February 2010

The Perfect Library

Alice Keller

Bodleian Library, Oxford University, alice.keller@ouis.ox.ac.uk

The Perfect Library

by **Alice Keller** (Assistant Director, Collections & Resource Description, Bodleian Library, Oxford University, UK) <alice.keller@ouis.ox.ac.uk>

Column Editor's Note: *This paper evaluates the question, what would a perfect library look like and which are the limiting factors when attempting to build such a library? The discussion focuses on well- and less-known images and visions from history, fantasy, and literature. Finally these images are compared to the Internet in order to compare whether services such as* **Google Book Search** *et al. are the perfect library. Particular attention is given to the assessment of the limiting factors of the online library: what are they, and how could they be solved? — AK*

Introduction

Lord Byron (1788-1824) is reported to have said, "I never travel without **Scott's** *Novels*, they are a perfect library in themselves; a perfect literary treasure. I could read them once a year with new pleasure."[1] And indeed, titles like *Ivanhoe, Rob Roy, The Lady of The Lake, Waverley, The Heart of Midlothian* or *The Bride of Lammermoor* certainly made good reading for a gentleman of the 19th century.

But what about the perfect library for a lady of the 19th century? In 1856 the publisher **Peck & Bliss**, Philadelphia, put together an entirely different list of books which was hailed by one of the newspapers as the "perfect library for a young lady."[2] The volumes came in a neat case and included titles such as: *The Flora Offering, The Book of Parlour Games, The Lady's Oracle, The Lady's Mentor, The Lady's Companion, The Vicar of Wakefield,* and *Paul and Virginia.* The advert ends with the following quote: "We know of no books which a father or mother may better bestow upon a daughter."

I don't assume that these are the sort of libraries or collections that the reader thinks of when he or she sees the title of this paper, but it is interesting to see how even a very small collection of volumes can be described as a perfect library.

Plans for the Perfect Library

In this paper I would like to reflect on the theme of the "perfect library" as a professional librarian and explore the history, the stories, and the images around this topic.[3] The history of perfect libraries cannot start anywhere else than with the **Ancient Library of Alexandria**. Our history books tell us that the **Library at Alexandria** was charged with collecting all the world's knowledge, and it can be seen as the first library we know of which gathered books (or the equivalent thereof) systematically from within Egypt and from abroad. It did so through an aggressive and well-funded royal mandate involving trips to the book fairs of Rhodes and Athens and a policy of pulling the books off every ship that came into harbour, keeping the originals and returning copies to their owners. No catalogue of the library sur-

vives, and it is not possible to know with any certainty how large the collection may have been. The ancient accounts agree that **Caesar** accidentally burned the library down during his visit to Alexandria in 48 BC. Clearly the vision and mission of the **Ancient Library of Alexandria** was to collect and hold the world's whole knowledge in one place.

Eighteen centuries later in France an architect had a similar vision of a building holding the entire world's knowledge under one roof. With his drawing "Deuxième projet pour la Bibliothèque du Roi" from 1785, the visionary French neoclassical architect **Etienne-Louis Boullée** designed a gigantic basilica, capable of holding the memory of the whole world. But, sadly, it was never built.
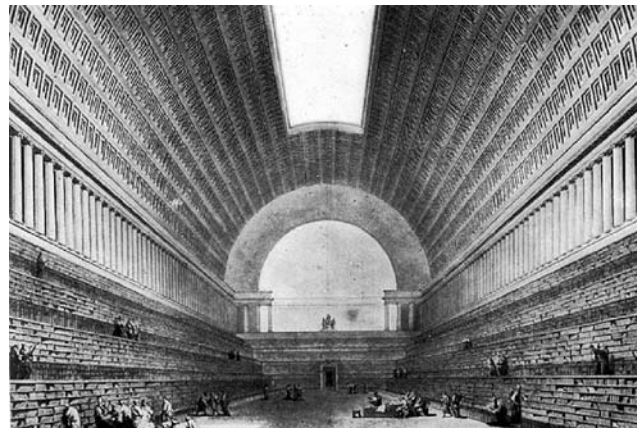

