

The Rare Books Catalog and the Scholarly Database

Anne Welsh

Department of Information Studies, University College London, Gower Street,
London WC1E 6BT

Abstract

A researcher's eye view of the value of the library catalog not only as a database to be searched for surrogates of objects of study, but as a corpus of text that can be analysed in its own right, or incorporated within the researcher's own research database. Barriers are identified in the ways in which catalog data can be output and the technical skills researchers currently need to download, ingest and manipulate data. Research tools and datasets created by or in collaboration with the library community are identified.

Keywords

Library catalogs / OPACs, Catalog indexing / display / design, Information retrieval, Bibliographic data - interoperability

Introduction

... 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 historia.

... 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.¹

This quotation, in Portuguese and English, on a wall in Lisbon's Gulbenkian Museum, cuts right to the heart of this article, which builds on a presentation at the Chartered Institute of Library and Information Professionals Cataloguing and Indexing Group Conference 2014² and a short communication in *Catalogue and Index*.³ It draws on original research into the Working Library of Walter de la Mare and a literature review of activities exploiting the catalogs of writers' libraries, which are *de facto* catalogs of rare materials. In doing so, this article explores the uses to which catalog data is being put by libraries in evaluating their collections and by researchers. Writers' libraries are presented as an example of an area of academic research into materials often held by special collections departments and for which bibliographic research and description is core. This article highlights the potential for library data to be more than a finding aid for researchers: to be also the foundation for scholarly databases.

Provenance and the History of Reading

In his seminal work *Provenance Research in Book History*, Pearson asserts that "The serious study of private libraries, and of the lessons which can be learned from book ownership, is a growth industry and one which has gained much ground in the recent past."⁴ He uses the example of St Cuthbert's

gospel-book to indicate that the roots of interest in provenance stretch back to the Middle Ages, when the gospel was included as a relic at Durham Cathedral.⁵ However, he makes a distinction between “venerating a book” and using its provenance information as evidence in the modern sense. The former is, of course, a spiritual experience, while the latter is both a tool for assessing collections and an academic research technique.

Interest in reading really flourished in the last decade of the 20th century,⁶ and Pearson’s first edition and its reprint with corrections spanned the rapid expansion of the field, so that he was able to observe in his “Introduction to the 1998 Reprint” that “There is a steadily growing literature on the ownership and use of books, embracing works on particular private libraries, studies of marginalia, and the new academic vogue for the history of reading.”⁷ This is not to suggest that readers were not of interest to scholars in earlier times. Within Literary Studies, reader response critics of the 1960s and 1970s introduced concepts including the intended reader and the unreliable narrator,⁸ interpretive communities,⁹ the ideal reader,¹⁰ the implied reader and the creation of meaning between the author and the reader,¹¹ the writer’s imagined audience,¹² and the role of the critic as mediator for the author and reader.¹³

However, Rose highlights a paucity of source materials for those interested in the physical evidence of the History of Reading. As a result, although Altick was able to write *The English Common Reader* in 1957,¹⁴ he was only able to spend one chapter on acts of reading themselves, and as Rose puts it, recent “scholars have, with considerable ingenuity, located and used a wide range of new materials that allow us to fill in the vast blank

spaces on Altick's map" to such an extent that "it is fairly astonishing to recall that just twenty years ago, the history of the common reader was widely believed to be unrecoverable."¹⁵

Today, classes on Book History often start with Darnton's Communications Circuit,¹⁶ with "Readers: Purchasers, Borrowers, Clubs, Libraries" towards 'the end' of the cycle, but with a dotted line from them to the Author at the start, representing (1) the iterative and reiterative influence that readers have on the creators of works and (2) that the processes associated with the life cycle of the book are not linear but circular in nature.

The centre of Darnton's circuit shows the major forces that shape and give rise to books: "Economic and Social Conjunction"; "Political and Legal Sanctions"; and "Intellectual Influences and Publicity." As Eliot and Rose have summarized, "Books are made by history: that is, they are shaped by economical, political, social and cultural forces," and at the same time, books can influence the wider world: "Readers can read the same book in a variety of different ways, with important consequences: after all, wars have been fought over differing interpretations of treaties."¹⁷

Modern researchers heed the warnings of earlier scholars not to be myopic in their studies. The publishing historian John Feather has been vociferous in reminding us of this point: "Book historians who are not at least aware of bibliographical techniques are ill equipped for their task, and it could be forcefully argued that a knowledge of historical bibliography should be the basis of their training as scholars. We should also, however, venture outside the confined space of the printing house into the world in which its products were used."¹⁸ As Secord has put it, the History of Reading encompasses "all

the diverse ways that books and other forms of printed words are appropriated and used,”¹⁹ while, as Jackson has pointed out, it does not simply have one sole aim to “recapture the mental processes by which readers appropriated texts.”²⁰ She highlights Darnton’s work on banned literature in 18th century France,²¹ Raven’s ongoing interest in the impact of mass market publications,²² St Clair’s study of access to reading material in the 19th century and Manguel’s wide-ranging essays,²³ which she describes as “tell[ing] us about the evolution of material accompaniments to reading.”²⁴

The tools and sources of information for discovering these histories are diverse, and sometimes prosaic: marginalia, notebooks, letters, autobiographies, library borrowers’ registers, booksellers’ lists and library catalogs – both private (such as the one Walpole made of his father’s collections) and institutional. In the last of these, our 21st century library catalogs could – and we might argue should – play a central role.

Writers’ Libraries as a Field of Study

The study of writers’ libraries lies at the nexus of Bibliography, Book History, Library History, Cultural and Literary Studies. Darnton suggests that “Most of us would agree that a catalog 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 catalog of the library in Monticello is to inspect the furnishings of Jefferson’s mind.”²⁵ Leah Price, in the introduction to her collection of photographs of authors’ bookcases and interviews with their owners asserts that “Bookshelves reveal at once our

most private selves and our most public personae. They can serve as a utilitarian tool or a theatrical prop.”²⁶

Herein lies the challenge and the promise of such study – in order to understand a writer’s library we must have a certain level of knowledge of their creative output and yet assessment of their book collection can enhance our appreciation of a writer’s working methods, and ultimately, their work. Gribben has asserted that librarians “may feel somewhat uncomfortable” in making judgments like these, whereas English professors may lack advanced technical skills in Bibliography. Given the number of literature graduates and doctorate holders who go on to train as information professionals, we may in fact feel decidedly uncomfortable with this segregation by person holding a role. However, we might fairly agree with his summary that “The study of an author’s library and reading is a borderline area between literary studies and library science”²⁷ as disciplines.

In his handbook *Collecting, Curating, and Researching Writers’ Libraries*, Oram provides us with a straightforward definition of 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.”²⁸ Within Cultural Studies, Museum Studies, Art History and Psychology there are many publications concerned with what we may call ‘the collecting habit’ and the motivations for collecting. In the library setting, we can consider Attar’s assertion that “Books in a library differ fundamentally from books anywhere else in that each one is part of a

collection, an aggregation that imposes its own meaning derived from the decisions and accidents that went into its formation.”²⁹

This view of ‘the collection’ is in line with Tanselle’s wide definition that “collecting is the accumulation of tangible things.”³⁰ It also places the agency of determining the “coherence of a ‘collection’”³¹ firmly in the hands of the librarian and sidesteps concerns about whether the previous, private owner styled themselves a collector or not.

