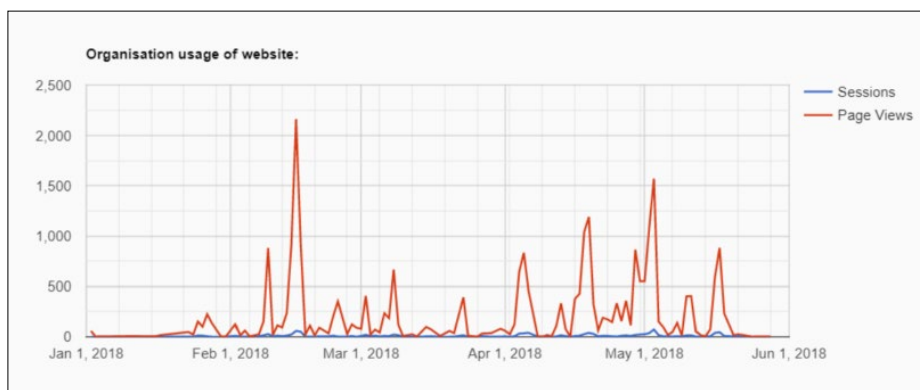
The associated usage numbers reflect a typical in-class research activity completed either on individual laptops or in a campus computer suite. Students have a limited amount of time and a designated period in the curriculum to explore a primary source archive online.

AMalytics (internal analytics platform)

METRIC	Q1 - Q2 2018
SESSIONS	1,558
PAGE VIEWS	28,423
AVERAGE PAGE VIEWS PER SESSION	18.24
AVERAGE SESSION DURATION	15:39
PDF DOCUMENTS DOWNLOADED	169
PDF PAGES DOWNLOADED	1,423

COUNTER 4

USER ACTIVITY	2019-07-01 to 2019-08-31
REGULAR SEARCHES	2,991
SEARCHES-FEDERATED AND AUTOMATED	0
RESULT CLICKS	4,870
RECORD VIEWS	11,698



The core usage metrics demonstrate what an in-class guided research task can produce. Students undertook independent research in-class, followed by a period of synthesizing their work with their peers outside of the classroom. The digital archive *Colonial America* was the central focus of this module's assessment task and the numbers reflect that centrality. Viewing collection usage on the timeline, the highest use came during weeks when in-class team research meetings and presentations were structured into the syllabus, as well as in the lead-up to the submission of final research papers and the optional oral presentations of these papers for extra credit in the third week of May.

The low download figure is also worth highlighting. Students were willing to come to the site and do short, sharp sessions of active research but kept everything online, either saving the discovered documents in browser Bookmarks or by using the 'My Archive' bookmarking feature within the

database. Also worth highlighting is the average session duration. While in class, students engaged in more intensive periods of usage, than the average user of *Colonial America*.

Most importantly, these large page views and 'pageviews per session' correlate with the positive qualitative feedback from both instructor and students. The students "were really proud of the work they did" and "felt like they were real historians". This class is a perfect example of a digital resource becoming the students' central focus after being introduced by the instructor and integrated into the curriculum through assignments and in-class research activities. Combining the qualitative feedback from users and the quantitative metrics for *Colonial America* helps build a picture of what successful and useful engagement with the digital resource looks like, for Adam Matthew as a product publisher and for libraries making purchasing decisions.

HOW DO WE ASSESS 'VALUE'?

Looking at usage case studies helps demonstrate how core product usage metrics correlate to the qualitative impact that access to primary sources can have on research and teaching. Combining these two sources of usage feedback presents further questions: How can we assess the impact and value of these types of collections? Are there further metrics libraries can use to understand how impactful primary source content is in comparison to other e-resources? Chassanoff, for example, looks to include more holistic measures of impact such as "improving access to research materials and decreasing the time and distance required to access materials" or "generating completely new audiences for previously inaccessible material" (Chassanoff, 2003, p.463). Moreover, beyond these standardized metrics, each academic library has a different funding, infrastructure and "acquisition style" (Levine-Clark, 2018) for serving the research and teaching needs of their institutions.

To build a better understanding of how librarians perceive digital primary sources, Adam Matthew interviewed librarians from two US universities to build a qualitative understanding of how they purchase digital primary sources, how their patrons make use of collections and how they assess their value and impact compared to other e-resources in the library.

Interview A was conducted with a subject liaison librarian for English and History at a large private university in Southwest United States. Details have been redacted on request. Interview B was conducted with a Special Collections Librarian and two E-Resource Librarians from a large state university in the Northwestern United States.