Fig. 1: Étienne-Louis Boullée, *Deuxième projet pour la Bibliothèque du Roi* **(1785).** *http://upload.wikimedia.org/wikipedia/commons/6/6d/Bibliotheque_nationale_boul.jpg*

The author **Alberto Manguel** points out that the design does not suggest a space for private reading or concentration.[4] He suggests it looks more like a space for rapid consultation. To me it is reminiscent of the ruins of Ancient Greece where philosophers held their conversations. However, it is unlikely that the French King would have let the man from the street into his private library.

Nonetheless, the important message for my purpose is that **Boullée** considered it possible to collect all the books of the world not only in one building, but in one huge room.

The next library I want to focus on in my search for the perfect library is the **New York Public Library** which opened its doors to the public on 24th May 1911. A full page image in the *Scientific American*[5] gives us a fantastic cross-section of the building. To me it is a fascinating representation of the vision of an ideal library at the dawn of the 20th century. The library had room for three million volumes and over 1,700 readers. The main reading room is situated on the top floor; the books are arranged in seven levels of stacks, occupying the greater part of the building. In the ante-room to the

main reading room there is a catalogue room with 6,000 catalogue trays. In order to obtain the book ordered by the reader as quickly as possible, a system of pneumatic tubes, lifts and conveyors was installed. The text accompanying the picture in the *Scientific American* tells us how the collections fulfil the requirements of *all* readers: "the man who desires a mere definition, a brief summary, or who desires simply to browse about, is given ample opportunity to satisfy his whims." Equally the scientific investigator, the engineer or the student of Oriental languages will find dedicated reading rooms and specialist collections. It was no longer considered possible, as with **Boullée** in 1785, to have all the books in one room. Also, the **NYPL** did not aspire to have every single book ever printed. However, the library planners did believe that three million volumes could satisfy the requirements of all readers. And these volumes were housed under one roof and delivered within just a few minutes thanks to the clever system of tubes, lifts and conveyors. But — as shown in the next paragraph — this does not mean that the idea of a complete collection of all published output in one location had been completely disbanded.

In 1929 **George Watson Cole** (Librarian of the **Huntington Library**) wrote a book chapter with the title "The Ideally Perfect Library."[6] His idea of a perfect library was a library holding copies of all products of the printing press from the earliest time to his present day. "Had such a collection been undertaken and carried out on the lines here suggested, properly and commodiously housed, systematically arranged and classified, and thoroughly indexed and catalogued by both authors and subjects, it would contain and make available to scholars, everything that could possibly be needed to pursue the investigation of any subject that has ever appeared in print. Such a collection would form an Ideally Perfect Library — a complete reservoir of knowledge so far as it has ever been embodied in print." At this point I would like to give the reader an idea of the number of books thought to be available in the early 20th century: **Cole** refers to **Iwinski** who published the results of an elaborate statistical study which estimated that there were 25 million "different books" in the world in 1911. This is admittedly a very large number, but not larger than what the greatest libraries of our time hold.

I think that despite **Cole's** dreams of an ideally perfect library, it was recognised in the first half of the 20th century that there could no longer be one single library big enough to hold all the books ever published. I see this realisation in the work of the Belgian lawyer **Paul Otlet** (1868-1944) and his colleague **Henri La Fontaine** (1854-1943), who set up the "Universal Bibliographic Repertory" shortly before the First World War.[7] The idea was to create a complete world bibliography, containing references to as many books, films, and sound recordings as possible. Together they established an international centre called "Palais Mondial" and later "Mundaneum" which was later taken over by the **International Federation for Information and Documentation (FID)**.

