

Books and Borrowing, 1730-1830: An Analysis of Scottish Borrowers' Registers

by

Brian Aitken, Katie Halsey and Matthew Sangster

with

Karen Baston, Maxine Branagh-Miscampbell, Alex Deans, Jacqueline Kennard Imrie, Gerard Lee
McKeever and Joshua J. Smith

Introduction

In this article, we introduce the research project 'Books and Borrowing, 1750-1830: An Analysis of Scottish Borrowers' Registers' (<https://borrowing.stir.ac.uk/>), with a particular emphasis on how the project went from conception to execution. We hope that this might be useful to future researchers in Digital Humanities.¹ The Books and Borrowing digital resource brings together the borrowing registers of eighteen historic Scottish libraries, and via the digitisation and transcription of these registers, and the analysis and interpretation of the data therein, provides a resource for researchers to establish which works were really circulating in the period 1750-1830, by whom they were borrowed, and when. It offers the opportunity to identify patterns and trends in book circulation, as well as to explore the borrowing habits of particular individuals, groups and socio-economic classes. The database is freely available to all.

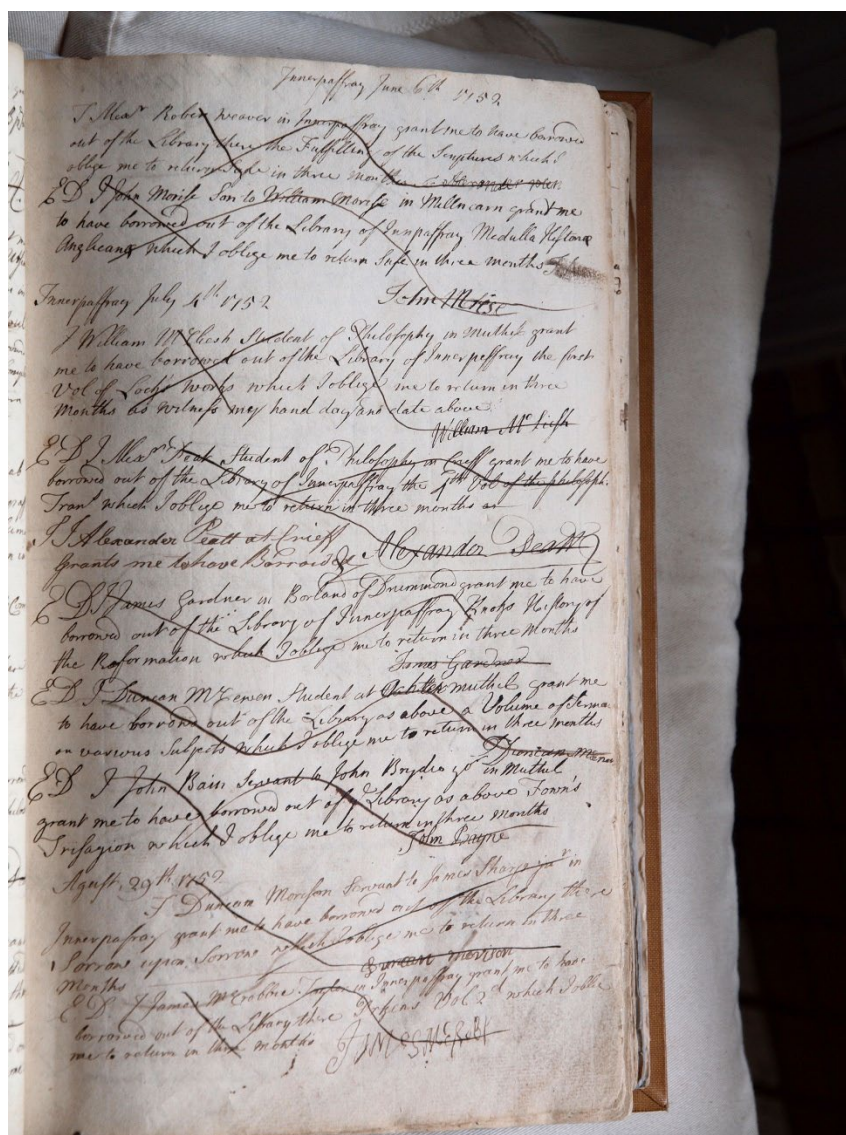
The project's initial key research question, as articulated to the funding body in our application, was: how can the untapped potential of Scottish borrowers' registers be employed to enrich our knowledge of the histories of books and reading? We sought to answer this main question via a series of further questions. These were:

- Which were the most popular forms, genres, titles and authors borrowed, and by which genders, occupations and socio-economic groups?
- How might borrowing data serve to challenge existing notions regarding the timing, scope, nature and influence of the Scottish Enlightenment?
- How might borrowing data serve to change our sense of the late eighteenth and early nineteenth centuries as the period of literary Romanticism?
- What can these registers tell us about shifts or continuities in the reading behaviour of particular communities, classes and individuals over time?
- To what extent do changing reading patterns encode the development of disciplinary and systematic knowledge over the course of the late eighteenth and early nineteenth centuries?
- What is distinctive about Scottish borrowing practices in local, national and international terms?
- In what ways can digital technologies serve to open up the quantitative data and qualitative experiences encoded in previously intransigent forms of manuscript evidence?

Because of the breadth and diversity of these questions, it was crucial that we should build a digital resource that was flexible enough to cope with them all. Our central aim from the beginning of the project, therefore, tied to our key research question, was to use existing digital technologies to make the valuable manuscript evidence (example pictured in *Figure 1*) accessible and searchable.

¹ Funded by the Arts and Humanities Research Council of the UK (Grant Number AH/T003960/1).

Figure 1: Innerpeffray Library, vol 1, page 5r.



