Transforming passive usage into actionable insights
A case study in data analytics from the Bibliotheca Alexandrina

1. Introduction and context

Public libraries have always adapted to reflect the needs of the communities they serve. One of the main tools traditionally used for collection management is circulation statistics, which records what materials are borrowed and how often. While this data provides valuable insight, it does not tell the full story, especially in institutions where many materials are used inside the library without being checked out.

At the Bibliotheca Alexandrina (BA), borrowing privileges are limited to specific user groups such as researchers. For the general public and most visitors, library use takes place entirely within the reading halls. Books are requested, read, and returned without leaving a trace in the circulation system. As a result, collection development decisions risk being based on incomplete data that fails to account for actual patterns of in-house use.

Research has found that a substantial portion of print book use occurs entirely within the library, significantly influencing overall demand,[1] a pattern this project sets out to capture systematically at BA. To address this limitation, the library introduced a method for capturing in-library usage more accurately. Using RFID technology, staff capture in-house items, such as items left on reading tables, before they are reshelved. These records are then combined with catalog metadata (subject, language, call number, publisher, and physical location) to understand which materials are being used, and when, allowing the library to recognize emerging subject areas, shifts in user interest, and underused parts of the collection.

This report presents a practice-oriented case study of how such data can be organized and interpreted using RFID, Excel, and Google Looker Studio[2] to track and visualize in-library usage trends and to support more responsive collection development strategies. It also demonstrates how libraries can adopt lean business intelligence (BI) practices as a cost-effective alternative to expensive smart-shelving technologies.[3]

1 Rose-Wiles, Lisa M.; Shea, Gerard; Kehnemuyi, Kaitlin: Read in or check out. A four-year analysis of circulation and in-house use of print books, in: The Journal of Academic Librarianship 46 (4), 2020. https://doi.org/10.1016/j.acalib.2020.102157.

2 Looker Studio was renamed Data Studio in April 2026. In this paper, however, the name Looker Studio is used as this was the name at the time the study was conducted.

3 This paper refers to the author's presentation with the title "From Passive Usage into Actionable Insights. A Case Study in Data Analysis at the Bibliotheca Alexandrina" on 22 May 2026 at the 114th BiblioCon in Berlin, Germany.

2. Background: Why circulation data is not enough

Circulation data has traditionally served as a primary indicator of collection usage, guiding decisions on acquisitions and resource allocation. Yet this metric fails to reflect on-site usage of materials consulted within reading halls or reference sections without formal checkout.[4] Ignoring this 'invisible' usage risks misjudging reader interests and misaligning collection development priorities.

Despite the wide availability of digital information, many authors stress the value of browsing physical collections, in particular the serendipitous discovery of material[5] Bethany Sewell notes that although the traditional browsing experience has changed with increasing digitization and rapid interlibrary loan services, the need for serendipitous discovery remains essential, especially in a research environment.[6]

Over the past two decades, libraries have increasingly adopted RFID technology and observational tracking to monitor in-library usage more effectively. RFID systems enable automated scanning of multiple items without line-of-sight, reducing staff workload.[7] Studies have shown that integrating RFID within academic library workflows can significantly improve efficiency and data collection for both circulation and in-library tracking.[8] RFID data provides detailed insights into reading hall activity, since each scan records when a specific item is removed from or returned to its location, providing information about actual usage that traditional circulation logs miss.[9] Manual shelf-reading records and reading hall visit logs have also been used in some libraries to complement digital methods.[10] These approaches demonstrate how libraries are evolving towards capturing passive usage, providing a richer picture of user engagement. Consequently, there remains a clear need for accessible methods that use routine RFID sweeps to capture high-quality, actionable data without requiring major infrastructure investment.

