

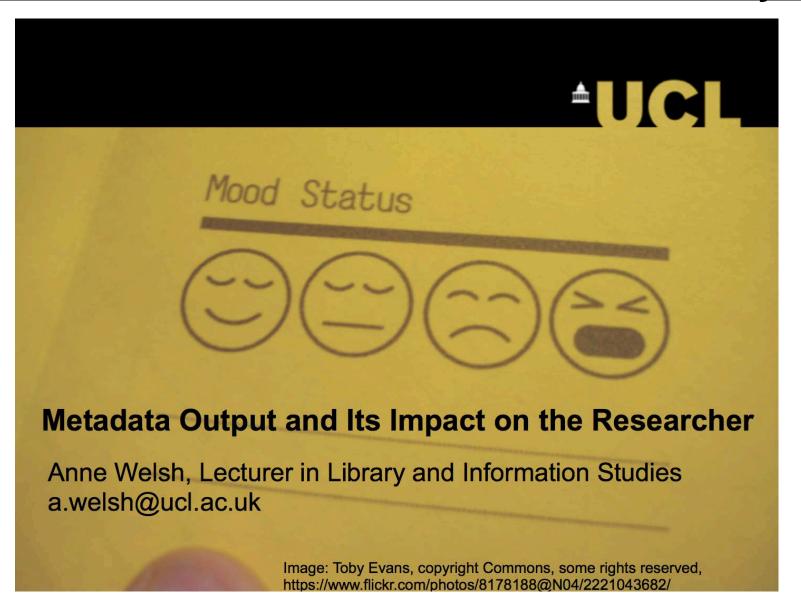
The Rare Books Catalogue as the Foundation of the **Scholarly Database**

Anne Welsh – a.welsh@ucl.ac.uk Lecturer in Library & Information Studies, University College London

IFLA Rare Books and Special Collections Section Midterm Lisbon, 22 February 2016



Preamble 1: At a Conference in Canterbury...





Preamble 1: At a Conference in Canterbury...





Preamble 1: At a Conference in Canterbury...



Published (Open Access) as:

Welsh, Anne. "Metadata Output and the Researcher." Catalogue and Index 178 (2015):

2-8,

http://www.cilip.org.uk/sites/default/files/documents/Catalogue%20and%20Index%20issue%20178,%20March%202015.pdf

s reserved, 13682/



Preamble 2: On an Exhibition Wall in Lisbon ...

... quando uma coleção se mantém una, a utilidade do catálogo é óbvia; quando a coleção se dispersa, o catálogo serve muitas vezes para confirmar a autenticidade de uma pintura, acrescentar-lhe valor imaginário e atribuir-lhe uma história.

Horace Walpole

... while a collection remains entire, the use of the catalogue is obvious; when dispersed, it often serves to authenticate a picture, adds to its imaginary value, and bestows a history on it.

Horace Walpole

Aedes Walpolinae. Rev. ed. 1752: xxxi, quoted in Paulino, Carla with Santos, Ana Teresa and Antunes, Ines. Calouste S. Gulbenkian and English Taste. Lisbon: Foundation Calouste S. Gulbenkian, 2015.

Writers' Libraries

"A set of books or other printed works owned by the author at a particular moment in time ... Writers' libraries in the possession of institutions are often (although not always) a collection of their books at the time of their death, or a subset thereof."

- Richard W. Oram. *Collecting, Curating, and Researching Writers' Libraries: A Handbook.* 2014: 1-2.

Image: National Portrait Gallery, copyright commons, http://www.npg.org.uk/collections/search/portrait/mw195731/



LORD BULWER LYTTON.

LORD MACAULEY T. CARLYLE

CHARLES DICKENS

GEO MAC-DONALD AJ.FROUDE WILKIE COLLINS ANTY TROLLOPE.



CATALOGUE OF THE LIBRARY OF THOMAS JEFFERSON



COMPILED WITH ANNOTATIONS BY
E. MILLICENT SOWERBY

Volume I

THE LIBRARY OF CONGRESS

WASHINGTON, 1952

Image: Library of Congress, https://www.loc.gov/rr/rarebook/coll/130.html

Writers' Libraries

"Most of us would agree that a catalogue of a private library can serve as a profile of a reader, even though we don't read all the books we own and we do read many books that we never purchase. To scan the catalogue of the library in Monticello is to inspect the furnishings of Jefferson's mind."

- Robert Darnton. *The Kiss of Lamourette: Reflections in Cultural History*. London: Faber, 1990: 162.