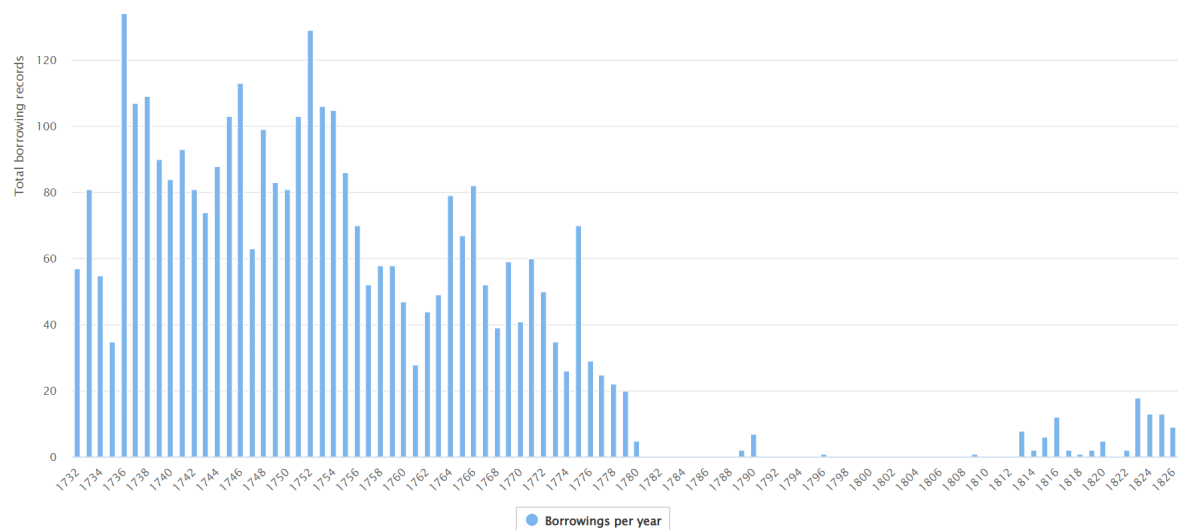
We knew from pilot projects that one of the most difficult things to do would be to create a system that could properly accommodate the complexity of the manuscript evidence, while remaining simple and intuitive to use. The ways in which we approached this challenge are detailed below, but we will now give a brief overview of the digital resource.

The Books and Borrowing online resource launched in April 2024. It contains 160 digitised ledgers of borrowing registers from eighteen libraries. Of these, 60 ledgers have been fully transcribed. The images of all ledger pages (more than 40,000 images) are all fully available in the digital resource, but the transcribed data, on which our statistical analysis is based, contains detailed information on 28,029 book holding records written by 12,835 authors, and 11,188 borrowers, of whom 95% were male, 4% female, and 1% unknown. The resource documents 161,905 separate borrowing events that took place between 1740 and 1840, making it, we believe, the largest corpus of data on historic book circulation in the UK to date.

The Browse and Search functions were designed to allow users to search and browse all information included in the resource, but, recognising that the quantity and complexity of the data might seem off-putting, we created a variety of different ways into the data. First, the Browse function allows users to

browse book editions, borrowers or libraries. The Browse Editions and Browse Borrowers functions allow users to browse books or borrowers across the whole dataset, including ways to limit alphabetically and by date. The Browse Libraries function is more complex. It takes users first to a list of all libraries in the database, including a map and bar chart. Users can then choose a library, at which point they are led to an individual home page for each library. These pages contain information about the library's history, location, holdings, and borrower demographics, as well as links to its catalogue and lists of sources relating to the library. The individual library home pages also include a bar graph that shows borrowings per year for the library in question, and an explanation of which registers are available for searching in our system, as demonstrated in the screenshot below, which shows the bar chart and explanation for Dumfries Presbytery Library.

Figure 2: Bar Chart and Explanation for Dumfries Presbytery Library
(<https://borrowing.stir.ac.uk/library/dumfries-presbytery>)

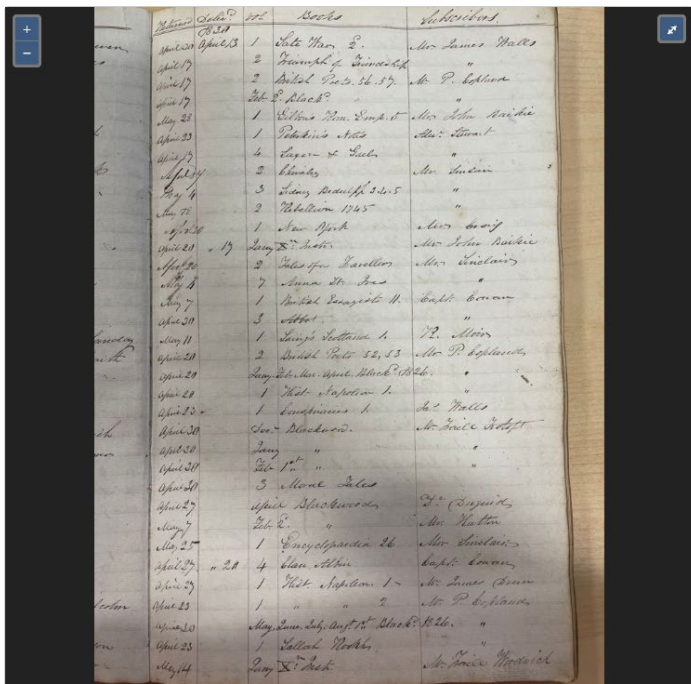


This bar chart shows total borrowings per year. The Dumfries Presbytery Library Issue Book is in one volume covering the period 1732-1826, and has been fully transcribed. Gaps in this bar chart therefore represent periods where no borrowings were recorded.