The Universal Bibliographic Repertory grew rapidly. By 1897 it contained 1.5 million entries. Two years later, the number was nearly three million, rising to nine million by 1912, and ultimately to nearly 16 million in 1930. The card catalogue contained two indexes: an author file and a classified subject file. An international search service was set up, and during 1912 over 1,500 requests for information were received and processed. Just like **Otlet**, I think all librarians of today recognise that it is no longer possible to collect all published output in one physical location. Our dream of more than twenty-one centuries has come to an end without ever having been fulfilled. But the perfect library was not only the dream of librarians; it has also been the theme of many literary works during the last centuries. These visions will be the topic of the next section.

Ideals of the Perfect Library in Literature

Whilst librarians restricted their dreams to the collection of published output, authors of fiction did not have this limitation. Instead, as the following examples will show, their visions did not only include books already written, but also *books yet to be written*. The oldest such vision which I came across in my literature search is from **Jonathan Swift's** *Gulliver's Travels*, in which **Gulliver** visits Glubbdubdrib, an island populated by sorcerers.[8] Here the author is permitted to visit the grand **Academy of Lagado** where a professor has invented a machine which generates knowledge and content for books in a semi-automated fashion. "Every one knew how laborious the usual Method is of attaining to Arts and Sciences; whereas by his Contrivance, the most ignorant Person at a reasonable Charge, and with a little bodily Labour, may write Books in Philosophy, Poetry, Politicks, Law, Mathematicks and Theology, without the least Assistance from Genius or Study."
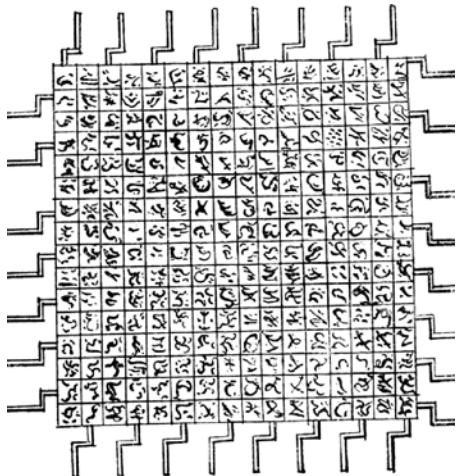

The literary machine consisted of a frame which held blocks of wood linked together with wires and covered with words in all moods, tenses and declensions. Young students held forty iron handles which were connected to the frame and the blocks of wood, and at the professor's command gave them a sudden turn. The layout of words in the frame changed and the students were asked to search for three or four words together that might make part of a sentence. Whenever they found such a fragment, they dictated it to four other boys who acted as scribes. The Professor suggested that if funds for five hundred such frames could be found "a compleat Body of all Arts and Sciences" could be produced and made available to the public.

Fig.2: Drawing of the literary machine described in Jonathan Swift's *Gulliver's Travels*, 1726. *http://www.jaffebros.com/lee/gulliver/bancroft/10.jpeg*

Swift clearly recognises that the limiting factor for human progress is human knowledge, or rather the laborious and slow process by which human knowledge is "produced." Using his ingenious literary machine he could automate the knowledge generation process and build a library comprising this complete body of all such knowledge.

Two different and yet similar approaches to building a library with *books yet to be written* were put forward in the first half of the 20th century. **Kurd Lasswitz's** short story "Die Universalbibliothek" was first published in Berlin in 1904.[9] He describes how four friends come up with the bright idea of a library which contains every single book imaginable. Their ideal book has 500 pages, 40 lines with 50 characters each, and any combination of 100 different symbols (letters, capitals and lower case, numbers, mathematical symbols, punctuation, and — most importantly — spaces). Every single book that could ever possibly be written with these specifications is held in this huge collection, so all needs are provided for. **Lasswitz**, a mathematician himself, spares the reader the effort of calculating the size of this library: his universal library would contain $10^{2\text{mio}}$ volumes.[10]