Writers' Libraries

"Within the ecosystem of a newly fertile subject, the biography is usually the first green growth to appear; then a bibliography grows up within its shade; several boldly broad critical surveys follow; and soon ... a grove of increasingly specialized studies takes root, affording productive habitats for modish critical approaches. Competing biographies and new editions of the works eventually flourish among the dense woodland vegetation of the climax stage ... Somewhere in the evolution of this delicate ecosystem of academic books, a study of the author's knowledge of others' writings - an examination of his or her library and reading – manages to thrust itself through the foliage of this timber into the sunlight."

- Alan Gribben. "Private Libraries of American Authors: Dispersal, Custody and Description." *The Journal of Library History* 21.2 (1986): 301.



Writers' Libraries

"Within the ecosystem of a newly fertile subject, the biography is usually the first green growth to appear; then a bibliography grows up within its shade; several boldly broad critical surveys follow; and soon ... a grove of increasingly specialized studies takes root, affording productive habitats for modish critical approaches. Competing biographies and new editions of the works eventually flourish among the dense woodland vegetation of the climax stage ... Somewhere in the evolution of this delicate ecosystem of academic books, a study of the author's knowledge of others' writings - an examination of his or her library and reading – manages to thrust itself through the foliage of this timber into the sunlight."

- Alan Gribben. "Private Libraries of American Authors: Dispersal, Custody and Description." *The Journal of Library History* 21.2 (1986): 301.



Cataloguing Writers' Libraries

"The mind of Man is recorded in his books, and the catalogues of the great libraries enable the individual to consult the universal mind on any limited subject that happens to interest him. The library collected by one man ... expresses only his own mind and special interests ... Usually, indeed, no material has existed for such attempts, unless it were an auctioneer's catalogue ...

- September 1939, Geoffrey Keynes."
 - Keynes. The Library of Edward Gibbon: A Catalogue. 2nd ed. 1980: 3.



Cataloguing Writers' Libraries

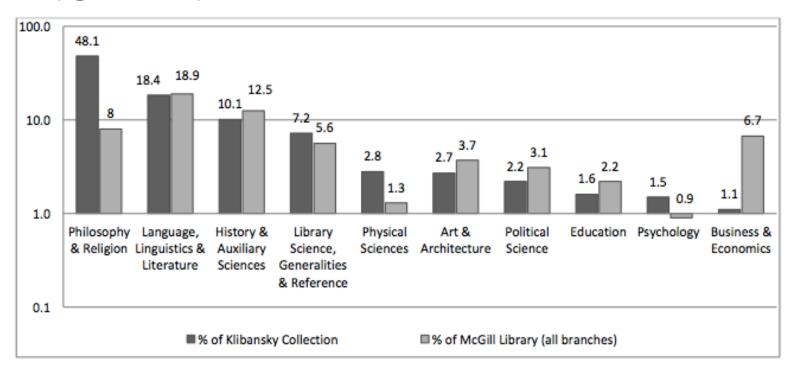
"The mind of Man is recorded in his books, and the catalogues of the great libraries enable the individual to consult the universal mind on any limited subject that happens to interest him. The library collected by one man ... expresses only his own mind and special interests ... Usually, indeed, no material has existed for such attempts, unless it were an auctioneer's catalogue ...

- September 1939, Geoffrey Keynes."
 - Keynes. The Library of Edward Gibbon: A Catalogue. 2nd ed. 1980: 3.



Quantitative Analysis from LMS Output

Figure 12. Top 10 RKC Broad Subject Areas as a percentage of the collection (logarithmic scale)



- Jillian Tomm. *The Imprint of the Scholar: an Analysis of the Printed Books of McGill University's Raymond Klibansky Collection: thesis.* Montreal: McGill University, 2012: 112.



Quantitative Analysis from LMS Output (with critical judgment)

"As we might expect from a man self-educated beyond secondary-shool level, we can see a mixture of university and popular presses. After establishing that most of the Oxford University Press entries on the catalogue actually represent offprints from the same journal, we can see that Dent (the publisher of the Everyman series) forms the largest proportion."