At the top of each individual library home page are five tabs, 'Introduction', Registers, 'Books', 'Borrowers' and 'Facts and Figures'. 'Books' and 'Borrowers' lead the user into browsable lists of books and borrowers for the library in question. 'Introduction' returns them to the individual library home page. 'Facts and Figures' presents a series of key facts and figures relating to the individual library. Beginning with a numerical summary, it then covers the most borrowed book titles, the most borrowed authors, the most borrowed genres, the most prolific borrowers, pie charts that show the breakdown of borrower occupations at the library and the breakdown of genres borrowed, and two bar graphs, showing borrowings through time by occupation type and by book genre. The 'Registers' tab allows the user to choose a register, and then view the digitised images, page by page, alongside our transcriptions and interpretations of the page in question. The screenshot below provides an example.

Figure 3: Page and Image View for Orkney Library Register C05-100-06 (1830-35) page 16

Tip: This view works best on a large screen.
[Download full image](#)



BORROWER

Mr James Walls

Gender: Male.

BOOK HOLDING

Genre: History, Genre: Lives

Memorials of the late war

Volumes borrowed: Volume 2

BOOK EDITION

Genre: History, Genre: Lives

Memorials of the late war

Language: English . Published: London | Edinburgh. Date of publication: 1828. Pagination: 2 vols.

Number of borrowings: Volumes associated with this edition were borrowed 29 times in 17 borrowing records

[Other authority](#)

BOOK WORK

Genre: History, Genre: Lives

Memorials of the Late War

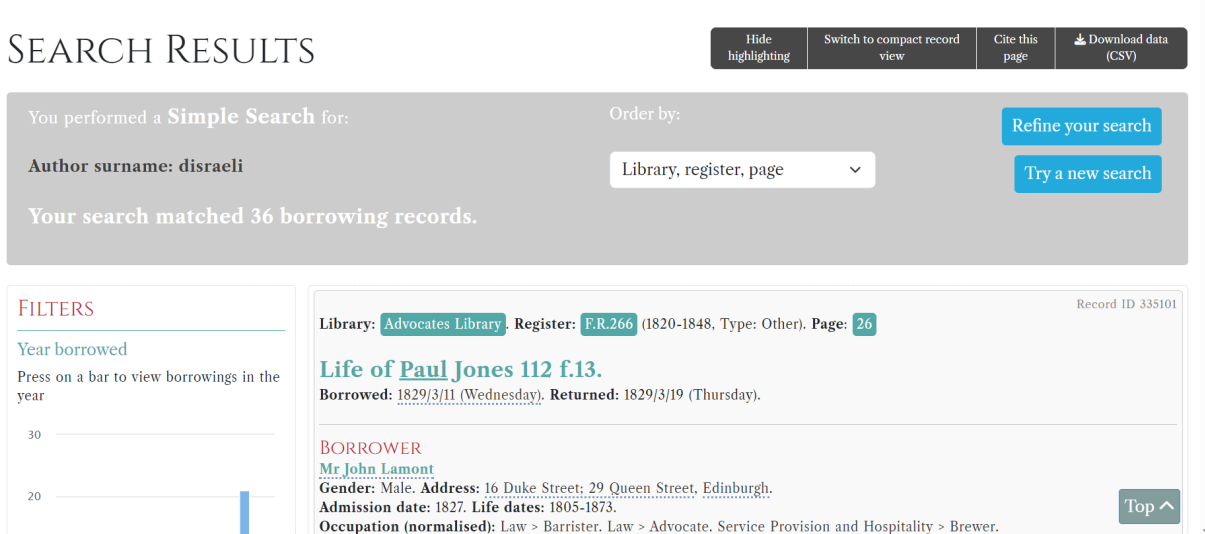
2 Triumph of Friendship

Borrowed: 1830/4/13 (Tuesday). Returned: 1830/4/17 (Saturday).

Record ID 423207
Top ^

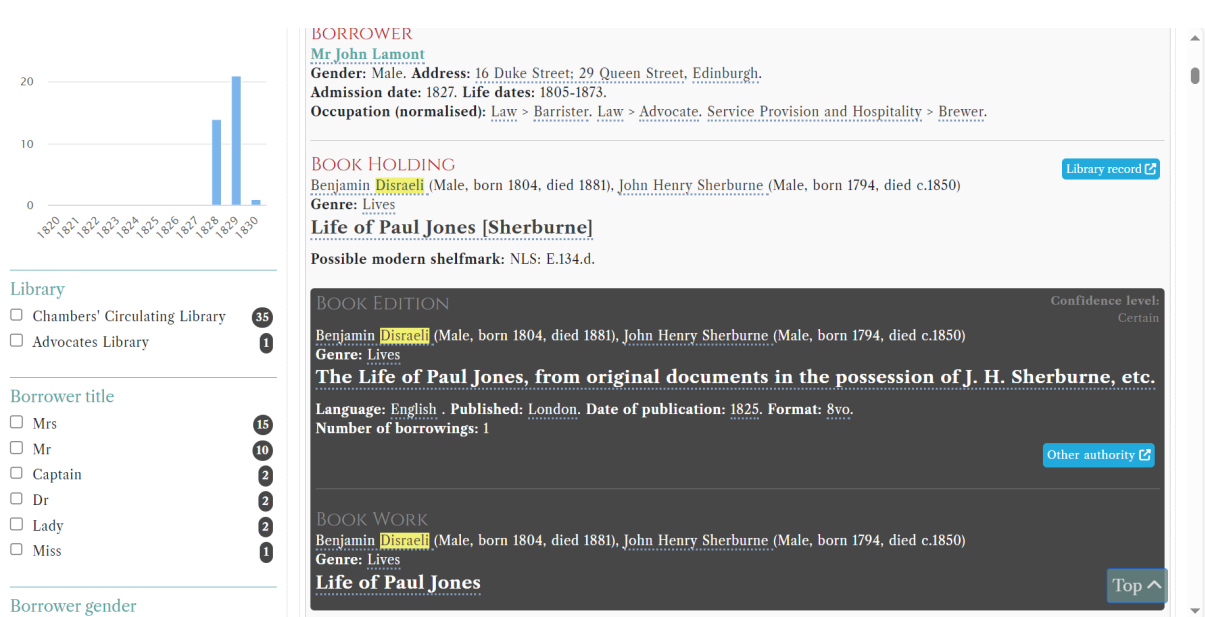
Secondly, there are three search options: a ‘Quick Search’ option, which offers a simple keyword search across all fields, an ‘Advanced Search’ which, in the ‘Simple Option’ presents users with what we consider likely to be the most popular options for searching (it is possible to delimit by library, author, book title, book genre, borrower name, and date(s) of borrowing), and finally a ‘Complete Option’, which allows users to search in extremely granular ways, including (but not limited to) by edition, format, language, date of publication of book; gender, occupation, address and title of borrower, and by particular holdings or editions of works held by individual libraries and across the entire dataset. Once a search has retrieved results, it is possible to filter further by a number of different categories: by year borrowed, by library, by borrower occupation, gender or title, by author, by *ESTC* number (where applicable), by book language, format, genre and place of publication. The screenshot below demonstrates the results retrieved by a search by author for ‘Disraeli’.