4 Rose-Wiles; Shea; Kehnemuyi: Read in or check out, 2020.

5 Barclay, Donald A.: The myth of browsing, in: American Libraries 41 (6/7), 2010, pp. 52–54; Bird, Ruth: Libraries: The value of just in case, not just in time, Peace Palace Library Blog, 21.08.2015, https://peacepalacelibrary.nl/blog/2015/libraries-value-just-case-not-just-time, last accessed: 27.08.2026; Kirk, Thomas: What has happened to browsing collections in academic libraries?, in: Library Issues 30 (5), 2020, pp. 1–4; Lederer, Naomi: Why libraries should retain a core print reference collection, in: The Reference Librarian 57 (4), 2016, pp. 307–322; McKay, Dana; Chang, Shanton; Smith, Wally; Buchanan, George: The things we talk about when we talk about browsing, in: Journal of the Association for Information Science and Technology 70 (12), 2019, pp. 1383–1394.

6 Sewell, Bethany B.: A bookless library, part I. Relocating print materials to off-site storage, in: Journal of Access Services 10 (1), 2013, pp. 43–50. https://doi.org/10.1080/15367967.2013.738392.

7 Mayfield, Kendra Tagging books to prevent theft, Wired, 20.05.2002, https://www.wired.com/2002/05/tagging-books-to-prevent-theft/, last accessed: 18.08.2026.

8 Zulfeen, Sana: RFID technology in academic libraries, in: International Journal of Research and Analytical Reviews 5 (1), 2018, pp. 1–4.

9 Shahid, Syed M.: Use of RFID technology in libraries. A new approach to circulation, tracking, inventorying, and security of library materials, in: Library Philosophy and Practice 8 (1), 03.10.2005, https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1069&context=libphilprac, last accessed: 28.08.2026; Indraji, Dominic, J.; Jaganbabu, J. Jothi, T.: Prospects and challenges of RFID applications in libraries. A case study, in: Journal of Advances in Library and Information Science 12, 2023, pp. 205–209. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5104995, last accessed: 28.08.2026.

10 Ferguson, Stuart; Thornley, Clare; Gibb, Forbes: How do libraries manage the ethical and privacy issues of RFID implementation?, in: Journal of Librarianship and Information Science 47 (2), 2015, pp. 117–130, https://doi.org/10.1177/0961000613518572.

In order to transform operational data into strategic value, libraries are increasingly adopting Business Intelligence (BI) and Evidence-Based Library Management (EBLM) approaches. Historically, library statistics were confined to static end-of-year spreadsheet reports or siloed within the Integrated Library Syssets (ILS), making it difficult to spot emerging trends or perform multi-dimensional analysis.[11] The adoption of data visualization tools and dynamic dashboards has changed the way in which library managers interact with usage metrics, enabling them to filter data interactively by subject, publication year, or collection.[12] Building on this need, the present project set out a practical workflow using routine RFID sweeps, Microsoft Excel, and Google Looker Studio to visualize non-circulation usage at the BA.

3. Project description

3.1 Scope and setting

The project was carried out in the Main Library reading halls of the Bibliotheca Alexandrina, covering six floor levels: F2, F1, the Entry level (E), and basement levels B1, B2, B3, and B4 (including the Multimedia section). Three specialized dependent libraries, the Taha Hussein Library for the visually impaired, the Children's Library, and the Young People's Library, were excluded from the scope due to their distinct operational models, user demographics, and collection policies, which would have introduced confounding factors into the analysis.

Data collection ran for a full 12-month operational cycle from June 2024 to May 2025, chosen to capture the full seasonal variation of patron attendance, including the high-traffic autumn academic semester (September–December), the mid-year examination period (January–May), and the non-academic summer period (June–August). It should be noted that this represents a single annual cycle and that findings may be influenced by year-specific external factors, such as institutional events, national holidays, or post-pandemic attendance recovery, that cannot be controlled for in a single-cycle design. Repeating the exercise across several annual cycles would strengthen the generalizability of the findings. The total dataset comprised 130,693 in-house item scan events, representing total usage events rather than unique titles; an item consulted on multiple occasions generates a corresponding number of scan records, providing a quantitative measure of the relative demand intensity for individual resources.[13]

3.2 Data collection workflow and privacy

The Bibliotheca Alexandrina enforces an institutional policy requiring patrons to leave consulted materials on reading tables rather than reshelving them independently. This policy, designed to maintain the integrity of the open-stack classification system, creates a centralized and consistent item-handling workflow through which all in-house usage is channeled. Three dedicated scanning stations,

11 Morton-Owens, Emily; Hanson, Karen L.: Trends at a glance. A management dashboard of library statistics, in: Information Technology and Libraries 31 (3), 2012, pp. 36–51.