A similar but much better known thought experiment originates from the Argentinean poet-librarian **Jorge Luis Borges**. It does not include any calculations, but instead gives us some visual aids which have captured the imagination of many artists. In his short story *The Library of Babel*,[11] the ideal library includes every single book which can be written with an alphabet of 25 orthographic symbols (22 letters, comma, full stop, and space) and on 410 pages. This library would contain everything — I mean every single book ever written or yet to be written: "Everything is there: the minute history of the future, the autobiographies of the archangels, the faithful catalogue of the Library, a demonstration of the fallacy of these catalogues … the Gnostic gospel of Basilides, the commentary on this gospel, the commentary on the commentary of this gospel, the veridical account of your death, a version of each book in all languages."

Borges summarizes the initial reaction: "When it was proclaimed that the Library comprised all books, the first impression was one of extravagant joy. All men felt themselves lords of a secret, intact treasure." But on second thought, this library also has its disadvantages. *The Library of Babel* also contains every nonsense imaginable: a completely empty book (only spaces), a book full of question marks, and any combination of true and wrong facts. **Lasswitz** also picks up on this point in his Universal Library: "The worst thing is when you think you have found a seemingly sensible volume. For example, you want to look up something in 'Faust' and you actually find the volume with the correct beginning. And when you have read on a bit, you suddenly find it continuing with [nonsense]." It is not difficult to realise that such a collection is completely useless; or even worse, it is misleading and dangerous.

I think that these examples of **Swift**, **Lasswitz** and **Borges** show very clearly that a collection can only be built at the speed at which human knowledge is acquired and processed. Books cannot be published in advance of us reaching that point of understanding. The limiting factor here is not space or money, but human knowledge and our ability to discern right from wrong.

Comparing these Ideals with the Internet

After reviewing visions of the perfect library in history and literature, it is now time to consider the opportunities offered by the Internet. After all, does not the Internet provide an ideal framework and platform for the perfect library?

Compared to the traditional print library, the Internet has many advantages. It does not suffer space problems, it is available 24x7 from any desktop, much of the content is available free of charge, and full texts are searchable. So is it the perfect library?

The Internet in the Real World

But before speaking about online libraries, I would like to mention a newspaper article I read earlier this year which encapsulates many features of the Internet transported into the physical world. On February 28, 2009 *The Daily Mail* reported how, after an **Amazon** supplier in Bristol abandoned its warehouse, the public was given free entry

to rummage the site and take away any books they wanted for free.[12] "Yesterday, Porsches and BMWs were parked alongside vans outside as the scavengers carried out their finds in crates and on trolleys. In an endearing display of how the British love the offer of something for nothing, someone had even come with a small trailer on the back of his car. Others were seen stacking books in prams." This picture combines many elements of the Internet: access was unrestricted and free; you could take any and as many books as you wanted for free; the quantities were vast; it was totally self-service; there were no rules and no order. And whilst the threshold to content was lowered to an absolute minimum, disorder was increased to a maximum — and the public loved it! What does this show us? I think it explains that while for us librarians a perfect library obligatorily includes the features of order and tidiness, the public loves randomness and chaos. And it is, of course, precisely this randomness and chaos which stands out for us librarians as one of the limitations — or the benefits(?) — of the Internet.

Limitations of the Library on the Internet

But in order to get a better grasp of the limitations of the library on the Internet, I would like to reflect on two recent articles of **Johanna Drucker** and **Robert Darnton**.

Johanna Drucker, Professor of Information Studies at the **University of California at Los Angeles**, takes the view of a scholar in the humanities.[13] Her point is that digitisation of content is not purely a technical matter of access. The online library cannot simply be created by mass digitization of printed material; what we need is a well designed digital working environment which serves the needs of scholarship. Otherwise "we will find ourselves in a future that doesn't work, without the methods and materials essential to our undertakings." One of the issues **Drucker** reflects on is the choice of copy for digitization: which version of a text should be digitized as a representative of a work? Or are there even differences between different physical items? For example, what about the copy-specific notes which can contain information of crucial importance to the future generation of scholars?