Private collections may have been deliberately constructed as such by their original owner. Such an example is the Phyllis T.M. Davies Collection of books by Walter de la Mare now at Cambridge University Library.

Alternatively, the previous owner may have acquired some of their books as working materials while others may have been specifically collected. In interviews with Oram and MacDonnell, Ted Kooser, Russell Banks and Jim Crace each make such a distinction in the books they own.³² Edmund White is an example of a living author who has sold manuscripts to the Beinecke Library and given them “books that helped me in my own work” but now has “the feeling the librarians don’t really want those research books anymore [sic].”³³ Instead, he talks about “go[ing] to the Strand and buy[ing] every book about Henry James’s letters, for instance, that I can find, but I ship them out as soon as I’m finished [writing] the essay.”³⁴ Here, there is a clear view of materials that are suitable for an institutional library, books he wants to retain, and books that, in his view, can be treated as ephemeral and disposable.

Other writers may have an egalitarian completest attitude to the books on their shelves. Junot Díaz, for example, reports “I have never liked the idea of a hidden book. It means no-one will ever randomly pick it up and have a

conversation with you about it.”³⁵ At the opposite end of the scale from the private collection that has been carefully constructed by its original owner, there are, of course, those who would agree with Ron Powers that “Books are to be read. Not to be resold, speculated in, sanctified, put on shelves as indicators of intellect or status, or otherwise violated. Read. Period.”³⁶ Turning to Tanselle again, “What one person accumulates haphazardly, another will regard as bearing a design; and even the product of a careful plan may turn out to be of interest to another person for an entirely different pattern that can be put into it.”³⁷

Writers’ Libraries in the Scholarly Record

An early example of a writer’s library that received scholarly attention is that of Edward Gibbon, the author of *The Decline and Fall of the Roman Empire*. William Beckford, the author of gothic bestseller, *Vathek*, bought Gibbon’s library in France wholesale, and although he was a famous bibliomaniac, he did not add it to his collection at home, but instead said that “I bought it to have something to read when I passed through Lausanne ... It is now dispersed, I believe. I made it a present to my excellent physician.”³⁸

Significantly, “when asked if the books were rare or curious ... He replied in the negative. There were excellent editions of the principle historical writers, and an extensive collection of travels. The most valuable work was an edition of “Eustathius;” there was also a MS. or two. All the books were in excellent condition; in number considerably above six thousand, near seven, perhaps.”

Here, we can see a later collector assessing a collection that, according to his bibliographer, was considered by its original collector “a

working library” and who himself declared “I am not conscious of having bought a book from a motive of ostentation.”³⁹ Beckford was clearly interested in the books – he claimed to have shut himself away for six weeks to read them – but having read them once, had no further desire to keep them, so he gave them away. Nor did he give any indication of considering the books special because such a famous author had been their original owner. Yet, a century later, Geoffrey Keynes went to great pains to piece together the bibliography not only of the library in Lausanne, but also of the other books Gibbon amassed in his lifetime.⁴⁰

In 1939, on compiling the first edition of *The Library of Edward Gibbon: A Catalogue*, Keynes reported that while bibliographers and librarians had worked on “the catalogues of the great libraries [which] enable the individual to consult the universal mind ... The library collected by one man ... expresses only his own interests and a catalogue of the books it contains can have no value unless the mind that it reflects is one of very universal distinction. Seldom, therefore, has it been thought worth while [sic] to attempt to reconstruct the individual libraries of the writers of the past.” Again, reflecting Rose’s claim for Altick’s experience of a paucity of data, Keynes writes, “Usually, indeed, no material has existed for such attempts, unless it were an auctioneer’s catalogue hastily compiled after the owner of the books had died.”⁴¹

Scholarly interest in writers’ libraries and the formulation of methods to reconstruct them awaited the rise not only of the reader response critics in Literary Studies, but a widening of focus from a canon of literature considered worthy of academic study to the full range of texts published from the 15th

century onwards, and the increasingly diverse group of people able to read them. Halsey has identified this work and the rise in the study of the book as object as leading “inexorably to a focus on readers, both contemporary and historical.”⁴²

At the same time, within Cultural Studies, Museum Studies and Art History, there have been accounts published of individual private collections, and within Psychology some work has focused on Collecting, on Creativity, and on the relationship between them. As Muensterberger asserted in his book *On Collecting: An Unruly Passion*, “We know from contemporary artists’ collections that they provide animation and inspiration, or may even sway his barely conscious susceptibilities, long before the artist himself is fully aware of the source.”⁴³

Different writers have different ways of working. As Lev Grossman has admitted, “I read obsessively when I’m writing. I think there are two kinds of fiction writers, those who read incessantly while they write, and those who can’t read at all, lest their individual voices get overwhelmed, or tainted somehow. I’m the first kind.”⁴⁴ Gary Shteyngart would seem to agree: “You read, then you write, then you read some more, then more writing, and so on in an endless wordy loop.”⁴⁵ However, Jim Crace’s experience is different: “I don’t see the books themselves as sources for my books. But ... whenever I am in the countryside and hiking, then I do feel creatively grand. As a landscape writer I feel deep and dirty amongst my sources ... I suppose I want to allow for the idea that our libraries reveal much but that other chambers of our lives can disclose a good deal more.”⁴⁶

In terms of scholarship of a particular writer, Gribben has described the process that can be observed: “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, if the quality of the literary canon is sufficient, a grove of increasingly specialized studies takes root, affording protective 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.”⁴⁷

The current article comes from work undertaken in the course of studying the Working Library of Walter de la Mare. We can see that Gribben’s pattern is roughly followed: de la Mare’s first publications were short stories in *The Cornhill Magazine* in 1896-7. His first book was the poetry collection *Songs of Childhood*, published in 1902 under the pseudonym Walter Ramal, with his first novel, *Henry Brocken*, following in 1904. In 1905 he began his career as a reviewer for *The Bookman* and started working for the *Times Literary Supplement* in 1908 and *The Westminster Gazette* in 1909, and so it makes sense that the first biographical study of him was a chapter in Adcock’s 1923 *Gods of Modern Grub Street*. The following year, R.L. Mégroz’s *Walter de la Mare: A Biographical and Critical Study* was published, to be joined in 1929 by Forrest Reid’s *Walter de la Mare: A Critical Study*.

In terms of bibliographies, Danielson's was published in 1921 and Murphy's in 1927. The latter was praised by Bowers in a footnote of his *Principles of Bibliographical Description* as a modern bibliography that "excels in the details of the descriptions."⁴⁸ Other biographical, bibliographical and critical works have been published, with the current key biography now that of Whistler in 1993⁴⁹ and the main critical work Bentinck's published in 2001.⁵⁰ Senate House Library completed cataloging the Working Library and Family Archive of the Printed Oeuvre in 2010,⁵¹ and work on my PhD focused on it began shortly thereafter. All fairly typical, according to Gribben's ecosystem.