12 Showers, Ben: Library Analytics and Metrics Using data to drive decisions and services, London 2015.

13 Merged Data Book Tracking, https://datastudio.google.com/reporting/b62fd32c-0dd1-4d0c-a074-f841fe04677a/page/Y8wHE?s=kDmbLvF0srE, last accessed: 28.08.2026.

each equipped with an RFID handheld device, were established at strategic positions across the reading hall floors.

A custom, lightweight mobile application was developed and deployed on each device, operating on a session-based architecture: each launch initiated a new scanning batch, and each barcode scan generated a timestamped log entry linked to the item's embedded RFID tag. This session-based design served as a primary safeguard against accidental duplicate scans within a session. Legitimate repeat consultations, meaning the same item scanned in separate sessions on the same or different days, were retained as valid, separate events, since they represent real, independently occurring usage.

To assess reliability, a cross-validation step was applied during data preparation: the barcodes recorded in the monthly CSV exports were cross-referenced against the library's Integrated Library System (ILS) catalog. Records that did not match a valid catalog entry were flagged and investigated. These typically corresponded to items with outdated barcodes, deaccessioned materials still physically present in the collection, or cataloguing anomalies, and were excluded from the final dataset[14] after review by the cataloguing department. All 130,693 retained events correspond to verified, active catalog items. With respect to patron privacy, the method is strictly item-centric: scan records are not linked to patron profiles, borrower cards, or any other personally identifiable information, in line with patron privacy principles and institutional data governance standards.

3.3 Data enrichment and dashboard development

Following a monthly CSV export from the three handheld devices, the raw scan logs were uploaded to a custom internal web application developed by BA's Information and Communication Technology (ICT) sector. This application queried the library's catalog database to automatically retrieve and append bibliographic metadata for each scanned barcode, including item title, Dewey Decimal Classification (DDC) call number, subject classification, publisher, language, and physical location (floor level, F2 through B4).

The enriched monthly datasets from the three devices were consolidated into a central master database hosted in Google Sheets. A custom Google Apps Script automated the import, appending, and cross-verification processes, applying a final deduplication pass to remove residual formatting inconsistencies across device exports.[15]

The master dataset was then connected in real time to Google Looker Studio, where an interactive dashboard was built to enable dynamic filtering and visual exploration of the 130,693 annual scan events across time, subject, language, and spatial location. This forms a lean, cloud-based Evidence-Based Library Management (EBLM) pipeline built entirely on existing infrastructure and freely available tools, with no additional software licensing costs.

14 A total of 241 records were excluded during data validation because they did not match a valid catalog entry.
15 Google Gemini was used to assist in the development of the Google Apps Script for RFID data collection (July 2025).

4. Results

4.1 Temporal and seasonal patterns

The temporal data revealed periods of intense in-house library usage, closely tied to the local academic calendar and seasonal shifts. The highest volume of passive usage occurred in November 2024, aligned with the mid-autumn academic semester, recording 32,629 item scans (24.9% of the entire year's in-house usage). This peak was followed by strong engagement in October (24,689 scans) and December (13,699 scans). A secondary, non-academic peak was identified in summer, with July recording 11,330 scans, indicating sustained public and researcher engagement outside standard university semesters. The lowest activity was recorded in May 2025 (739 scans), consistent with end-of-academic-year closure periods.