FRAMING ONE-TIME PURCHASE AND VENDOR RELATIONSHIPS

There is an extensive range of different access and acquisition models for e-resources in 2020, and both interviews indicated that one-time purchase and perpetual access is fitting and advantageous for digital archival content. Having perpetual access allows libraries to “lock it in... because you don’t know what future budgets are going to look like” (Interview B, 2020). Michael Levine-Clark, Dean and Director at University of Denver Libraries, approaches one-time purchasing as a useful model for purchasing “as many of these collections as we could afford, since it is impossible to predict what sorts of topics students will choose to write on. More collections provide more options and give students more opportunities to conduct original research on topics that excite them.”

One-time purchase was the preferable route for all librarians interviewed, but each institution had their own acquisition procedure. Interviewee A discussed how “any purchase over \$1000 has to run through the collection development committee for review and approval”, whereas interviewee B “just manages requests when they come in” from faculty and library colleagues. Availability of data mining and the “quality of previous purchases...and support we’ve received” (Interview A) were also discussed as factors when purchasing digital primary source content. With the publication of more clean digitized historical datasets every year, data and text mining opportunities are increasingly popular with graduate students, faculty and even some undergraduate classes. When making purchasing decisions, products’ suitability for these needs are increasingly taken into consideration.

ROLE OF CURATION & ORGANIZATIONAL TRUST IN COLLECTION DEVELOPMENT

Both interviewees stated how librarians selecting and curating archival purchases is an essential part of the acquisition of digital primary source

content. Digital archival products are “all sorts of shapes and sizes” (Interview A) and serve a range of research and teaching needs at different times. To purchase content that is “going to fit our collection” (Interview B) and “fit our future needs” (Interview A), librarians must work cohesively with colleagues across disciplines and teams to ensure they are purchasing with direction. When looking at a purchase, both interviewees talked about assessing “what classes this will support” and referenced instances in which new faculty joining or departmental strengths were the primary factor for purchasing. In the words of Interviewee A, purchases “are qualitative decisions with the hope [that] there will be quantitative back up”.

Secondly, all the interviewees talked about trust that had been built up in cohesive Collection Development and/or Acquisitions teams where “it’s very cooperative for us” (Interview A) and “everyone’s a professional...who can ask [each other] for guidance” (Interview B). The librarians were unanimous in emphasizing how that cohesion in the library staff allows them to purchase primary source content with confidence.

IMPORTANCE OF QUALITATIVE FEEDBACK FROM FACULTY AND STUDENTS

In discussing the factors contributing to purchasing decisions and management of digital primary source content, the importance of qualitative feedback became evident, not only in the trial period pre-purchase, but also in the measurement or reporting process, too. Interviewee B praised how insightful their structured open trials process is, for example. Faculty and students “can give feedback so they can say ‘yeah this has really great content that we continue to see value in’ which then “helps inform whether or not we want to purchase [the database]”. This process is becoming a common occurrence in academic libraries, such as Liverpool, Alberta, and Leeds Beckett. Open feedback and comments on library trials serve as a “transparent and open

way” for library patrons to “build a qualitative and nuanced picture of how a certain resource would impact their teaching and learning...in view of their peers, classmates and librarians” (Shelbe & Wildemuth, 2009, p.172).

Qualitative feedback is also essential for measuring impact and use post-purchase. Both interviewees highlighted how archival sources rely on anecdotal feedback, case studies and in-person experience to build a picture of impact, which tends to be more useful than more “linear, secondary sources” (Interview A):

If users are saying “this is really great”, there is a lot of value in that feedback. Some of the metrics we get from vendor COUNTER data, sometimes there are a lot of questions about whether it’s accurate, what does it mean? How are they counting this? Was the object downloaded or did someone just quickly look at it? (Interview B).

Both interviews drew on the narrow window that COUNTER and usage analytics provide for digital primary sources. Focusing on numbers that appear “to be small raises the risk that a resource will be seen as insignificant” (Hutchings et al, 2011) without expanding to further metrics. Primary source research is difficult to pin down with pure usage statistics, as the case studies above show. It is not a ‘seek and retrieve’ research method that can be easily tracked through ‘number of sessions’ or record views, as seen in the variance of the case studies in the previous section. Primary source research requires time, patience and, most importantly, does

not always produce the research results or answers hypothesized. Chassanoff goes further to suggest that numeric data does little to aid librarians in “understanding aspects of information use” in digital archival products, as they “focus primarily on quantifying use through transaction log analysis or citation analysis, rather than exploring scholarly use of, and satisfaction with, digital resources.” (Chassanoff, 2003, p.463).