The Scholar and the Computer Catalog

Gribben was also fairly prescient, in 1986, in foretelling an increase in the use of the computer within research. Although "doubt[ing] whether such permanent records of library borrowings" as Kesselring used for her work on Hawthorne and his family's loans from the Salem Athenaeum⁵² "will survive from our contemporary computer-assisted libraries,"⁵³ he was optimistic about new methods and tools for research: "The technology and determination that enable us to penetrate outer space will most likely also give us better means to explore the intellectual lives of our cherished authors. 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."⁵⁴

Early bibliographies of writers' libraries often relied on the merging of various lists produced by the writers themselves, booksellers' inventories and

library catalogs. This was a complex and time-consuming task, and reliance often had to be placed on the bibliographic skills of library staff. Keynes identified early on that “It was necessary to obtain professional assistance in the compilation of the catalog from the various sources ... 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.”⁵⁵ For the first edition of *The Library of Edward Gibbon* in 1939, R.A. Skelton spent “158 hours of his spare time during the course of six months”⁵⁶ while for the second in 1980, David McKitterick produced a list of “61 entries with one correction of an entry in the original Catalogue and one addition of a book now in my library.”⁵⁷

Bibliographers report the usefulness of library catalogs for checking bibliographic details of books appearing in authors’ correspondence and personal lists. As Harding puts it in his 1957 study *Thoreau’s Library*, “When all these sources were exhausted, I resorted to a search of the Library of Congress Catalogs, the British Museum Catalogs, Sabin, Roerbach, and the catalogs of such nineteenth-century libraries as the Boston Athenaeum.”⁵⁸ Again, direct involvement of library staff was necessary: “I should add here that Miss Sarah Bartlett has checked my list against the Concord Library holdings,”⁵⁹ since in 1957 it was not possible for the researcher to check the paper donations files himself.

By 1981, when Reynolds published *Hemingway’s Reading, 1910-1940* he was able to report on computerized searching techniques: “The Ohio College On-Line Catalog contains a data base [sic] of over four million book entries. From our local terminal we needed only to punch in the first words of

the title and part of the author's last name. Within seconds, a list of possibilities would appear on our screen. The mindless computer loved nothing better than to search for Browns. If, however, either author or title were the least misspelled, we drew a blank, for computers do not think or guess."⁶⁰ However, not all their searching could be conducted online: "Books published only in England, for example, might not be in the data base [sic]. We turned back to the *British Museum Catalogue* ... The last resort was the *Catalogue of English Books*. Year by year, we thumbed through until we found the entry."⁶¹

The Catalog Record and the Library User

It is important to note that the computer catalog data being searched by scholars from the 1960s onwards were mostly MARC records, and to acknowledge the inherent limitations in their structure.⁶² As Avram pointed out in *The MARC Pilot Project Final Report*, the main aim of the original project was "to test the feasibility of a distribution service of centrally produced machine-readable cataloging data."⁶³ Although Information Retrieval was a consideration, it was third in a list of four assessment criteria for the format, with printing at the top of that list – the production of catalog cards was a key driver for the project.

Crucially, the designers of MARC state their own awareness of their limited fore-knowledge of how researchers might search for information: "Since so little is known about how a bibliographic record will be used in machine-readable form for retrieval, it was only possible to anticipate future applications."⁶⁴ Although we have seen MARC go through several versions to

become MARC 21, the underlying structure of the records, and the bibliographic information we record within them, remains much the same even today.⁶⁵ It was not until the 1980s that keyword searching of the whole bibliographic record became possible,⁶⁶ and so the search experiences described by Reynolds, in which strict knowledge of author and / or title was needed, relay an authentic first-person report of researcher interactions with the library catalog at the time.

Indeed, in the early 1980s, although public catalogs were designated OPACs – Online Public Access Catalogs – “online” did not have the same connotation in Pre-Internet days as it does now, and the catalog terminals were merely networked within their own local area, usually just the library building itself.⁶⁷ So the simplest and most convenient way to obtain information about private collections now dispersed across institutions was to contact them, as Edel and Tintner encouraged those interested in Henry James’s Library to do: “The several university collections have complete bibliographic details, including the number of volumes in each title, whether or not the pages have been cut, and what, if any, exact marginal annotations appear. Transcripts of their lists are available from the libraries.”⁶⁸

Within the 21st century catalog, the need, or otherwise, for separate listings of special collections is largely a question of (1) local cataloging decisions concerning the quantity and quality of information recorded in the catalog (2) catalog search options and (3) options for users to output search lists themselves. As Joseph Nicholson has highlighted, in the modern catalog and discovery layer, “the vagaries of keyword searching and a lack of uniformity in note fields can make it painfully difficult for users to track down

books belonging to particular private libraries in online catalogs.”

Nevertheless, “catalog records should be constructed in such a way as to allow patrons to identify books that belonged to a particular writer and, equally important, to retrieve them as part of a group.”⁶⁹

Although FRBRization (the implementation of IFLA’s *Functional Requirements for Bibliographic Records*’s bibliographic model in cataloging, or an approximation of it within the discovery layer) places “relationships at the heart of the catalogue,”⁷⁰ in real terms the collocation of materials that formerly shared the same private ownership in ways that assist in their retrieval is complex. Senate House Library provides a good example of ways in which researchers can retrieve records from named collections in which they are interested: (1) by searching by author name with the addition of the phrase “former owner”; (2) by locating a known item and then using a hypertext link to search for the other items in the collection (3) by searching by local classmark. It is also possible to approach the search via a collection description on the special collections web page. Each of these options has strengths and weaknesses and might suit researchers arriving at the library catalog with different levels of information retrieval experience and expertise.

Collection Description

As Nicholson has summarized, “The private libraries of writers pose a number of peculiar challenges to catalogers in special collections units due to a hybrid identity that incorporates aspects of both archival collections and books.

Though the fundamental unit of the private library is the book ... the textual bedrock of such collections often serves as a substratum on which layers of

materials commonly considered archival in nature have been deposited.”⁷¹

This “quasi-archival” nature of writers’ libraries⁷² makes description at collection level a *sine qua non* for best practice, and Nicholson suggests both the creation of collection-level MARC records, following *Descriptive Cataloging of Rare Materials (Books)* (DCRM(B)) Appendix B, and the creation of a collection description on a library web-page.

Senate House Library has not created collection-level MARC records, but it has provided a useful summary of its main special collections on the library website. The page on the *Walter de la Mare Collection* provides an overview of the Working Library and the De la Mare Family Archive of Walter de la Mare’s Printed Oeuvre, a note on de la Mare, with a link to his entry in the *Oxford Dictionary of National Biography*, brief information on acquisition, an indication of related holdings, and a select list of publications about the collection. Crucially, the information on access provides a summary of when the materials were cataloged and first became available to researchers, and the instruction to carry out a local classmark search.