4.2 Language of materials used

The metadata enrichment allowed a precise breakdown of patrons' linguistic preferences. Arabic-language materials dominated in-house usage, accounting for 64.8% of total scan events. English-language materials formed the second-largest segment at 25.1%, followed by French, Spanish, and other languages together comprising the remaining. This distribution provides concrete evidence for collection managers on where to allocate foreign-language acquisition budgets, based on actual, localized demand rather than assumed preferences.

4.3 Subject areas and floor usage

Cross-referencing the scanned DDC call numbers with the physical reading hall locations revealed which subject areas were attracting the most visitors. The most frequently consulted DDC classifications were: 932 (History of Ancient Egypt) with 2,361 scans; 892.737 (Modern Arabic Fiction and Literature) with 2,129 scans; 297.14 (Islamic Studies and Sources) with 1,372 scans; and 892.736 (Arabic Literary Criticism) with 1,095 scans. Political Science (DDC 320.917) and World History (DDC 909.097) also registered significant demand, with 999 and 985 scans, respectively.

Spatially, this thematic demand correlated directly with activity levels across the different library floors. Level B2 recorded the highest activity, with 55,880 scans (42.8% of the total), followed by Level F1 with 26,528 scans (20.3%) and the Entry Level (E) with 16,238 scans (12.4%). Together, these three zones accounted for 75.5% of all in-library usage data that can be used to inform library operations directly enabling management to plan staff deployment, reading-desk availability, and daily reshelving resources for these specific, high-traffic zones.

5. Discussion

5.1 Evidence-based collection development

The project shows that relying solely on circulation statistics gives an incomplete picture of library use. By capturing and visualizing more than 130,000 in-house usage events, it offers a scalable blueprint for Evidence-Based Library Management (EBLM). Its most significant impact lies in its direct application to collection development and acquisition strategy: the thematic and linguistic data provide an

objective, patron-driven basis for acquisitions. The volume of in-house consultations for Arabic literature (DDC 892) and ancient Egyptian history (DDC 932) signals clear, localized demand not reflected in formal checkout logs. Collection managers can use this data to justify increasing acquisition budgets for these DDC ranges, while also identifying chronically underused subjects or languages where purchases might be reduced or weeding projects initiated.

5.2 Optimizing floor operations and space management

Beyond acquisitions, capturing spatial and temporal usage data directly improves the daily physical management of the library. Knowing that November and October are the peak months for reading-hall activity, and that Levels B2, F1, and E account for 75.5% of all in-library activity enables management to deploy staff strategically. Reshelving staff, reference librarians and security personnel can be scheduled and directed to high-traffic zones during peak months to prevent reshelving backlogs, protect the integrity of the open-stack classification system and improve the overall user experience.

5.3 A cost-effective, sustainable BI architecture

This project shows that public and academic libraries do not need to invest in costly 'smart shelving' infrastructure to achieve modern data analytics. Using existing RFID handhelds, a custom session-based mobile app, and a cloud-based EBLM architecture built on Google Apps Script and Looker Studio, the library has created a sophisticated, automated data pipeline at virtually no additional software costs. This lean approach offers libraries elsewhere a sustainable, replicable model for transforming raw operational logs into strategic business intelligence.

6. Conclusion

This practice report demonstrates that relying exclusively on traditional circulation metrics is no longer sufficient for understanding real user engagement, particularly in institutions that focus on reference and research. By integrating routine RFID sweeping workflows with a lean, automated Business Intelligence pipeline, using a custom mobile app, internal metadata enrichment, and Google Looker Studio, the Bibliotheca Alexandrina was able to illuminate the 'invisible' usage of its physical collections. This approach provides collection managers with detailed, evidence-based data to align acquisitions with actual patron demand and offers a cost-effective alternative to expensive smart-shelving infrastructure. The methodology and dashboard design can be transferred to other libraries facing similar constraints on circulation-based assessment.

Almoataz Mohamed Gad Elkarim, Bibliotheca Alexandrina, Egypt, https://orcid.org/0009-0007-9588-2905

Citable link (DOI): https://doi.org/10.5282/o-bib/6313

This work is licensed under Creative Commons Attribution 4.0 International.