There is no universal formula to where the balance between qualitative and quantitative sits for all e-resources, but both interviews have shown so far that qualitative feedback from patrons, understanding usage case studies and relying on “trained librarians” who “know their faculty and departments” (Interview A) is a significant factor in value or impact assessment of digital primary sources.

ARCHIVAL CONTENT SERVES A DIFFERENT RESEARCH NEED TO SECONDARY SOURCES.

Lastly, the interviews touched on taking different approaches to primary and secondary sources in the purchasing process. Primary sources serve a unique role to library patrons and a different research purpose to secondary sources such as eBook or journal content. One librarian raised the distinction between secondary sources, which need to be assessed on “what percentage of the content is scholarly” and cost, in comparison to primary sources, which “don’t necessarily apply” this scrutiny as “archival sources are what creates scholarly content” (Interview A).

CONCLUSION

The case studies represented in this paper, and the various styles of collection development discussed show that there is no single way of measuring value in digital primary sources that would be satisfactory to all libraries or publishers. The very nature of archival sources is their uniqueness. Experienced subject specialists, working in cohesive and communicative teams are best placed to interpret the value of certain primary sources, based on both qualitative and quantitative metrics and discussion in the library.

Trusting in librarians' experience to select and utilize primary sources was cited in both interviews as the most positive step an administration can take. This is also reflected in case study two and three. Both case studies involved the subject specialist to oversee and advise on resource implementation in class and to capture the feedback and impact on students' learning experience. Of course, usage statistics serve a vital role, but the interviewees underlined the importance of framing usage numbers as a tool

to serve librarians at their discretion, rather than making decisions for them.

Lastly, drawing on high-impact case studies and discussing how librarians must utilize qualitative metrics to justify acquisitions underlined the importance of faculty and student feedback. Libraries understand the value of building qualitative and detailed feedback and case studies from users. Collating this qualitative feedback not only gives libraries a clear view of the impact a resource can have on teaching or research; it also builds important bridges with faculty. These positive relationships result in more-informed librarians, who in turn can continue to make nuanced collection development decisions on digital primary sources

REFERENCES

- Chassanoff, Alexandra, "Historians and the Use of Primary Source Materials in the Digital Age", *American Archivist*, 76 (2): 458-80
- Cleverley, Julie, and Samantha Heeson. 2019. "Usage, Engagement and Impact: Evaluating the Usage of and Measuring Impact and Engagement with Library Resources at Leeds Beckett University Library". *Insights* 32 (1): 25. DOI: <http://doi.org/10.1629/uksg.474>
- Comité des Sages. (2011). *The New Renaissance: Report of the 'Comité des Sages' Reflection Group on Bringing Europe's Cultural Heritage Online*. Report. Brussels: The European Commission.
- Levine-Clark, Michael, "Going to the Source: Understanding Usage of Primary Source Materials Worldwide", Adam Matthew Website, 2018 <<https://www.amdigital.co.uk/images/product-casestudies/pdf/Going-to-the-Source---Understanding-Usage-of-Primary-Source-Materials-Worldwide---White-Paper.pdf>
- Meyer, E. T., Eccles, K., Thelwall, M., & Madsen, C. (2009). *Final Report to JISC on the Usage and Impact Study of JISC-funded Phase 1 Digitisation Projects & the Toolkit for the Impact of Digitised Scholarly Resources (TIDSR)*. Retrieved from http://microsites.oii.ox.ac.uk/tidsr/system/files/TIDSR_FinalReport_20July2009.pdf
- Sheble, Laura and Barbara Wildemuth, "Transaction Logs," in *Applications of Social Research Methods to Questions in Information and Library Science*, ed. Barbara Wildemuth (Westport, Conn.: Libraries Unlimited, 2009), 174.
- Steedman, Carolyn, *Dust: The Archive and Cultural History*, New Brunswick, Rutgers University Press, 2002
- Tanner, S. & Deegan, M. (2011). *Inspiring Research, Inspiring Scholarship: The value and benefits of digitised resources for learning, teaching, research and enjoyment*. Report. London: JISC. Retrieved from http://www.kdcs.kcl.ac.uk/fileadmin/documents/Inspiring_Research_Inspiring_Scholarship_2011_SimonTanner.pdf