Locating brief collection descriptions on a webpage is also useful for the researcher since such websites are indexed by search engines, while there is much within library catalogs that still resides in the deep web. As Oram and Nicholson point out in the introduction to their directory of writers’ libraries, “basic information on library holdings of writers’ libraries is difficult to obtain.”⁷³ They indicate institutional cataloging and differing practices as a key issue here, but to this we may also add the general difficulties in even carrying out general web searches to look for the books a writer once owned versus the books she wrote. A collection description on a library webpage, as well as

entry into whichever published directories seem appropriate is a great assistance to researchers in this field.

Search is Only Elementary

So far we have devoted a fair amount of time to tracing searches and search strategies, and this reflects a major way in which researchers interact with libraries and their catalogs. In our attempts to fulfill Ranganathan's *Five Laws of Library Science*, we look towards FRBR's four user tasks to help us in our aims "Every reader his book" and "Every book its reader."⁷⁴ These tasks, familiar to every cataloger, are "to find entities that correspond to the user's stated search criteria ...; to identify an entity ... ; to select an entity that is appropriate to the user's needs ... ; to acquire or obtain access to the entity described [underlining in original]."⁷⁵

First published in 1998, these tasks have become the objectives of those of us involved in Cataloging and Information Retrieval, and, when we think about how a particular group of researchers may use our facilities and tools, it is natural for us to have care for the ease with which they are accomplished by researchers. Indeed, it might be argued that until our Library Management Systems were capable of offering full search capabilities, it is right that we should focus our efforts in this area. Certainly, the rise in ethnographic research into catalog use⁷⁶ reflects a widening of our interest in how people search.

However, it has never been argued that search is the only use to which library users may put the catalog. Even within the FRBR report, the wording introducing the user tasks is "Four generic user tasks have been defined ...

The tasks are defined in relation to the elementary uses that are made of the data by the user.”⁷⁷ So the tasks are “generic” – the ones that most, if not all, users will undertake; the uses themselves are “elementary” – the beginning or starting point (note that the word “primary” was *not* used); and what is being used is “data” – not limited to records, nor to results lists – data.

Perhaps we are so used to the fundamental concept that the catalog record is a surrogate for the material itself and to the argument that what users want is full-text access that it becomes easy for us to overlook the status of the catalog itself as data. Smiraglia has pointed out that the catalog is a cultural artifact,⁷⁸ while Whaite has grounded this in history – “A catalog that is in use is a finding tool, but when a newer version is introduced, the old catalog becomes a relic of its time.”⁷⁹ We might extend this into contemporary history: while a catalog’s elementary use is as a finding tool, its data can tell us of its time. As Andersen has made explicit, there is a materiality that we can explore in the “bibliographic record as text.”⁸⁰

If we accept that our catalog data is not solely paratext for the item it describes, but is also a text in its own right, we can open it up to any of the research techniques for text, including any of the tools that have been devised for the Computational Linguistics that is often seen to be the genesis of the Digital Humanities. If text-mining techniques can be used to reveal the structure of Pynchon’s novel *V*,⁸¹ how might they be used to examine the structure of library catalogs? What quantitative analyses might we, or the people who use our catalogs for access, wish to carry out on our data, and how might this be possible?

Quantitative Analysis of the Catalog

Within libraries, we are already beginning to exploit quantitative tools. OCLC's Worldshare Collection Evaluation⁸² provides reports on collections, including "comparisons to individual libraries, peer groups and benchmark library groups."⁸³ As it is marketed to libraries for collection management purposes, the three uses that OCLC have chosen to highlight are acquisition, deselection and accreditation, but some of the other features could be put to interesting work on individual library collections, such as those of writers' libraries. For example: "View detailed information including title, subject analyses and local circulation data; export comparison data for offline analysis and reuse; [and] visualize comparison data."⁸⁴

Copac has recently made their Collection Management Tools⁸⁵ available "via single sign on (Shibboleth), so existing users no longer need their CCM [Copac Collection Management] Tools username."⁸⁶ This development also means that use of the tools is no longer restricted solely to the librarians responsible for collection development – anyone whose institution is part of the Shibboleth consortium⁸⁷ can use their university login to access the tools. As with OCLC's Worldshare Collection Evaluation, Copac Collection Management Tools have, naturally, been envisaged for use by libraries to develop their collections: "By providing a suite of search and visualization options, users can make use of the rich holdings data in Copac: to support difficult decision making about what materials to keep, remove, conserve or indeed purchase."⁸⁸

Case studies and user stories highlight collection management activities. However, there is clear scope for researchers to use the tools in

academic work. To pick one example, Copac reports that “At St. Andrews University the Tools were used to assess a significant donation to the library, confirming its value.”⁸⁹ Such value statements are of obvious use to researchers in deciding whether to travel to see a collection and, on occasion, in writing the introduction to an article. Knowing how rare within public collections the contents of a formerly private library are is of clear scholarly interest, as is being able to determine quickly and easily the extent to which an author’s books are held within UK research libraries.

Such approaches could be said to belong to the academic field of Digital Bibliography – the creation and / or application of computational tools and techniques to the study of the book as object. Other projects we might claim for this interdisciplinary area might include *Early Modern Print*,⁹⁰ which has developed and made available tools such as EEBO-TCP Keywords in Context and the EEBO-TCP N-grams Browser, and the many projects of the Consortium of European Research Libraries (CERL), including its Heritage of the Printed Book (HPB) database⁹¹ and its Material Evidence in Incunabula,⁹² built on the Incunabula Short Title Catalogue with added provenance and annotation information and links to the CERL Thesauri (of provenance and place names).

Many of these projects are built on MARC records: library records in MARC are plentiful and it does not take a great training in quality control for researchers to gain an understanding of the sources of the best quality, most detailed records for their needs. OCLC, Copac and CERL are examples of consortia that provide their member libraries with many different services – not only tools and data hosting, but also training, conferences and

opportunities to take part in research and development projects. As a result, there is a great deal of good-quality data about collections hosted in tools that researchers can use – good quality data derived from library catalog records. Are we beginning to see the flowering of Attar’s prediction from 2004 of “the developing function of a catalogue record as a research tool in itself, instead of a mere finding aid”?⁹³

Current limitations in library data and systems

The basic structure of MARC pre-dates the Internet, and its consequent limitations and lack of flexibility in terms of sharing with systems outside libraries has been well-documented, from Tennant’s original opinion piece that “MARC Must Die”⁹⁴ through to the Bibframe Primer⁹⁵, in which MARC’s structural failings are highlighted as a rationale for the Library of Congress to explore future possibilities for data exchange formats. The growth of the semantic web and correspondingly in researchers using semantic web technology, standards and schema, such as the *Resource Description Framework* (RDF) on which Bibframe is based, makes this an attractive successor, with the promise that we can move from a situation in which the data in our web catalogs is essentially hidden in the deep web where standard search engines do not penetrate, through what Willer and Dunsire have termed “a manifesto for a paradigm shift”⁹⁶ so that our data is an integral part of the semantic web.