A good example of such research can be seen in the work of **Peter Thompson**, Lecturer in Early American History at the **University of Oxford**. His research focuses on the marginalia in a 1649 tract "Virginia Impartially Examined," where a single, relatively neat, seventeenth-century hand either strongly agrees or disagrees with the content of the tract.[14]

Most other scholars who do research on this tract will use the copy available on the Internet via **Early English Books Online**. But this surrogate is based on the

Fig. 3: Handwritten marginalia in William Bullock's *Virginia Impartially Examined* (1649), Bodleian Library.

British Library's copy which does not contain the additional information which comprised an essential part of **Thompson's** research. Indeed, **Drucker** is right: if all research is based on one identical digitized copy, it will lack the vibrancy which can be achieved by comparing different copies which have each undergone different treatment by past generations of readers.

This issue of quality, choice, and completeness was explored by **Paul Duguid** in his paper "Inheritance and loss?"[15] **Duguid** observes several shortfalls in **Google's** digitization programme. He emphasizes

Prior to **OCLC**, he spent seven years in the soft drink industry including marketing positions with **The Coca-Cola Company**. He has a BS from the **Georgia Institute of Technology** and an MBA from the **University of Michigan**. What a guy!

http://www.lyrasis.org

Got the most wonderful greeting card the other day from Bishop, Georgia. It wasn't one of those electronic virtual thingies, it was a real card with real handwriting and typing and pictures and all that. It was from the incredibly amicable **Danny Overstreet** and his partner, **Monica Lynch**. **Danny** is still looking for the right job. Any employers out there?

Speaking of snail mail, got a copy of the **2007/2008 annual report** of **The U.S. Naval Academy's Nimitz Library** from the incredibly clever and smart **Richard Hume Werking**, Library Director/Associate Dean for Information. Knowing that my son **Raymond** graduated from **West Point** several years ago, **Richard** thought I'd be interested in the building of the new library at the **U.S. Military Academy**. The library is called **Jefferson Hall** in honor of the third President of the U.S. who established the Military Academy in 1802.

Speaking of history, was interested to see that the awesome **David Ferriero**, the 10th Archivist of the U.S., is the first librarian to serve in this position. The **National Archives** contains ten billion items (monumental to miniscule). **Ferreiro** says, in an interview with **Liane Hanson** of **NPR**, that one of the most exciting things that he has seen there is a check for $7.2 million to purchase Alaska from the Russians.

http://www.npr.org/templates/story/story/php?storyId=122416367

Moving right along in history, I have been continually captivated **by Robert Darnton's** *The Case for Books* (Public Affairs, 2009). It was especially interesting to read his discussion of **Nicholson Baker's** *Double Fold* ("A Paen to Paper," pp. 109-129). One wonders how what libraries and librarianship are doing now with the printed and virtual record of humanity will be judged twenty-five or fifty or even a hundred years from now. Speaking of which**, Robert Darnton** will be one of our keynote speakers at the **2010 Charleston Conference**!

www.katina.info/conference

Speaking of the **Charleston Conference,** just got an email from the incredible **Edna Laughrey**, **Edna** said that a recent article in *The Charleston Report* by **Bob Holley**

<aa3805@wayne.edu> ("Ineffective Advertising," **TCR**, v.14#4, p.1) hit the nail on the head. Check it out.

www.charlestonco.com/

And, speaking of **Edna** <elaughrey@verizon.net>, we are trying to get her to come to the **30th Charleston Conference** this November. When you see her, tell her to do that or send her an email! Much fun will be had by all.