Quantitative Analysis from LMS Output

```
Record 1 of 937LOCATIONS SENATE HOUSE LIBRARYAUTHOR
                                                    Shakespeare, William, 1564–1616.TITLE
                                                                                                Complete works.IMPRINT
                                                                                                                              [Edinburgh? : Printed by
autograph comments in the
margins. Inscribed inside front
                                         cover: 'Christmas 1864. Presented to Thomas Bayly Splatt. Ist
                                                                                                                  Class, Highest Marks, Principal Dr
Quicke.'NOTE
                  ULL copy bound in black cloth, with gilt edges. UK-LoURL.NOTE ULL copy comprises vol. 2 only and lacks preliminaries. UK-LoURL.ADD AUTHOR
De la Mare, Walter, 1873–1956, former owner. UK–LoURL.1 > SHL SPECIAL COLLECTIONS WdlM
                  1873-1956, former owner. UK-LoURL.1 > SHL SPECIAL COLLECTIONS WdlM AVAILABLERecord 2
Adam international review : our Dylan Thomas memorial number / [edited by Miron Grindea].IMPRINT
                                                                                                         AVAILABLERecord 2 of 937LOCATIONS
                                                                                                                                             SENATE HOUSE
LIBRARYTITLE
                                                                                                                             London : Adam International
Review, 1953,DESCRIPT. 80p. : ill. ; 23 cm.SERIES Adam international review : 21,no. 238.NOTE ULL copy is from the working library of Walter de la
                     LOURL.SUBJECT Thomas, Dylan, 1914-1953.ADD AUTHOR De la Mare, Walter, 1873-1956, former owner. UK-LoURL.ADD TITLE Our Dylan Thomas
Mare.UK-
memorial number.1 > SHL SPECIAL COLLECTIONS [WdlM] 1
                                                                   AVAILABLERecord 3 of 937LOCATIONS SENATE HOUSE LIBRARYAUTHOR Andrade, E. N. da C.
(Edward Neville da Costa). 1887-1971.TITLE
                                              Isaac Newton.IMPRINT
                                                                   London : M. Parrish. 1950.DESCRIPT. 111 p. : ill., ports : 19 cm.SERIES
                                                                                          inscribed by the author on the front free endpaper. UK-LoURL.NOTE
portraits.NOTE
                    ULL copy is from the working library of Walter de la Mare,
ULL copy has dust wrapper. UkLU.SUBJECT Newton, Isaac, Sir, 1642-1727.SUBJECT
                                                                               Physicists -- Great Britain -- Biography.ADD AUTHOR De la Mare, Walter,
1873-1956, former owner. UK-LoURL.1 > SHL SPECIAL COLLECTIONS [WdlM] 10
                                                                                        AVAILABLERecord 4 of 937LOCATIONS SENATE HOUSE LIBRARYAUTHOR
                A Naturalist and immortality / by C de B.DESCRIPT. 22 p.; 23 cm.NOTE ULL copy is from the working library of Walter de la Mare,
inscribed by Williams Repton on the front free endpaper. UK- LoURL.NOTE
                                                                                       ULL copy has dust wrapper.. UkLU.SUBJECT Hudson, W. H. (William
Henry), 1841–1922.ADD AUTHOR De la Mare, Walter, 1873–1956, former owner. UK–LoURL.1 > SHL SPECIAL COLLECTIONS [WdlM] 100
                                                                                                                                        AVAILABLERecord 5 of
937LOCATIONS SENATE HOUSE LIBRARYTITLE The old clay patch : a collection of verses written in and around Victoria University College,
                                          de la Mare and S. Eichelbaum.EDITION
                                                                                     2nd ed.IMPRINT Auckland: Whitcombe & Tombs, 1920.DESCRIPT. xi, 153
Wellington, N.Z. / edited by F.A.
             First ed.: Wellington, N.Z.: Whitcombe & Tombs, 1910.NOTE Includes index.NOTE
                                                                                                       Bagnall, D228.NOTE ULL copy is from the working
library of Walter de la Mare. inscribed by F. de la Mare on the front free endpaper. UK
                                                                                                             -LoURL.SUBJECT Victoria College (Wellington,
                        New Zealand poetry -- 20th century.SUBJECT College verse, New Zealand -- 20th century.ADD AUTHOR De La Mare, F. A. (Frederick
N.Z.) -- Poetry.SUBJECT
Archibald)ADD AUTHOR - Eichelbaum, S.ADD AUTHOR - De la Mare, Walter, 1873–1956, former owner. UK-LoURL.1 > SHL SPECIAL COLLECTIONS [WdlM] 101
AVAILABLERecord 6 of 937LOCATIONS SENATE HOUSE LIBRARYTITLE Poems / by four authors.IMPRINT Cambridge : Bowes & Bowes, 1923.DESCRIPT. 119 p. ; 20
        Poems by J.R. Ackerley, A.Y. Campbell, Edward Davison and Frank
                                                                                                            "This volume comprises four independent books of
                                                                                         Kendon.NOTE
                            authors did not consider suitable for separate publication. ...
                                                                                                     Most of the poems included have already appeared in The
verse which the
                    mercury, The new statesman, The outlook, The Westminster
                                                                                      gazette, The spectator, The Saturday review, Form, The
                                                                                                                                     inscribed to
challenge, The Cambridge review and The eagle." -- Note, p. 5.NOTE
                                                               ULL copy is from the working library of Walter de la Mare and is
him from Edward Davison, Feb. 1923. UK-LoURL.SUBJECT 💮 English poetry -- 20th century.ADD AUTHOR Davison, Edward Lewis, 1898-, inscriber. UK-LoURL.ADD AUTHOR De
la Mare, Walter, 1873-1956, former owner. UK-LoURL.1 > SHL SPECIAL COLLECTIONS [WdlM] 102
                                                                                                       AVAILABLERecord 7 of 937LOCATIONS
                                                                                                                                          SENATE HOUSE
                  Coleridge, Mary E. (Mary Elizabeth), 1861-1907.TITLE Poems / by Mary E. Coleridge.EDITION
                                                                                                                  8th ed.IMPRINT
                                                                                                                                     London : E. Mathews &
Marrot, 1927.DESCRIPT. xxi, 214 p. ; 18 cm.NOTE ULL copy is from the working library of Walter de la Mare. UK- Lourl.ADD AUTHOR De la Mare,
                                                                               AVAILABLERecord 8 of 937LOCATIONS SENATE HOUSE LIBRARYAUTHOR
Walter, 1873–1956, former owner. UK–LoURL.1 > SHL SPECIAL COLLECTIONS [WdlM] 103
Williamson, Henry, 1895–1977.TITLE The gold falcon, or, The haggard of love : being the adventures
                                                                                                                 of Manfred, airman and poet of the World War,
and later husband and father, in search of freedom and personal sunrise in the city of New York and of the consummation of his life.IMPRINT
[England] : The author, 1933 (Plymouth : Latimer Trend)DESCRIPT. 415 p. ; 21 cm.NOTE Published anonymously. By Henry Williamson. Cf. BLC.NOTE
copy is a presentation copy to Richard de la Mare, inscribed beneath the printed dedication to him. UK-LoURL.ADD AUTHOR De la Mare, Richard, former owner. UK-LoURL.1 > SHL SPECIAL COLLECTIONS [WdlM] 104 AVAILABLERecord 9 of 937LOCATIONS SENATE HOUSE LIBRARYAUTHOR Barnes, William,
                    Poems of rural life in the Dorset dialect / by William Barnes.IMPRINT
                                                                                           London: C. Kegan Paul, 1879.DESCRIPT. xi, 467 p.; 20 cm.NOTE
ULL copy is from the working library of Walter de la Mare. Pencilled inscription on half-title: S. Spencer Smith. UK —LOURL.NOTE
ULL copy is bound in green morocco with gilt border. Top edge gilt. UK—LOURL.NOTE ULL copy has pencilled crosses beside some titles. Press
```