In practical terms, as well as being less open than we might like it to be in the 21st century, the library data that we create in MARC records requires a high overhead of systems work. As has Pūrongo summarized, “Library

systems experts now spend time managing MARC data – manipulating it, doing quality management activities, and mining it, and all the while keeping the MARC data moving through systems in client server architecture.”⁹⁷

Beyond the catalog itself, in order to make our data available on the Internet, we have to publish it, and that too requires technical skills. Many national libraries, who have an ethos of openness and, therefore, of wanting to make their data available for use by researchers, have web pages dedicated to download options, so that those outside the library with the technical skills can ingest their data and manipulate it.

Recently, national libraries including the British Library, the Library of Congress, the Deutsche Nationalbibliothek and the Bibliotheque Nationale de France (BnF) have made data available in RDF for researchers, allowing those with linked data skills to incorporate these datasets in their work. An important consideration in this work that has been expressed by researchers from the BnF is that “as the main purpose of the library is to give access to documents for patrons, the HTML publication ha[s] to be coherent with the RDF publication, the data in RDF being just a different view from the same data that is in the HTML page.”⁹⁸ In this brief statement we can see the major challenges of presenting our catalog data to the world in the 21st century: data that is created in MARC has to be published in XML and also in RDF so others can reuse it.

Digital Bibliography

There have been several projects that have used both programming and data skills in order to manipulate library data in order to answer bigger Humanities

questions. For example for the ALA Midwinter Hackathon in 2014, Mitch Fraas created a network visualization of former owners of manuscripts at University of Pennsylvania Libraries “in the hopes that it [would] be not only useful to scholars but also might generate some conversation over how libraries and archives distribute their valuable descriptive information.”⁹⁹ While demonstrating the techniques he used to create the visualization, he makes a remark part-way through his account, which is worthy of notice: “I realize now that this task would have been near to impossible at most libraries where the online catalogs and back-end databases don’t easily allow public users to batch download full records. Fortunately at Penn all of our catalog records are available in MARC-XML form.”

Issues in downloading data from public catalogs recur in the literature. In her PhD thesis on the Raymond Klibansky Collection at McGill University Library, Tomm writes about having to feed catalog data through reference management software so that she could obtain the data she needed in a format she could manipulate and use,¹⁰⁰ while Baker, writing about a small project working with data from the British Cartoon Archive, reported that he had to run programming scripts to cleanse the metadata before he could run it through the quantitative analysis tool he was using.¹⁰¹

Even in downloading records in basic formats, like CSV, which should allow easy importation to spreadsheet programs, researchers working outside the library can encounter issues. If data has not been specifically published for use by researchers, it is not uncommon to encounter issues with the standard download options offered by Library Management System providers. In 2010, for example, it was not possible to download the entire set of records

for the Working Library of Walter de la Mare from Senate House Library's public catalog. The CSV file, once imported into Excel, located all the information in the first cell. Massaging the data through reference management software, as Tomm did for her project, resulted in records which lacked their provenance notes. The only ways forward were to either re-key the data by hand, or download a file and manually tab delimit it, or ask the systems team at the library to publish the data. Although larger libraries are beginning to look into bespoke data publishing for researchers, there are, of course, many workflow and cost issues surrounding supply and demand for any one-to-one services.

As things stand currently, download options from most standard library catalogs are targeted towards reference management. A show of hands at the Cataloguing and Indexing Group Conference in 2014 revealed that only one of the delegates present had tested that all the download options offered by their catalogs (not just the reference management ones) resulted in the download of complete records.¹⁰² The other catalogers present in the room had, until that point, been content to input high-quality catalog records without double-checking how readers might output them. Essentially, we had been creating surrogate records to satisfy FRBR's elementary user tasks, all of which are focused on search.

Exploitative Power in Bibliographical Control

In 1968, when computerized cataloging was in its infancy, Wilson published his seminal work of cataloging theory, *Two Kinds of Power: An Essay on Bibliographical Control*.¹⁰³ In it, he argued that our power is of two kinds –

“descriptive” and “exploitative.” We could make a case that from the birth of MARC and the computer more generally, as a profession we have been focused on the first of these, and that now, with greater opportunities to manage and manipulate data, we can focus on the second, which, in Wilson’s own argument, was always the greater. Smiraglia’s summary of “exploitative power” is still pithiest and clearest – “the power of a scholar to make the best possible use of recorded knowledge.”¹⁰⁴

If the drive in the heritage sector to open up its collections data is to result in a large uptake by researchers, it is necessary to be aware that there are barriers to use¹⁰⁵ and to work to overcome them. As well as practical issues such as interoperability and fragmentation of resources and tools, there can be a lack of technical skills among Humanities researchers, and even a lack of awareness of the possibilities that exist.¹⁰⁶ An acknowledgement of the need for encouragement to experiment can be seen in the creation of departments such as the British Library’s Digital Scholarship team, with a remit to encourage use of datasets published by the Library, including running outreach events and competitions. Meanwhile, the presence of academics within libraries through awards such as the Kluge Fellowship in Digital Studies at the Library of Congress further encourages collaboration on digital projects using library data.

Collaboration has not always been the history of projects involving or recreating writers’ libraries. In keeping with his account of early OPAC searching, Reynolds has given a summary of how, having used the catalogs to identify Hemingway’s books and “carded each: author, title, date, genre, source, subject and contents,” the academic team used “the massive tri-

university computer to record the data, using a program specially written by George Van den Bout.”¹⁰⁷ They were able to sort by category as well as by year, and this assisted in their critical thinking about the topic of Hemingway’s reading.

Today there are projects in which teams of academics create the entries for their databases from start to finish. Some, such as *Melville’s Marginalia Online*¹⁰⁸ take as their starting point a published bibliography – in this case Sealts’s “Check-list of Books Owned and Borrowed” and the supplements published by Sealts and Olsen-Smith.¹⁰⁹ The most prominent database in the History of Reading is arguably the *Reading Experience Database*, which collects information on online forms from volunteers in Australia, Canada, the Netherlands, New Zealand and the UK.¹¹⁰

However, there are also projects in which libraries and academics collaborate. Davies and Fichtner’s *Freud’s Library: A Comprehensive Catalogue* published a CD-ROM with full catalog in English and accompanying text in German and English describing the project and reporting on its findings on Freud’s books and reading habits, including qualitative and quantitative analysis of the subjects in his collection, the provenance of the books, and the number of dedication copies he received.¹¹¹

Another project which utilizes a Library Management System at its core is The Gladstone Reading Database (GladCat) – a project between the University of Liverpool and St Deiniol’s Library to document the books owned by Prime Minister William Ewart Gladstone, including indicating his annotations.¹¹² Originally funded by the Arts and Humanities Research Council (AHRC), the catalog has continued to be maintained and added to by

the Library after the project finished in 2009.¹¹³ There have been publications and a funded PhD associated with the project, and University of Liverpool Department of English submitted it as an impact case study in the Research Excellence Framework (REF 2014) exercise. The submission claimed “The primary impact of the Gladstone project has been to preserve an important part of the UK’s cultural heritage and to make it available to audiences outside the academy. Specifically, it has stimulated tourism to Gladstone’s Library, a significant commercial heritage institution in the area, and in Wales second only to the National Library of Wales in Aberystwyth, with holdings of over 250,000 volumes. In particular, it has enabled a significant re-orientation of the Library’s marketing strategy, emblemized by the change of name, from ‘St Deiniol’s Library’ to ‘Gladstone’s Library’ in 2010 reflecting the opportunity for the Library to market itself as a Gladstone heritage institution.”¹¹⁴