Returning to the **Charleston Conference**, was corresponding with the awesome **John Tagler** <jtagler@publishers.org> (Vice President & Executive Director of the **AAP/Professional and Scholarly Publishing Division**) about a recent press release. Nine million illegal downloads of copyright-protected books were documented during the closing months of 2009, according to a new study. The independent study, conducted by the online monitoring and enforcement service **Attributor**, looked at illegal downloads of 913 popular titles. On average, each of the titles tracked was downloaded approximately 10,000 times. "…Those nine million pirated books should be a call-to-arms for policymakers, educators, and every reader who cares about the future of digital and printed books," said **Tom Allen**, President and CEO of the **Association of American Publishers**. The study can be found online at *http://www.attributor.com/docs/Attributor_Book_Anti-Piracy_Research_Findings.pdf*.

that the intellectual tasks of vetting editions and assessing scholarly value for future generations should be taken into account from the very design of the project, not reverse-engineered later. "Even with some of the best search and scanning technology in the world behind you, it is unwise to ignore the bookish character of books." Following the logic of **Drucker** and **Duguid** the Internet can only provide a perfect library if it either includes a rigorous selection process — or if there are sufficient resources to scan every single print copy of a work in order to replicate not only the work on the Internet, but the actual physical item with its copy-specific features.

But the library on the Internet also has other much more serious gaps. The serious problem of the actual gaps in content is described by **Robert Darnton**, **Carl H. Pforzheimer University** Professor at **Harvard**, in his paper "**Google** & the Future of Books."[16] According to copyright legislation, digitized content enters the public domain seventy years after the author's death. In the United States this means that free access to the cultural heritage ends on January 1, 1923, the date from which great numbers of books are subject to copyright laws.[17] In Europe, **Google** took a much more conservative approach and calculated the 70 years after the author's death to mean a 140-year moving wall, which pushes the end date for the public domain material back to 1869 (for 2009). This means that books which are freely available on **Google Book Search** in the US are not available to the public in the UK.

To **Darnton**, restricted access due to copyright legislation is the fundamental flaw and the great limitation of the Internet as a library. His conviction is that libraries exist to promote a public good: "the encouragement of learning," learning "Free To All." He is not totally opposed to the for-profit business approach — after all, the public good depends on a profitable economy — however, if we permit the commercialization of the content of our libraries, there is no getting around a fundamental contradiction.

Yes, libraries must digitize, but not on any terms. It must be done in the interest of the public, in order to give the public open access to our cultural heritage. The way to do so, according to **Darnton**, has to be by rewriting the rules of the game, by subordinating private interests to the public good: "if we get the balance wrong at this moment, private interests may outweigh the public good for the foreseeable future, and the Enlightenment dream may be as elusive as ever." It is truly tantalizing. We are on the brink of creating something like a perfect library on the Internet; the vaults of the network hold millions of pages, but we don't have the key to access them.

Conclusion

So, what is the conclusion of my review of history, fantasy, and reality? I still have not answered the question of what the perfect library looks like. I could refer to **George Watson Cole** for an answer: "We must therefore conclude that the best all-round library is that in which the scholar is most likely to find the greatest number of answers to such questions as may arise in connection with his work."[18] But this is not a true answer; it is rather a loophole to avoid answering the key question. In the physical world, the limitations to build a perfect library are space constraints and lack of resources, but also the fact that throughout history so many published works have been lost to humanity through negligence, accident, or wilful intention. In many ways the Internet addresses the issue of space and resources. But now it seems that copy-specific selection and copyright legislation are preventing us from building the perfect library.

There remains the option of **Swift**, **Lasswitz**, and **Borges**: we could use computer power to produce all the books ever written and yet to be written and display them on the Web. Here that limitation is our incomplete human knowledge, or rather our limitations in discerning right from wrong. And it is precisely in order to discern right from wrong that we are building libraries in the first place. In many ways it is only through humanity's effort to build the perfect library that we will ever reach it — if ever. The perfect library is both a means to an end and an end itself.

I would like to finish with a poem called "The Perfect Library" by **David Drake**[19] which encapsulates very well the problem that the perfect library will always remain at least one step ahead of us.