of poem, 'Roben's coorten', from the Dorset chronicle (?1880) pasted on verso of first divisional half-title, with pencilled

cutting



Quantitative Analysis from LMS Output (with manual editing)

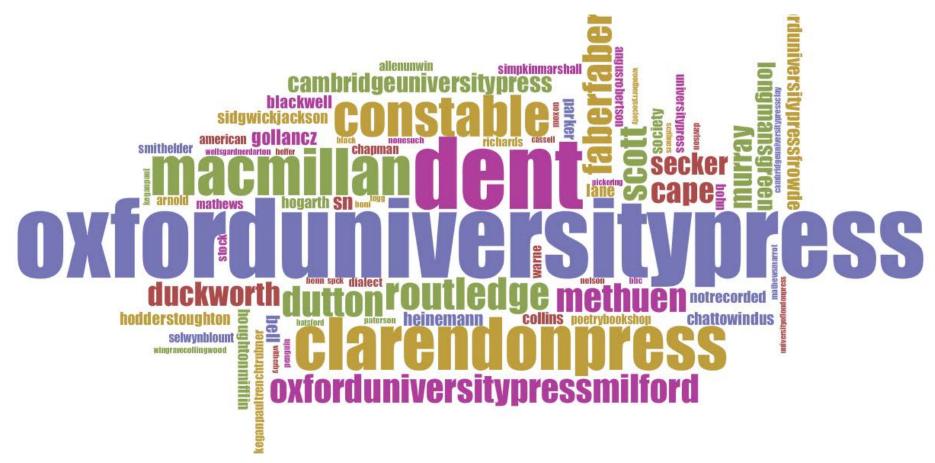
```
Record 1 of 937LOCATIONS SENATE HOUSE LIBRARYAUTHOR Shakespeare,
William, 1564-1616.TITLE Complete works.IMPRINT [Edinburgh?:
Printed by Ballantyne, 186-?].DESCRIPT. pp.3-919: ill.; 18 cm.NOTE

ULL copy is from the library of Walter de la Mare with his
autograph comments in the margins. Inscribed inside front cover:
'Christmas 1864. Presented to Thomas Bayly Splatt. Ist Class,
Highest Marks. Principal Dr Quicke.'NOTE ULL copy bound in black cloth,
with gilt edges. UK-Lourl.NOTE ULL copy comprises vol. 2 only and lacks
preliminaries. UK-Lourl.ADD AUTHOR De la Mare, Walter, 1873-1956, former
owner. UK-Lourl.1 > SHL SPECIAL COLLECTIONS WdlM
```