A Wish Upon A Catalog

The claims made for GladCat’s impact are large and far-reaching, and, in a different, more technological way, the aims of the national and other libraries publishing their collections data in RDF for uptake by the linked data community may also seem quite large. Twenty-first century catalogers are used to big visions – we are living through the era of RDA implementation with Bibframe (and a complete data structure change) on our horizon. In some ways, this article makes a smaller case: the case for the researcher working in a field that is not only of interest to academics but also, presumably, to libraries who want to know more about their individual collections.

While RDF data will allow for progress in the research into writers' libraries and provenance more generally, and while much has been achieved by consortia like OCLC, Copac and CERL in publishing and utilizing XML data based on MARC records, there is still plenty that can be achieved through a researcher's being able to download complete records into a basic format – something as simple as CSV that displays correctly when imported to a spreadsheet would be enough.

To come back to the question in the title of the conference at which this work was originally presented, “A common international standard for rare materials?” the answer this article offers is that the input of data by libraries to their catalogs has not been a stumbling block for researchers – not even when computerization and data sharing was relatively new; not even when catalogs were manual. Research into writers' libraries held now by institutional libraries is dependent on several processes – discovery, access, collocation of materials, consistent catalog input and reliable output. The missed opportunity, it is the contention here, is not input so much as output (beyond search and display).

In some ways, this is a smaller case than the cases for international standards and linked data. In another way, it is much larger, leading us back round to Wilson's philosophy of the exploitative power of bibliographic control. If we can meet the needs of researchers who want to engage with our data not as a route through to 'the real' objects of their research – full-text files, books, the item for which catalog data is a surrogate – but as an integral part of their own research, then, surely, we are assisting not simply in an

“elementary” user task, but something that is fundamental to scholarship: “the best possible use of recorded knowledge.”¹⁵

Acknowledgement

Some of the research in this article was undertaken as part of a PhD in Cultural Studies at UCL. The author would like to thank Gladstone’s Library for the award of a Revd. Dr. Murray MacGregor Scholarship in 2016, which provided her with the peace and space to write up part of her thesis.

ORCID

Anne Welsh <http://orcid.org/0000-0002-5621-7490>

¹ Horace Walpole, *Aedes Walpolinae*. (Rev. ed. 1752), xxxi, quoted in Carla

² Anne Welsh, “Metadata Output and the Researcher” (paper presented at the CILIP Cataloging and Indexing Group Conference, Canterbury, Kent, September 8-10 2014).

³ Anne Welsh. “Metadata Output and the Researcher,” *Catalogue and Index* 178 (2015), 2-8.

⁴ David Pearson, *Provenance Research in Book History: A Handbook*. Reprinted with a new introduction (London: The British Library; New Castle, Delaware: Oak Knoll, 1998), 2.

⁵ C.F. Battiscombe, ed., *The Relics of St Cuthbert: Studies by Various Authors Collected and Edited with an Historical Introduction* (Oxford: Oxford University Press for the Dean and Chapter of Durham Cathedral, 1956), 360.

⁶ Jonathan Rose, “Altick’s Map: The New Historiography of the Common Reader” in *The History of Reading. Volume 3. Methods, Strategies, Tactics*, ed. Rosalind Crone and Shafquat Towheed (Basingstoke: Palgrave Macmillan, 2011), 15-26.

⁷ Pearson, *Provenance Research*, xi.

⁸ Wayne C. Booth, *The Rhetoric of Fiction* (Chicago; London: University of Chicago Press, 1961).

- ⁹ Stanley Fish, *Surprised by Sin: The Reader in Paradise Lost* (London: Macmillan, 1967 ; *Is There a Text in this Class?: The Authority of Interpretive Communities*. Cambridge, Massachusetts: Harvard University Press, 1980).
- ¹⁰ Gerard Genette, *Figures III* (Paris: Seuil, 1972).
- ¹¹ Wolfgang Iser, *The Implied Reader: Patterns of Communication in Prose Fiction from Bunyan to Beckett* (Baltimore; London: John Hopkins University Press, 1974).
- ¹² Walter Ong, "The Writer's Audience is Always a Fiction," *PMLA* 90.1 (1975), 9-21.
- ¹³ Jonathan Culler, *Structuralist Poetics: Structuralism, Linguistics and the Study of Literature* (London: Kegan, Paul, 1975).
- ¹⁴ R.D. Altick, *The English Common Reader: A Social History of the Mass Reading Public, 1800-1900* (Chicago: University of Chicago Press, 1957).
- ¹⁵ Rose, "Altick's Map," 21.
- ¹⁶ Robert Darnton, "What Is the History of Books?" *Daedalus* 111.3 (1982), 65-83.
- ¹⁷ Simon Eliot and Jonathan Rose, introduction to *A Companion to the History of the Book*, ed. Simon Eliot and Jonathan Rose (Oxford: Wiley-Blackwell, 2009), 1.
- ¹⁸ John Feather, "The Book in History and the History of the Book," *The Journal of Library History* 21.1 (1986), 24.
- ¹⁹ James Secord, *Victorian Sensation: The Extraordinary Publication, Reception, and Secret Authorship of 'Vestiges of the Natural History of Creation'* (Chicago: University of Chicago Press, 2001), 3.
- ²⁰ H.J. Jackson, "'Marginal Frivolities': Readers' Notes as Evidence for the History of Reading," in *Owners, Annotators and the Signs of Reading*, ed. Robin Myers, Michael Harris and Giles Mandelbrote (New Castle Delaware: Oak Knoll; London: British Library, 2005), 149.
- ²¹ Robert Darnton. *The Corpus of Clandestine Literature in France, 1769-1789* (New York; London: Norton, 1995).
- ²² William St Clair, *The Reading Nation in the Romantic Period* (Cambridge: Cambridge University Press, 2004).
- ²³ Alberto Manguel, *A Reading Diary* (Edinburgh: Canongate, 2005).
- ²⁴ Jackson, "Marginal Frivolities," 149.
- ²⁵ Robert Darnton, *The Kiss of Lamourette: Reflections in Cultural History* (London: Faber, 1990), 162.
- ²⁶ Leah Price, introduction to *Unpacking My Library: Writers and Their Books*, ed. Leah Price (New Haven, Connecticut; London: Yale University Press, 2011), 3.
- ²⁷ Alan Gribben, "Private Libraries of American Authors: Dispersal, Custody and Description," *The Journal of Library History* 21.2 (1986), 310.
- ²⁸ Richard W. Oram, introduction to *Collecting, Curating, and Researching Writers' Libraries: A Handbook*, ed. Richard W. Oram and Joseph Nicholson (Lanham, Maryland: Rowman & Littlefield, 2014), 1-2.
- ²⁹ Karen Attar, "Books in the Library," in *The Cambridge Companion to the History of the Book*, ed. Leslie Howsam (Cambridge: Cambridge University Press, 2015), 17.
- ³⁰ G. Thomas Tanselle, "A Rationale of Collecting," *Studies in Bibliography* 51 (1998), 2.