Some day we may see
The perfect library,
But it seems to me
We should be quite wary
Of claiming perfection
For any one yet –
Who knows, on reflection,
How good we can get? 🐾

Endnotes

1. *The London Literary Gazette* and *Journal of Belles Lettres, Arts, Sciences, Etc.* (London, 1824): pp. 659.

2. **T.S Arthur**, *Friends and Neighbours, or, Two Ways of Living in the World* (Philadelphia: Beck & T. Bliss, 1856) [301].

3. My perspective is from the collection's point of view — not the services point of view. I entirely agree that you would need perfect services in a perfect library, but within the context of this conference, I am viewing the perfect library exclusively from the point of view of the collection.

4. **Alberto Manguel**, *The Library at Night* (New Haven: Yale University Press, 2008), 139.

5. *Scientific American* (27 May 1911).

6. **George Watson Cole**, "The Ideally Perfect Library," in: Essays offered to **Herbert Putnam**: by his colleagues and friends on his thirtieth anniversary as librarian of Congress, 5 April 1929, ed. by **William Warner Bishop** and **Andrew Keogh** (New Haven: Yale University Press, 1929), 113-127.

7. **W. Boyd Rayward**, The Universe of Information: the Work of **Paul Otlet** for Documentation and International Organisation (VINITI for the International Federation for Documentation, 1975); available from *http://hdl.handle.net/2142/651*; accessed 12 April 2009.

8. **Lemuel Gulliver**, *Travels into Several Remote Nations of the World*; or *Gulliver's Travels* by **Jonathan Swift**, first published London, 1726 (amended 1735). Part III: A Voyage to Laputa, Balnibarbi, Luggnagg, Glubbdubdrib and Japan, Chapter V; available from *http://www.jaffebros.com/lee/gulliver/*; accessed 12 April 2009.

9. **Kurd Lasswitz**, Die Universalbibliothek (Berlin 1904); available from *http://gutenberg.spiegel.de/*; accessed 12 April 2009.

10. **Lasswitz** calculates that if a librarian flew along the shelves at the speed of light, it would take him two years to pass a quintillion volumes (that is a 1 with 18 zeros).

11. **Jorge Luis Borges**, *The Library of Babel*, (transl. **James Irby**, 1962); available from *http://jubal.westnet.com/hyperdiscordia/library_of_babel.html*; accessed 12 April 2009.

12. "Thousands in scramble for free books after **Amazon** supplier abandons warehouse," Mail Online, 28 February 2009; available from *http://www.dailymail.co.uk/news/article-1156973/Thousands-scramble-free-books-Amazon-supplier-abandons-warehouse.html* accessed 12 April 2009.

13. **Johanna Drucker**, "Blind Spots: Humanists must plan their digital future," *The Chronicle of Higher Education*. Section: The Chronicle Review 55, Issue 30 (2009): B6; available from http://chronicle.com/free/v55/i30/30b00601.htm; accessed 12 April 2009.

14. **Peter Thompson**, "William Bullock's 'Strange Adventure': A Plan to Transform Seventeenth-Century Virginia," *The William and Mary Quarterly* 61, no. 1 (2004): 16; available from *http://www.historycooperative.org/journals/wm/61.1/thompson.html*; accessed 12 April 2009.

15. **Paul Duguid**, "Inheritance and loss? A brief survey of Google Books," *First Monday* 12, no. 8 (2007); available from *http://outreach.lib.uic.edu/www/issues/issue12_8/duguid/*; accessed 12 April 2009.

16. **Robert Darnton**, "Google & the Future of Books," *The New York Review of Books* 56, no. 2 (February 12, 2009); available from *http://www.nybooks.com/articles/22281*; accessed 12 April 2009.

17. **Darnton** explains how this precise date was arrived at in a footnote.

18. **George Watson Cole**, *The Ideally Perfect Library*, 127.

19. **David Drake**, *Each of us is a book: Poems for the library minded* (Jefferson, NC: McFarland & Co., 2003), 38.