Figure 4: Search Results for ‘Disraeli’



The first borrowing to appear is the *Life of Paul Jones* (1825), borrowed from the Advocates Library on 11 March 1829 and returned on 19 March by Mr John Lamont, a barrister. The first filters (by year and by library) are visible in the left-hand side of the screen. Further details, and more options for filtering, appear lower down the screen, as demonstrated in the screenshot below.

Figure 6: Options for Filtering



If we delimit by borrower gender, choosing only female borrowers of Disraeli’s works, we discover that *Paul Jones* and the Advocates library disappear from the list of options, leaving only Chambers’ Circulating Library, and only two of Disraeli’s works, *The Voyage of Captain Popanilla* (1828) and *Vivian Grey* (1826-7), remain. Filtering in different ways (for example by borrower occupation) would allow for analysis of the data from various different angles.

Figure 7: Filtering by Gender

☐ Chambers' Circulating Library 18

Borrower title

☐ Mrs 15

☐ Lady 2

☐ Miss 1

Borrower gender

☒ Female 18

Borrower occupation

☐ Miscellaneous 5

☐ Wife/Spouse 5

Author

☐ Benjamin Disraeli 18

ESTC

Book language

☐ English 18

BOOK EDITION

Benjamin **Disraeli** (Male, born 1804, died 1881)
Genre: Fiction
The Voyage of Captain Popanilla.
Language: English . **Published:** London. **Date of publication:** 1828. **Format:** 12mo.
Number of borrowings: 11

Other authority

BOOK WORK

Benjamin **Disraeli** (Male, born 1804, died 1881)
Genre: Fiction
Voyage of Captain Popanilla

Library: Chambers' Circulating Library **Register:** Library Register (1827-1830, Type: Town). **Page:** 44

Record ID 284483

Popanilla
Borrowed: 1828/8/19 (Tuesday). **Returned:** 1828/9/18 (Thursday).

BORROWER

Mrs Boyce
Gender: Female. **Address:** 53 Castle Street, Edinburgh.
Subscription Type: New Books II Month. **Latitude:** 55.9523045. **Longitude:** -3.2034351. **Subscription Type (Map):** New

Day: 19. **Start Month:** 8. **Start Year:** 1828. **End Day:** 18. **End Month:** 9. **End Year:** 1828.

Top

The borrowing records in both the search results and the ‘browse register pages’ view feature a subtle dotted line underneath any item that is searchable, such as place of publication or book genre. Pressing on such items performs a search for the item in question, and this ‘click through’ feature makes it considerably easier to follow interesting pathways through the data: when finding something of interest, rather than having to remember this and navigate to the search page, a simple press of a button leads the user directly to the results they are interested in.

Across the whole site, but in particular on our Facts and Figures pages (which pertain both to individual libraries and the whole dataset), we have provided a number of visualisations relating to key information: tables, bar graphs, pie charts, treemaps and line graphs. In addition, we created an interactive map (of the Chambers Circulating Library borrowers) and an online exhibition (based on Edinburgh’s reading culture) to try to find different ways into the data, and to present our own research findings in ways that different audiences might find accessible.

The development of the resource

Developing the systems to effectively support a large and complex dataset requires planning and consideration; without sufficient planning the development process risks resulting in a system that does not meet the needs of the project. The Books and Borrowing project commenced in early 2020 and the online resource was launched in April 2024. The first few months of the project, before the team of researchers who would transcribe the many tens of thousands of borrowing records came onboard, was devoted to defining the data structures, documenting the requirements and building the collaborative content management system (CMS) that would enable the researchers to work efficiently and the project to reach its goals.

This preliminary phase of the project involved close collaboration between the project’s principal and co-investigators, Katie and Matt, and the project’s systems developer, Brian. At this stage we needed to consider the data structures that would be required to store not just the data, but also the many varied relationships between the different types of data. Even at this early stage we also needed to consider what the eventual uses of the data might be: why exactly we were intending to record particular types of data and how this data might be used in future.

The project was fortunate to have access to existing smaller datasets from earlier projects covering several of our target libraries, such as Glasgow University Library and Innerpeffer Library. These datasets, while by no means as extensive as data produced by Books and Borrowing, provided a useful starting point for mapping out the data structures the project would need. Each of the datasets had different structures and fields; for example the data for Innerpeffer Library featured fields for library fines and librarian names while the data for Glasgow University Library featured the names of each student's classes and professors. Using these existing datasets as a starting point, and in conjunction with Katie and Matt's knowledge of the types of data they wished the project to cover, we were able to map out a 'core' set of fields that would be common to all of our target libraries and an extensible 'additional' set of fields that could be created to meet the needs of individual libraries. The former would provide the basis for the project's searchable fields while the latter would be displayed in the records but would not be searchable.