- ³¹ Tanselle, "Rationale of Collecting," 3.
- ³² Richard W. Oram and Kevin MacDonnell, "Writers on Their Libraries: Interviews," in *Collecting, Curating and Researching Writers' Libraries: A Handbook*, ed. Richard W. Oram with Joseph Nicholson (Lanham, Maryland: Rowman & Littlefield, 2014), 85-105.
- ³³ Edmund White and Leah Price, "An Interview with Edmund White," in *Unpacking My Library: Writers and Their Books*, ed. Leah Price (New Haven, Connecticut; London: Yale University Press, 2011), 186.
- ³⁴ Edmund White and Leah Price, "Interview," 185.
- ³⁵ Junot Díaz and Leah Price, "An Interview with Junot Díaz," in *Unpacking My Library: Writers and Their Books*, ed. Leah Price (New Haven, Connecticut; London: Yale University Press, 2011), 45.
- ³⁶ Oram and MacDonnell, "Writers and Their Libraries," 104.
- ³⁷ Tanselle, "Rationale for Collecting," 3.
- ³⁸ Cyrus Redding, "Recollections of the Author of *Vathek*," *New Monthly Magazine and Humorist* 71.283 (1844), 307.
- ³⁹ Edward Gibbon, *The Memoirs of Edward Gibbon by Himself*, ed. George Birkbeck Hill (London: Methuen, 1900), 121.
- ⁴⁰ Geoffrey Keynes, *The Library of Edward Gibbon: A Catalogue*, 2nd ed. (Winchester: St Paul's Bibliographies, 1980).
- ⁴¹ Geoffrey Keynes, *The Library of Edward Gibbon*, 3.
- ⁴² Katie Halsey, "'Folk Stylistics' and the History of Reading: A Discussion of Method," *Language and Literature* 18.3 (2009), 231.
- ⁴³ Werner Muensterberger, *Collecting: An Unruly Passion: Psychological Perspectives* (Princeton: Princeton University Press, 1994), 211.
- ⁴⁴ Lev Grossman and Leah Price, "An Interview with Lev Grossman," in *Unpacking My Library: Writers and Their Books*, ed. Leah Price (New Haven, Connecticut; London: Yale University Press), 90.
- ⁴⁵ Gary Shteyngart and Leah Price, "An Interview with Gary Shteyngart," in *Unpacking My Library: Writers and Their Books*, ed. Leah Price (New Haven, Connecticut; London: Yale University Press), 171.
- ⁴⁶ Richard W. Oram and Kevin MacDonnell, "Writers and Their Libraries," 91.
- ⁴⁷ Gribben, "Private Libraries," 301.
- ⁴⁸ Fredson Bowers, *Principles of Bibliographical Description* (Winchester: St Paul's Bibliographies, 1994), 361.
- ⁴⁹ Theresa Whistler, *Imagination of the Heart: The Life of Walter de la Mare* (London: Duckworth, 1993).
- ⁵⁰ Anne Bentinck, *Romantic Imagery in the Works of Walter de la Mare* (Lewiston, NY; Lampeter: Edwin Mellen Press, 2001).
- ⁵¹ "Walter de la Mare Library, Senate House Library," last accessed May 6, 2016, <http://www.senatehouselibrary.ac.uk/our-collections/special-collections/printed-special-collections/walter-de-la-mare-library>
- ⁵² Marion L. Kesselring, *Hawthorne's Reading 1828-1850: A Transcription and Identification of the Titles Recorded in the Charge-Books of the Salem Athenaeum* (New York: New York Public Library, 1949).
- ⁵³ Gribben, "Private Libraries," 304.
- ⁵⁴ Gribben, "Private Libraries," 311.
- ⁵⁵ Keynes, "Preface," 4.
- ⁵⁶ Keynes, "Preface," 4.

- ⁵⁷ Keynes, "Preface to the Second Edition" in *The Library of Edward Gibbon: A Catalogue*, 2nd ed. (Winchester: St Paul's Bibliographies, 1980), 10.
- ⁵⁸ Walter Harding, *Thoreau's Library* (Charlottesville: University of Virginia Press, 1957), 8.
- ⁵⁹ Harding, *Thoreau's Library*, 10.
- ⁶⁰ Michael S. Reynolds, *Hemingway's Reading, 1910-1940: An Inventory* (Princeton, New Jersey: Princeton University Press, 1981), 9.
- ⁶¹ Reynolds, *Hemingway's Reading*, 9.
- ⁶² Welsh, "Metadata Output and the Researcher" (2015)
- ⁶³ H.D. Avram, *The MARC Pilot Project: Final Report on a Project Sponsored by the Council on Library Resources, Inc.* (Washington, DC: Library of Congress, 1968), 1.
- ⁶⁴ H.D. Avram, J.F. Knapp and L.J. Rather, *The MARC II Format: A Communications Format for Bibliographic Data* (Washington, DC: Library of Congress, 1968), 4.
- ⁶⁵ Welsh, "Metadata and the Researcher" (2015).
- ⁶⁶ J.H. Bowman, "OPACs: The Early Years and User Reactions," *Library History* 23 (2007), 317-329.
- ⁶⁷ J.H. Bowman, "OPACs."
- ⁶⁸ James Edel and Adeline Tinter, "The Library of Henry James, from Inventory, Catalogs and Library Lists," *The Henry James Review* 4.3 (1983): 158-190.
- ⁶⁹ Joseph Nicholson, "Cataloging Writers' Private Libraries," in *Collecting, Curating, and Researching Writers' Libraries: A Handbook*, ed. Richard W. Oram with Joseph Nicholson (Lanham, Maryland: Rowman & Littlefield, 2014),
- ⁷⁰ Anne Welsh and Sue Batley, *Practical Cataloguing: AACR, RDA and MARC 21* (London: Facet, 2012), 8.
- ⁷¹ Nicholson, "Cataloging," 29.
- ⁷² Nicholson, "Cataloging," 31.
- ⁷³ Oram and Nicholson, "Location and Bibliographical Guide to Writers' Libraries," in *Collecting, Curating, and Researching Writers' Libraries: A Handbook*, ed. Richard W. Oram with Joseph Nicholson (Lanham, Maryland: Rowman & Littlefield, 2014), 107.
- ⁷⁴ S.R. Ranganathan, *The Five Laws of Library Science* (Madras: Madras Library Association; London: Edward Goldstone, 1931).
- ⁷⁵ IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records Final Report* (Munich: K.G. Saur, 1998), 82.
- ⁷⁶ Victoria Wilson, "Catalog Users 'in the Wild': The Potential of an Ethnographic Approach to Studies of Library Catalogs and Their Users," *Cataloging and Classification Quarterly* 53.2 (2015), 190-213.
- ⁷⁷ IFLA Study Group on FRBR, *Functional Requirements for Bibliographic Records*, 82.
- ⁷⁸ R.P. Smiraglia, "Rethinking What We Catalog: Documents as Cultural Artifacts," *Cataloging and Classification Quarterly* 45 (2008), 25-37.
- ⁷⁹ Katharine Claire Whaite, "New Ways of Exploring the Catalogue: Incorporating Text and Culture," *Information Research* 18.3, <http://www.informationr.net/ir/18-3/colis/paperS09.html#.VwITTXi9ZUQ>

⁸⁰ Jack Andersen, "Materiality of Works: The Bibliographic Record as Text," *Cataloging and Classification Quarterly* 33 (2002), 39-65.