Record	Location		Author				Title		Edition	
Record 1 of 937	LOCATIONS	SENATE HOUSE LIBRARY	AUTHOR	R Shakespeare,	William,	1564-1616.	TITLE	Complete works.		
Record 2 of 937	LOCATIONS	SENATE HOUSE LIBRARY					TITLE	Adam international review	w : our Dylan Th	iomas memorial num
Record 3 of 937	LOCATIONS	SENATE HOUSE LIBRARY	AUTHOR	R Andrade,	E. N. da C. (Edward Neville	e 1887-1971.	TITLE	Isaac Newton.		
Record 4 of 937	LOCATIONS	SENATE HOUSE LIBRARY	AUTHOR	. De B.,		1000-000-000-000-000-00-00-00-00-00-00-0	TITLE	A Naturalist and immortali	ity /	Sec. Art. Sec.
Record 5 of 937	LOCATIONS	SENATE HOUSE LIBRARY					TITLE	The old clay patch: a collect	ec EDITION	2nd ed.
Record 6 of 937	LOCATIONS	SENATE HOUSE LIBRARY		Augustus			TITLE	Poems /		
Record 7 of 937	LOCATIONS	SENATE HOUSE LIBRARY	AUTHOR	R Coleridge,	Mary E. (Mary Elizabeth),	1861-1907.	TITLE	Poems /	EDITION	8th ed.
Record 8 of 937	LOCATIONS	SENATE HOUSE LIBRARY	AUTHOR	R Williamson,	Henry,	1895-1977.	TITLE	The gold falcon, or, The ha	aggard of love :	being the adventure
Record 9 of 937	LOCATIONS	SENATE HOUSE LIBRARY	AUTHOR	R Barnes,	William,	1801-1886.	TITLE	Poems of rural life in the D	Dorset dialect /	