Our understanding of the data structures was then recorded in a 'data description' document, which provided a name and description of every type of data, every field and every relationship between data types that the system would need to contain.² The document also featured an entity-relationship diagram that offered a visual representation of the data types and their relationships:

² <https://borrowing.stir.ac.uk/wp-content/uploads/2024/04/BandB-data-description-v2-1.pdf> . NB as the project developed, we decided that the genre needed to be associated with Works, Holdings and Items, not just Editions.

The diagram is an Entity-Relationship (ER) model for a library database. It consists of the following entities and their relationships:

- Entities:**
 - Page**: Connected to **Ledger** (1:N) and **Book Item** (1:N).
 - Ledger**: Connected to **Page** (N:1) and **Borrowing** (1:N).
 - Book Item**: Connected to **Page** (N:1) and **Borrowing** (M:N).
 - Borrowing**: Connected to **Book Item** (N:M), **Borrower** (M:N), and **Additional FieldData** (1:N).
 - Borrower**: Connected to **Borrowing** (N:M), **Occupation** (M:N), and **Additional Field** (1:N).
 - Occupation**: Connected to **Borrower** (N:M).
 - Additional FieldData**: Connected to **Borrowing** (N:1) and **Additional Field** (N:1).
 - Additional Field**: Connected to **Additional FieldData** (1:N).
 - Book Holding**: Connected to **Book Item** (1:N) and **Book Edition** (M:N).
 - Book Edition**: Connected to **Book Holding** (N:M) and **Book Work** (N:1).
 - Book Work**: Connected to **Book Edition** (1:N) and **Author** (M:N).
 - Author**: Connected to **Book Work** (N:M).
 - Available sources**: Connected to **Author** (N:1).
 - Borrower Xref**: Connected to **Borrower** (1:N) and **Library** (N:1).
 - Library**: Connected to **Borrower Xref** (1:N).
- Relationships:**
 - Page** to **Ledger**: 1:N
 - Ledger** to **Page**: N:1
 - Page** to **Book Item**: 1:N
 - Book Item** to **Page**: N:1
 - Book Item** to **Borrowing**: M:N
 - Borrowing** to **Book Item**: N:M
 - Borrowing** to **Borrower**: M:N
 - Borrower** to **Borrowing**: N:M
 - Borrowing** to **Additional FieldData**: 1:N
 - Additional FieldData** to **Borrowing**: N:1
 - Borrower** to **Additional Field**: 1:N
 - Additional Field** to **Borrower**: N:1
 - Additional FieldData** to **Additional Field**: N:1
 - Additional Field** to **Additional FieldData**: 1:N
 - Book Holding** to **Book Item**: 1:N
 - Book Item** to **Book Holding**: N:1
 - Book Holding** to **Book Edition**: M:N
 - Book Edition** to **Book Holding**: N:M
 - Book Edition** to **Book Work**: N:1
 - Book Work** to **Book Edition**: 1:N
 - Book Work** to **Author**: M:N
 - Author** to **Book Work**: N:M
 - Available sources** to **Author**: N:1
 - Author** to **Available sources**: 1:N
 - Borrower Xref** to **Borrower**: 1:N
 - Borrower** to **Borrower Xref**: N:1
 - Borrower Xref** to **Library**: N:1
 - Library** to **Borrower Xref**: 1:N
 - Borrower** to **Occupation**: M:N
 - Occupation** to **Borrower**: N:M

The data structure for book records is the most complex part of the system, and was developed in tandem with colleagues at the University of Liverpool's *C18th Libraries Online* project.³ A library contains any number of book holdings: these are records of the specific physical objects held by a library and may consist of one or more book volumes (labelled as 'items' in the above diagram). Individual volumes of a book, such as volume 2 of Adam Smith's 1776 *Wealth of Nations*, may be borrowed separately and the structure facilitates the recording of this. Different libraries may contain their own copies of the same book, and book edition records are the means through which these connections are tracked. An edition is a record of a specific printing of a book, such as the 1776 quarto edition of Adam Smith's *An Inquiry into the Nature and Causes of the Wealth of Nations* printed in London for William Strahan, Thomas Cadell and William Creech. One single book edition record is stored for each such printing, and the individual holding records of this particular edition at

³ <https://www.liverpool.ac.uk/history/research/research-projects/history-libraries/>

each library are then connected to the edition record. The final and most conceptual book record is the book work. This is a record of book as a conceptual object, such as Adam Smith's *Inquiry into the Nature and Causes of the Wealth of Nations*. There may be any number of different editions of a conceptual work, for example printed in different years, or by different publishers. This four-level hierarchical structure enables a wide range of book searches to be undertaken: a researcher who is broadly interested in *Wealth of Nations* can search at book work level, retrieving data relating to all individual editions, specific copies of these editions held at libraries and the individual volumes that were borrowed. Another researcher may be interested in a particular edition of this work and can limit their search accordingly. A further researcher may only be interested in the borrowings of the edition, or individual volumes of it, held at a specific library and the structure enables such a fine-grained search to be performed.

Any number of book authors and genres can be associated with any of the four levels of book record, and their association automatically cascades down from the level of association to the lower levels. For example, if an author is associated with a book work, then all connected editions, holdings and items will also feature this association. However, this association can be augmented or replaced at a lower level if necessary, for example if a particular edition of a book happens to include additional content by a different author.