⁸¹ Christos Iraklis Tsatsoulis, "Unsupervised Text Mining Methods for Literature Analysis: A Case Study for Thomas Pynchon's *V*," *Orbit* 1.2 (2012), <https://www.pynchon.net/articles/10.7766/orbit.v1.2.44/>

⁸² "OCLC Worldshare Collection Evaluation Overview," last accessed, May 6, 2016, <http://www.oclc.org/collection-evaluation.en.html>

⁸³ "OCLC Worldshare Collection Evaluation Brochure," last accessed, May 6, 2016, http://www.oclc.org/content/dam/oclc/services/brochures/215186usf_WorldShare_Collection_Evaluation.pdf

⁸⁴ "Worldshare Collection Evaluation Brochure."

⁸⁵ "Copac Collection Management Tools," last accessed, May 6, 2016, <https://ccm.copac.jisc.ac.uk>

⁸⁶ "Copac Collection Management Tools."

⁸⁷ "Shibboleth," last modified, May 4, 2016, <https://shibboleth.net>

⁸⁸ "About CCM Tools," last accessed, May 6, 2016, <https://ccm.copac.jisc.ac.uk/about/>

⁸⁹ "CCM Tools User Stories," last accessed, May 6, 2016, <http://blog.ccm.copac.ac.uk/user-stories/>

⁹⁰ "Early Modern Print: Text Mining Early Printed English," last accessed, May 6, 2016, <http://earlyprint.wustl.edu>

⁹¹ "Heritage of the Printed Book Database," last modified, November 27, 2015, <https://www.cerl.org/resources/hpb/main>

⁹² "Material Evidence in Incunabula," last modified, November 27, 2015, <https://www.cerl.org/resources/mei/main>

⁹³ Karen Attar, "Cataloguing Early Children's Books: Demand, Supply and a Seminar," *Catalogue and Index* 151 (2004), 11.

⁹⁴ Roy Tennant, "MARC Must Die," *Library Journal* 15 October (2002), http://lj.libraryjournal.com/2002/10/ljarchives/marc-must-die/#_

⁹⁵ Eric Miller, Uche Ogbuji, Victoria Mueller and Kathy MacDougall, *Bibliographic Framework as a Web of Data: Linked Model and Supporting Services* (Washington, DC: Library of Congress, 2012).

⁹⁶ Mirna Willer and Gordon Dunsire, *Bibliographic Information Organization in the Semantic Web* (Oxford: Chandos, 2013), xxvii.

⁹⁷ Ngā Pūrongo, "MARC 21 to Bibframe: Outcomes, Possibilities and New Directions," *New Zealand Library & Information Management Journal* 55.1 (2014), 17.

⁹⁸ Agnes Simon, Romain Wenz, Vincent Michel and Adrien Di Mascio, "Publishing Bibliographic Records on the Web of Data: Opportunities for the BnF (French National Library)," in *The Semantic Web: Semantics and Big Data*, ed. Philipp Cimiano, Oscar Corcho, Valentina Presutti, Laura Hollink and Sebastian Rudolph (Heidelberg: Springer, 2013), 565.

⁹⁹ Mitch Fraas. "Charting Former Owners of Penn's Codex Manuscripts," *Mapping Books*, January 24, 2014, <http://mappingbooks.blogspot.co.uk/2014/01/charting-former-owners-of-penns-codex.html>

¹⁰⁰ Jillian Tomm, "The Imprint of the Scholar: An Analysis of the Printed Books of McGill University's Raymond Klibansky Collection" (PhD diss., McGill University, 2012).

¹⁰¹ James Baker, "On Metadata and Cartoons," *British Library Digital Scholarship Blog*, May 16, 2013, <http://britishlibrary.typepad.co.uk/digital-scholarship/2013/05/on-metadata-and-cartoons.html>

¹⁰² Welsh, "Metadata Output" (2015).

¹⁰³ Patrick Wilson, *Two Kinds of Power: An Essay on Bibliographical Control* (Berkeley, California: University of California Press, 1968).

¹⁰⁴ Smiraglia, "Rethinking What We Catalog," 35.

¹⁰⁵ Melissa Terras, James Baker, James Hetherington, David Beavan, Anne Welsh, Helen O'Neill, Will Finley, Oliver Duke-Williams and Adam Farquhar, "Enabling Complex Analysis of Large-Scale Digital Collections: Humanities Research, High Performance Computing and Transforming Access to British Library Digital Collections," *Digital Humanities 2016* (Alliance of Digital Humanities Organizations, 2016).

¹⁰⁶ Simon Mahony and Elena Pierazzo, "Teaching Skills or Teaching Methodology?" *Digital Humanities Pedagogy: Practices, Principles and Politics*, ed. Brett D. Hirsch (Cambridge: Open Book, 2012).

¹⁰⁷ Reynolds, *Hemingway's Reading*, 10.

¹⁰⁸ "Melville's Marginalia Online," last accessed, May 6, 2016, <http://melvillemarginalia.org/front.php>

¹⁰⁹ Merton M. Sealts, *Pursuing Melville, 1940-1980: Chapters and Essays* (Madison, Wisconsin: University of Wisconsin Press, 1982).

¹¹⁰ "Reading Experience Database," last accessed, May 6, 2016, <http://www.open.ac.uk/Arts/reading/>

¹¹¹ J. Keith Davies and Gerhard Fichtner, *Freud's Library: A Comprehensive Catalog* (London: The Freud Museum; Tübingen, 2006).

¹¹² "The Gladstone's Reading Database," last accessed, May 6, 2016, <https://www.liverpool.ac.uk/english/research/gladstone-library/>

¹¹³ "Gladstone's Library Library Catalogues," last accessed, May 6, 2016, <https://www.gladstoneslibrary.org/reading-rooms/library-catalogues>

¹¹⁴ "Gladstone's Library, Gladstone's Reading: Impact Case Study," last accessed, May 6, 2016, <http://impact.ref.ac.uk/casestudies2/refservice.svc/GetCaseStudyPDF/7748>

¹¹⁵ Smiraglia, "Rethinking What We Catalog," 35.