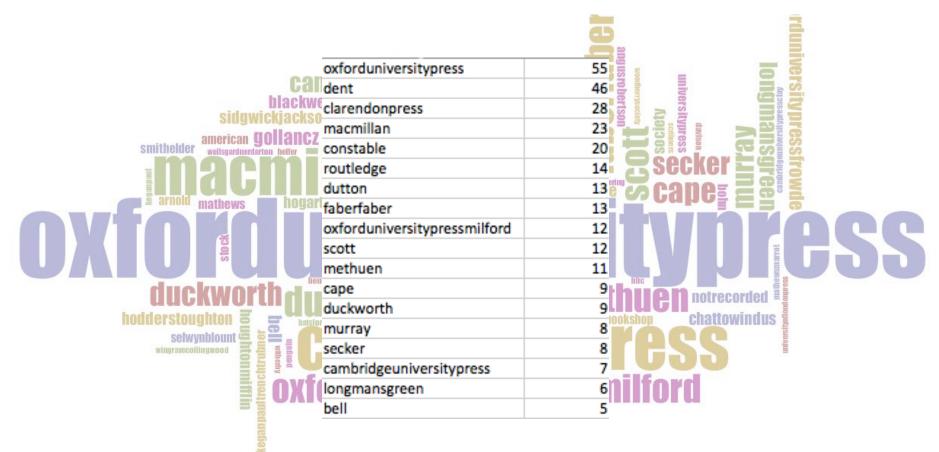


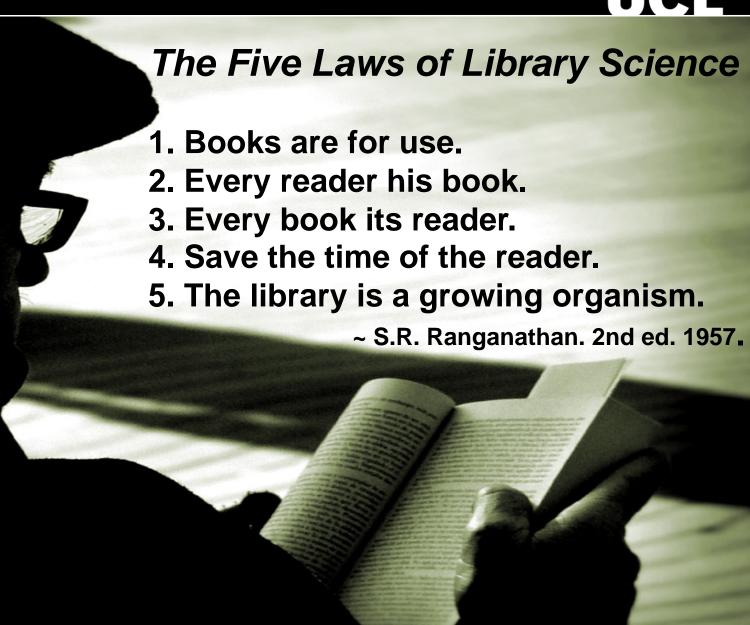
Quantitative Analysis from LMS Output (with critical judgment)





Quantitative Analysis from LMS Output (with critical judgment)









Four generic user tasks have been defined for the purposes of this study. The tasks are defined in relation to the elementary uses that are made of the data by the user:

- to <u>find</u> entities that correspond to the user's stated search criteria (i.e., to locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity);
- to <u>identify</u> an entity (i.e., to confirm that the entity described corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics);
- to <u>select</u> an entity that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs);
- to acquire or <u>obtain</u> access to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an entity electronically through an online connection to a remote computer).



- Four generic user tasks have been defined for the purposes of this study. The tasks are defined in relation to the elementary uses that are made of the data by the user:
- to <u>find</u> entities that correspond to the user's stated search criteria (i.e., to locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity);
- to identify an enit (i.e. the ruinn that the enity describe corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics);
- to <u>select</u> an entity that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs);
- to acquire or <u>obtain</u> access to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an entity electronically through an online connection to a remote computer).



Professional Assistance

"It was necessary to obtain professional assistance in the compilation of the catalogue ... The identification of the books hidden in the very inaccurate entries of the different lists being an arduous and difficult task which I could not possibly assume myself. I was fortunate in obtaining the expert services of Mr R.A. Skelton, Assistant Librarian in the British Museum, who expended upon the catalogue 158 hours of his spare time.

- September 1939, Geoffrey Keynes."
 - Keynes. The Library of Edward Gibbon: A Catalogue. 2nd ed. 1980: 3.



THOREAU'S LIBRARY

BY

Walter Harding



Professional Assistance

"I should add here that Miss Sarah Bartlett has checked my list against the Concord Library holdings." (p. 10)

"Although the editors of Emerson's *Journals* state that Thoreau bequeathed ... books to Harvard, Mr. Robert H. Haynes, Assistant Librarian ... assures me that there is no record." (p. 23)

- Walter Harding. *Thoreau's Library*. University of Virginia Press,1957.



Collaboration

"It occurs to me that professional librarians should often team up with literary scholars for joint efforts on these projects – the resulting studies would be less eccentric in catalog form, would offer fuller cross-indexes and bibliographies, and yet could link the association volumes to the authors' writings."

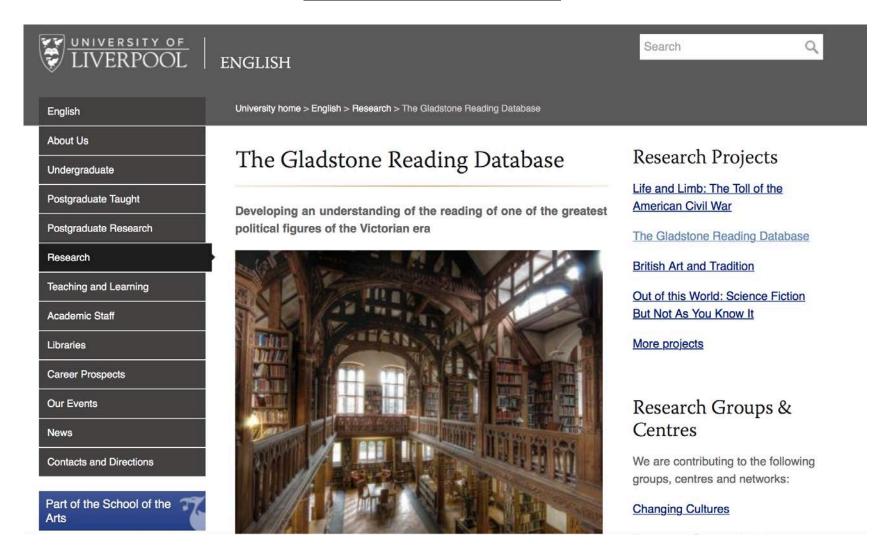
- Alan Gribben. "Private Libraries of American Authors: Dispersal, Custody and Description." *The Journal of Library History* 21.2 (1986): 311.