We considered a hierarchical model for genre but decided this was unnecessarily complicated and instead selected 21 genres that could be applied to the books in our system. This classification system, including details of the subjects each genre covers, is available online.⁴ Our borrower occupation classification system is also available online.⁵ This hierarchical system allows any number of occupations to be assigned to a borrower (for example student and advocate) and enables searches to be performed at any hierarchical level (for example borrowings by all borrowers that have an 'Arts and Letters' occupation, or borrowings made specifically by authors).

Defining the data was the foundation on which everything else in the project was built. It was vital to have the structure mapped out, clearly defined, documented and agreed upon before anything else was developed, otherwise we risked developing a system that would not meet the needs of the project. However, it is important to bear in mind that while it is vital to map out the data structures prior to development, such structures cannot and should not be entirely set in stone at this stage. It is inevitable that as a project progresses everyone's understanding of the data and its potential uses will evolve and changes are likely to be proposed. A mechanism for discussing these changes and, if feasible, implementing them, must be incorporated into the development process. For example, our original data description document did not include author gender as a field, an oversight that was spotted once the systems had been developed. The project team agreed that it was essential that this field was incorporated into the data structure and be included in the search facilities and therefore the system was updated to reflect this.

The de-facto standard when transcribing historical texts and creating digital equivalents is TEI XML⁶ but during the planning phase of the project we decided against using this technology and instead chose to store the project's data in a relational database (the free and open-source MySQL system).⁷ The principal reason for this was because the project was primarily interested in extracting and querying the data contained within the borrowing records on the written pages rather than the creation of a digital representation of the page. As shown in the photograph of Innerpeffray Library's register above, the borrowing register pages can be very difficult to decipher, often featuring nonlinear layouts, difficult handwriting, extensive abbreviations and multiple crossings-out. The primary goal of

⁴ <https://borrowing.stir.ac.uk/wp-content/uploads/2023/12/Genre-Classifications.xlsx>

⁵ <https://borrowing.stir.ac.uk/wp-content/uploads/2023/12/Occupation-Classifications.xlsx>

⁶ For more information see <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/SG.html>

⁷ <https://www.mysql.com/>

the researchers was to extract dates, book titles, authors and borrowers and to be able to store this data and its relationships in a structured manner that would enable querying; documenting the exact structure of each page was a lesser concern, not least because of our decision to include the photographs of the registers which any user would be able to see. For this reason, a relational database was deemed to be a better fit for the project.

With the data structures and storage platform agreed upon, we could then set up the database on a server located at Stirling University. Once this was in place we could then import the existing sample datasets from around 12 libraries. As previously mentioned, the sample data differed markedly between libraries in terms of structure and format, and bespoke import scripts needed to be developed and tested that would extract the data and map it onto our database system. Once this process had been successfully completed, we had the beginnings of our dataset housed in an online database, but no interface through which a team of researchers could access, add or edit the data. The next stage was to design and develop the content management system through which the researchers could complete these tasks.

The CMS is an online tool that the researchers can log into and manage the project's data. It connects to the underlying MySQL database and enables the data to be managed via a series of easy to use web-based forms. As with the process of data definition, the design of the CMS was shaped by regular meetings involving Katie, Matt and Brian. It is essential when developing such a tool that the people who will be using it and who know the intricacies of the data are involved in the design process otherwise the resulting software risks being unsuitable; jumping straight into the development process with insufficient consultation will almost certainly lead to inefficiencies and confusion.

The meetings led to the formulation of a set of requirements for the CMS, which were documented by Brian in a requirements document that detailed all of the features the system would need to incorporate.⁸ This document was written in non-technical language and presented a high-level summary of the desired features, describing what needed to be developed rather than going into detail about exactly how it would be developed. This ensured that Katie and Matt could easily understand the document, and, following a series of iterations of writing, sharing, meeting and updating, we finalised the plans for the CMS. It should however be noted that the requirements for the CMS were not permanently fixed at this stage; it is inevitable that as a system is used potential improvements or omissions will be noticed and this eventuality should be factored into the development process.

The CMS was developed using the PHP scripting language⁹, the Bootstrap user interface framework¹⁰ and the jQuery JavaScript library.¹¹ The resulting system enables the project researchers to add, edit and delete records for libraries, registers, pages, borrowings, borrowers, books and authors through a series of forms that are easy to use and minimise the risk of duplication of data. Records such as books, authors and borrowers only need to be added into the system once and can then be linked to any number of borrowing records. For example, adding an existing borrower to a new borrowing record is a straightforward process involving typing a few letters of the borrower's name, which then brings up a list of possible matching borrowers. The researcher then simply needs to select the appropriate borrower from the list to make the association.

The following screenshot features a section of the CMS, displaying the borrowing records found on a specific library register page at Chambers' Circulating Library. The data associated with each borrowing record is displayed, featuring information about the borrowing, the borrower, the book holding, including which specific volumes were borrowed, the associated book edition and work, plus

⁸ <https://borrowing.stir.ac.uk/wp-content/uploads/2024/04/BandB-CMS-requirements-v2-1.pdf>

⁹ <https://www.php.net/>

¹⁰ <https://getbootstrap.com/>

¹¹ <https://jquery.com/>

the author. Using the icons to the left of the record it is possible to edit any of the information or (if necessary) delete the record. Deleting a record makes it inactive in the system but the record continues to be stored in the underlying database and can be reinstated if required. The buttons above the borrowing records enable the researcher to navigate between pages in the register, view the digitised image of the page and add a new record to the page.