LIBRARIAN WANTED!

Image: Wikimedia, https://upload.wikimedia.org/wikipedia/commons/8/88/Public_librarian.jpg

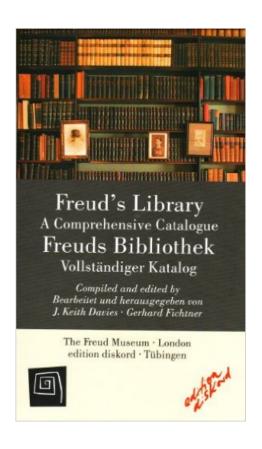


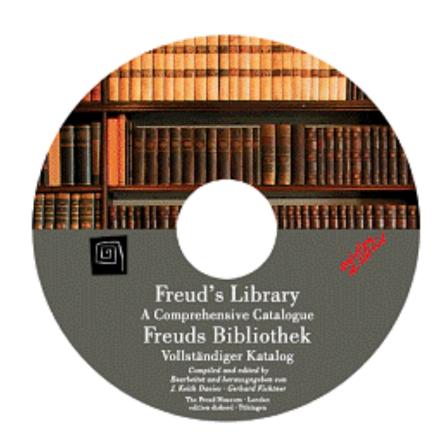
Collaboration





Collaboration





J. Keith Davies and Gerhard Fichtner (comp.). Freud's Library: A Comprehensive Catalogue = Freud's Bibliothek: Vollständiger Katalog. London: The Freud Museum; Tübingen: Diskord, 2006.





Automation

"Word-processors, as well as other apparatuses now beyond our ken, will ultimately supplement the researcher's notecards and fileboxes, but an unquenchable curiosity about the creators and backgrounds of great literary manuscripts will continually bring forth dauntless scholars in each generation."

- Gribben (1986): 162.

Image: Simon_sees, copyright commons, some rights reserved
https://www.flickr.com/photos/39551170@N02/12217141633/



Programming Skills Required

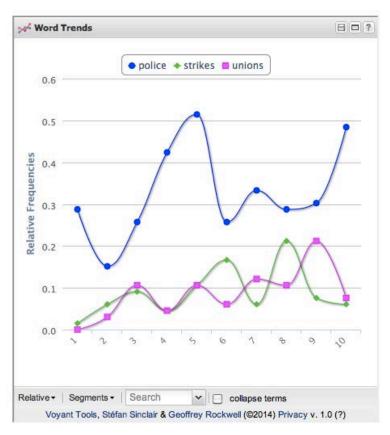
```
500 | a Ms. codex.
546 __|a Italian.
500 __ |a Title from title page (p. 1).
500 | a Incipit: Ruggiero conte di Sicilia vendendo ... (p. 3)
500 [a Pagination: Paper, i (modern paper) + i (earlier paper) + 10 + i (modern paper);
     [1-20]; modern pagination in pencil, upper outer corners.
500 | a Script: Written in a cursive script by a single hand.
500 la Decoration: Illustrated with seven pen-and-ink portraits labeled: Ruggiero I.
     Guglielmo II, Ruggiero III, Guglielmo IV, Tancredi V, Ruggiero VI, and Guglielmo VII, "sette
     rè di Napoli della stirpe de' Normandi" (p. 2).
500 __ a Binding: Modern boards.
500 __|a Origin: Written in Italy, circa 1700.
561 1_|a Formerly owned by Sir Thomas Phillipps ms. 6271 (note, p. 1).
510 4 la Described in Zacour, Norman P. and Hirsch, Rudolf, Catalogue of Manuscripts in the
     Libraries of the University of Pennsylvania to 1800, Supplement A (2), The Library
     Chronicle 36, no. 1 (1970), |c p. 10 (Ms. Italian 241).
524 la UPenn Ms. Codex 465
```

- Mitch Fraas.