Figure 9: Section of the Content Management System

BOOKS AND BORROWING CMS

Home
Add Library
Browse Libraries
Add Book
Browse Books
Add Author
Browse Authors
Add Genre
Browse Genres
Add Occupation

Browse Occupations

← Return to Library

CHAMBERS' CIRCULATING LIBRARY, LIBRARY REGISTER, PAGE 16 (ID 79395)

← Previous (15)
Page Image
+ New Borrowing Record
Next (17) >
Show editors' notes

Order	Borrowing	Borrower	Book Holding / Items	Book Edition / Works	Authors
✎ 🗑	1 Columbus-1-2-3-4 Borrowed: 1828-6-12 (Thursday) Returned: 1828-8-1 (Friday)	Mr Robert Cameron Borrower type: Main Gender: Male Settlement: Edinburgh Street: 7 St Vincent Street Normalised occupation: Accountant Subscription Type: New Books III, Half-Year 0, 20, 0. Latitude: 55.9579264 Longitude: -3.2033956 Subscription Type (Map): New Start Day: 12 Start Month: 6 Start Year: 1829 End Day: 11 End Month: 12 End Year: 1829	HOLDING Standardised title: History of the Life and Voyages of Christopher Columbus Catalogue Number: 849 ITEMS Part number: 1 Part number: 2 Part number: 3 Part number: 4	EDITION Confidence level: Very likely Title: A history of the life and voyages of Christopher Columbus / by Washington Irving. Other authority: http://explore.bl.uk/BLVU1:1SCOP-ALL:8LL01001821049 Language: English Place of publication: London Date of publication: 1828 Format: 8vo Pagination: 4 vols WORK Title: History of the Life and Voyages of Christopher Columbus	WORK Name: Washington Irving 1783 1859
✎ 🗑	2 Croppy-v2 Borrowed: 1828-6-12 (Thursday) Returned: 1828-8-1 (Friday)	Mr Robert Cameron Borrower type: Main Gender: Male Settlement: Edinburgh Street: 7 St Vincent Street Normalised occupation: Accountant Subscription Type: New Books III, Half-Year 0, 20, 0. Latitude: 55.9579264 Longitude: -3.2033956 Subscription Type (Map): New Start Day: 12	HOLDING Standardised title: Croppy: a Tale of the Irish Rebellion. Catalogue Number: 864 ITEMS Part number: 3	EDITION Confidence level: Very likely Title: The Croppy; a Tale 1798. Other authority: http://explore.bl.uk/BLVU1:1SCOP-ALL:8LL01000824952 Language: English Place of publication: London Date of publication: 1828 Format: 12mo Pagination: 3 vols WORK Title: Croppy; A Tale 1798	WORK Name: Michael Banim 1796 1874

The development of the CMS was completed before the team of researchers began working for the project in June 2020 and was successfully used by more than 15 researchers over the course of the project.

With the research team in place and the CMS fully functioning, the role of the developer was greatly reduced for the following two years and was primarily limited to fixing bugs, implementing requested feature updates and managing the process of preparing more than 40,000 library register page images for inclusion in the resource. This task involved liaising with several digitisation units, writing scripts to rename image files to fit the project's naming scheme, uploading the images to the project's server and writing scripts to batch create page records for the page images in the project's database.

The project's large collection of images is stored in a IIIF (International Image Interoperability Framework) server.¹² The server enables one high resolution image per register page to be stored,

¹² <https://iiif.io/>

with the server generating any required derivatives on demand. This is considerably more efficient than generating and storing all potentially required derivatives and facilitates the ‘Google Maps’ style zoom and pan interface to the images that the front-end offers.

It was not until 2022 that preparations began for the development of the public-facing front-end to the data. As with the phases of data definition and the development of the CMS, this process began with the creation of a requirements document, the creation of which was a collaborative effort involving the whole project team. The team would meet to discuss the features they would like the website to offer and Brian would make notes, write these up and disseminate them. The team would then meet again and discuss these requirements and the process would then repeat until all aspects of the website had been covered and a concrete set of requirements had been documented. This iterative process gave everyone time to consider the proposals and ensured everyone had a chance to make suggestions about what the front-end would feature. As with the CMS, the requirements were written in non-technical language and at a fairly high level: a description of what features would be included and how they would function rather than specific details of how they would be developed. A final version of the requirements document was completed in March 2022, covering all the features that are now available on the project website, although it should be noted that the requirements were not entirely fixed and immovable at this point and further additions and refinements were made as development proceeded.¹³

A fundamental design decision that was made at this stage was to develop both a front-end and an API (Application Programming Interface) in parallel. An API is a “set of rules or protocols that enables software applications to communicate with each other to exchange data”.¹⁴ It sits between a data source and other applications and provides clearly defined pathways to datasets. These are called ‘endpoints’ and take the form of URLs and an application can connect to an API endpoint to request data in a specific format, such as JSON or CSV.

There are numerous advantages to building an API as opposed to incorporating all of the required data querying functionality directly into the front-end itself. For example, other applications (or indeed humans) can connect directly to the API and reuse the data for other purposes, such as incorporating borrowing records into a new research project for different purposes. It also means the dataset is not intertwined with the current front-end and developing a new front-end in future would be much more straightforward. In addition, data can easily be downloaded in different formats (for example CSV for use in Excel) to the format the front-end is set up to work with (in our case JSON). The API developed for the project is publicly accessible and fully documented, and the front-end connects to this any time data is required; there are no parts of the front-end that bypass this API or use undocumented methods to retrieve and display data.¹⁵

As with the CMS, both the API and the front-end were developed using the PHP scripting language, with the front-end also using Bootstrap and jQuery. In addition, interactive maps in the front-end were developed using the Leaflet JavaScript mapping library¹⁶ and visualisations were developed using the HighCharts JavaScript library.¹⁷

During the development phase Brian would focus on one specific feature in the requirements document, such as ‘browse a chosen library’s borrowers’. He would develop the API endpoints the feature would require in order to access the relevant data, for example an endpoint to return counts of borrowers at a specified library with surnames beginning with each letter of the alphabet, and another

¹³ <https://borrowing.stir.ac.uk/wp-content/uploads/2024/04/BandB-front-end-requirements-v3.pdf>

¹⁴ <https://www.ibm.com/topics/api>

¹⁵ <https://borrowing.stir.ac.uk/api/v1/>

¹⁶ <https://leafletjs.com/>

¹⁷ <https://www.highcharts.com/>

endpoint to return borrower details for all borrowers at a library with surnames beginning with a specified letter.

With the endpoints in place Brian would then develop an initial version of the feature in the front-end that would connect to the API and then retrieve, format and display the data. Once this initial version of the feature was complete Brian would then share the URL with the rest of the project team, enabling them to test it and give feedback, which Brian would then incorporate and then share. This process would be repeated until everyone was satisfied with the feature, at which point this would be ticked off in the requirements document and Brian would move on to the development of another feature. This iterative workflow was followed until all features detailed in the requirements document had been developed and a fully functioning first version of the website was complete.

As the development of the search facilities proceeded it became apparent that while our relational database structure was perfect for modelling and storing the project's data and its relationships, the complexity of the data structure coupled with the volume of data meant that some of the queries we needed to perform were taking far too long to execute, and some kind of data caching method would be required. It also became apparent that the search results returned could be overwhelming for users. A search for a specific author or book title, for example, may bring back thousands of borrowing records and in the early iterations of the front-end's search facility all a user could do with the results was navigate through them, with 100 results listed per page. We realised that some form of results filtering would be highly beneficial and began investigating solutions that could cache the data to make searches more efficient while offering facilities to filter the results.

The project decided to adopt the Apache Solr search platform to meet these needs.¹⁸ Solr can index a dataset and provides lightning-fast search facilities, and it also offers faceted searching. When configured, faceted searching returns not just the search results, but a breakdown of results by specified criteria, such as borrower gender as shown in *Figure 7* above. This then enables the user to filter their results, for example, as shown above, by selecting results for only female borrowers, a feature that is very helpful for users of the front-end. Implementing the Solr-powered search facilities was the last major development task of the project and its inclusion greatly enhanced the search options we were able to offer to users.

Conclusion

The online resource has been very well received since its launch in April 2024. For future projects intending to develop an online resource featuring a comparably large and complex dataset, there are a number of lessons that we have learned. Producing and sharing documentation is of particular importance as over the course of a multi-year project it is inevitable that if details are not written down something will be forgotten. To this end, the project produced data definition and requirements documents; their creation was a collaborative effort, and the team could reference these documents throughout the project.

Developing systems for such a project must be a collaboration between the developer and the rest of the project team and benefits greatly from being an iterative process: it is important to have meetings to discuss the requirements, document and share them, meet again to discuss them and repeat the process until all features are covered and there is an agreement on what needs to be produced. When it comes to the development process the team should be given access to early versions of features so they can test them out and give feedback. If the developer works in seclusion for six months and only unveils a supposedly finished system to the team after this period it is likely that the system will at best need major reworking and at worst will be entirely unsuitable.

¹⁸ <https://solr.apache.org/>

The developer must be prepared to collaborate closely with the project team and should ideally have some understanding of and interest in the source material, but it is equally important that the project team involves the developer throughout the project. If the developer is kept at arm's length and (for example) isn't involved in project meetings then their understanding of both the data and the processes the team will be following will be restricted; being part of the team makes it more likely that the developer will gain a better understanding of the project and its goals and will be more likely to be able to produce the system that the team actually needs.

To give a very quick snapshot of how the resource has so far been used: many users have been most interested in our headline 'top ten' lists, a number of which fundamentally disrupt existing notions of canonicity and literary influence. To give just one example, not one of the Big Six Romantic poets (Keats, Shelley, Byron, Blake, Wordsworth or Coleridge) appears in the twenty most-borrowed books of the Romantic period. (Top of the list is Sir Walter Scott; this is not a surprise, but a valuable reminder of quite how heavily Scott dominated the cultural imagination of our period.) There are, of course, potentially a number of different reasons for the absence of the Big Six; these are explored in the monograph arising from the project, *Books and Borrowing in Scotland 1750-1830* which will come out with Edinburgh University Press in 2026. Other researchers have been interested primarily in the reception of particular authors, texts or genres, at different times and across the entire period, while some very interesting new research is currently underway on the relationships between class mobility and library access in the early nineteenth century, using our dataset. Colleagues involved in Masters teaching in both Digital Humanities and Book History have made use of our database for a variety of purposes, and for undergraduate teaching it has proved helpful in exploring the reception of particular authors in relation to their peers, and in foregrounding the questions of canon-formation, canonicity and literary influence touched on above. For those closely involved in the project, it has categorically shifted the ways we think about our period.

For those working in later periods, we hope the resource might also prove useful, if only in providing some solid statistical evidence relating to the books that were really circulating in the period when individuals who came to maturity in the Victorian age were in their formative years, and hence those that were likely to have shaped their minds and thinking. A Victorian literary culture shaped by the works of Scott, Charles Rollin, William Robertson, David Hume, Oliver Goldsmith, William Fordyce Mavor, Tobias Smollett, Maria Edgeworth, Plutarch and the Comte de Buffon (the most-borrowed authors of the period 1780-1840) looks rather different to one shaped by Wordsworth, Coleridge, Shelley and Keats, for example. Conversely, the current canonicity of Keats, Shelley, Wordsworth *et al* is owed almost entirely to the ways in which they were interpreted and understood in the Victorian age; this fact might shed interesting new light on how particular constructions of poetic identity were formulated in this period.

Our hope is that in this brief essay we have given some insight into the development, workings and potential uses of the *Books and Borrowing* digital resource, and we would very much welcome questions and feedback from those working on other comparable Digital Humanities projects.