Charting former owners of Penn's Codex Manuscripts. Mapping Books: [Blog], 24 January 2014, http://mappingbooks.blogspot.co.uk/ 2014/01/charting-former-owners-of-penns-codex.html



Programming Skills Required



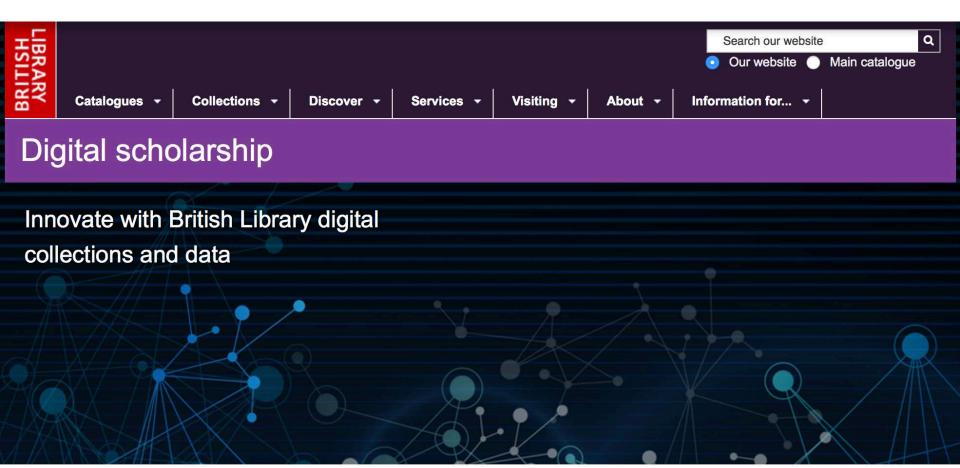
Word trend graph of 'Title' field for data exported from British Cartoon Archive database for years between 1960 and 1979 (dataset, Voyant)

James Baker.

On metadata and cartoons. *Digital Scholarship Blog*, 16 May 2013, http://britishlibrary.typepad.co.uk/digital-scholarship/2013/05/on-metadata-and-cartoons.html



Data Skills Required



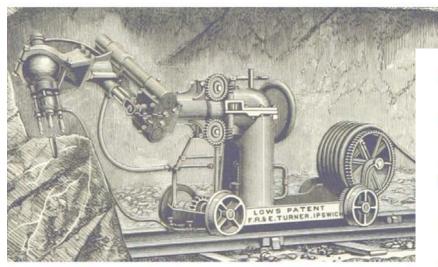
Undertake innovative research with our digital collections and data with the help of our Digital Research experts.



Data Skills Required

Collection guides

Datasets for content mining



Books from the early modern period and the 19th century are among thousands of digitised items available for content mining.

About the collection

A variety of digitised content, including some 68,000 digitised volumes of 19th-century books, including literature, historical and philosophical works are available for content mining.

If you have an interesting research idea in mind that relies on content mining, please contact digitalresearch@bl.uk to discuss.

What is available online?

19th Century Printed Books: 68,000 digitised volumes (around 25 million pages) published between 1789 and 1900 cover a wide range of subject areas including philosophy, history, poetry and literature. The 19th-century books are on a 30TB storage device with the original JP2000, PDFs, XML and OCR files in Alto format. We are currently in the process of extracting all OCRd text from the 25 million pages.

Early English Books Online (EEBO) 1473-1700: includes thousands of digitised texts from British Library's collections. 25,000 texts in EEBO have now been fully transcribed as part of the Text Creation Partnership Phase I.



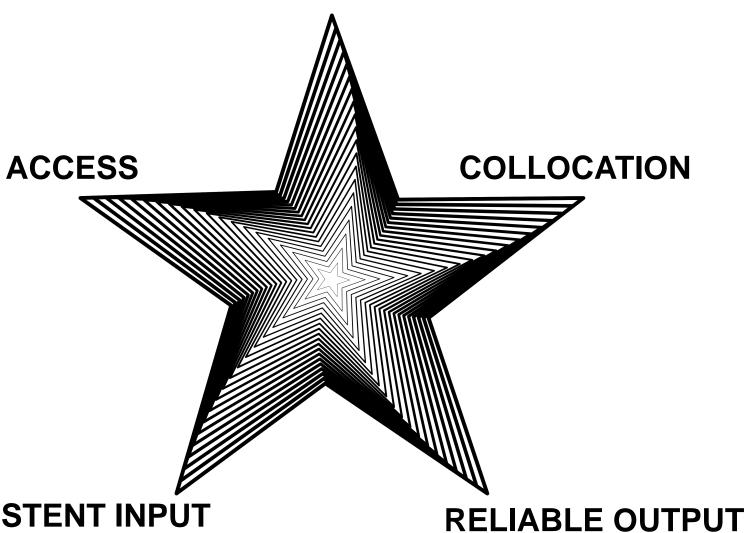
Data Skills Required





A Wish Upon A Catalogue

DISCOVERY



CONSISTENT INPUT