

[Early English Books Online \(an Internet service offered in a variety of institutional access packages, typically \\$31 250 for an annual subscription, \\$171 000 one-time payment for permanent access\) http://www.lib.umi.com/eebo](http://www.lib.umi.com/eebo)

Gabriel Egan, Shakespeare's Globe London, ge@totus.org
John Jowett, Shakespeare Institute, University of Birmingham, j.d.jowett@bham.ac.uk

John Jowett and Gabriel Egan. "Review of the Early English Books Online (EEBO)." *Interactive Early Modern Literary Studies* (January, 2001) 1-13: <URL: <http://purl.oclc.org/emls/iemls/reviews/jowetteebo.htm>>

1. *Early English Books Online* (EEBO) is an online version of the University Microfilms International (UMI) microfilm products *Early English Books 1475-1640* and *Early English Books 1641-1700*. It is already a major resource for the study of early modern literature, culture, and history. When complete it will provide digitized images of every page of the works listed in Pollard & Redgrave's *Short-Title Catalogue* (STC) of printed materials published in the English language from 1475 to 1640, Wing's *Short-Title Catalogue* of works dating from 1641 to 1700, and the *Thomason Tracts 1640-1661* collection of Civil War publications.
2. Searches can be simple or advanced, the latter allowing use of the Boolean operators 'AND', 'OR', and 'NOT'. The search fields include 'author', 'subjects' (by Library of Congress subject headings), 'title', 'keywords' (which combines 'title' and 'subjects'), 'reel position', 'STC number', and 'imprint'. A comprehensive Help section provides excellent guidance. Consultation of Help can be necessary, for instance to identify search fields such as 'DATE' that are not displayed on the Advanced Search screen. Sometimes the searches proceed by mysterious operations. For instance, even though the imprint search is supposed to cover the stated date of publication, the search for a work by Shakespeare with an imprint of 1623, 'PUB(1623) AND AU(shakespeare)', produces no results. The expected book, the 1623 Folio, is located by altering 'PUB' to 'DATE'. The difference between the two fields is that 'PUB' supposedly retrieves information that actually appears on the imprint, as '1623' does in the case of the Shakespeare Folio, whereas 'DATE' reflects the actual date of publication where there is a difference. This is a recurrent failure of the imprint ('PUB') field.
3. Curiously, the successful search for the Shakespeare Folio also retrieved the 1623 reprint of Thomas Lodge's *Euphues*. Further enquiry shows that *Euphues* has under its subject heading 'Shakespeare, William, 1564-1616.—Sources'. In some sense this is how Shakespeare becomes identified as author, though the slippage between fields is unexpected.
4. Users will of course need to bear in mind the vagaries of early modern spelling when conducting author, title, or keyword searches. A truncation facility for leaving unspecified the end of a word is helpful here; for instance 'TI(owl?)

almana?') successfully identifies the title *The owles almanacke* without the need to check the spelling variants for the end of each word. But truncation cannot assist with forms such as 'drad' for 'dread', although more sophisticated search software can accommodate such complexities. This example can be stated as 'dr' followed by 'a' or 'e', in any numbers or combinations, followed by 'd', and in the underlying software of search engines — built using the computational commonplace of 'regular expressions' — such searches are quite routine. The difficulty which no manufacturer has overcome is the provision of an easily-understood grammar and syntax for such search expressions; current grammars cannot even agree whether '?' or '*' should be the symbol used as the wildcard in truncation expressions.

5. EEBO's citation file derives from the same source as the English Short Title Catalogue (ESTC), an electronic database containing records for all the items in Pollard and Redgrave's and Wing's STCs^[1]. ESTC is not immune to error, and this can make searching difficult or impossible in a significant number of cases. One example we noted is that Cervantes' *Exemplarie Novells* is entered as 'Exemparie nouells', where each word contains a typographical error. Similarly, the title of the 1613 edition of Montaigne's *Essays* is given as 'Essays vwritten' instead of 'Essayes written', an error of spelling in the first word and of typography in the second. Where transcription is accurate, the user conducting searches needs to be mindful of typographical conventions such as the representation of 'vv' for 'w' as the two letters.
6. Once a book is located, the user can proceed to a full citation, illustrations, or document images. The full citation gives bibliographical information following that in ESTC, though it has been redesigned to meet the needs of the project. Where ESTC issues notes on a number of different copies, in EEBO details of only the copy microfilmed are provided. The editing has resulted in a few oddities; for instance the note on the format has often (though not always) been deleted, leaving a hanging semicolon before the point where it originally appeared. Format is in fact crucial information, as there is often no other indication of the physical size of the book in question.
7. Our attempts to consult several titles were met with the Bodleianesque message 'Could not load image from vault'. Where there is no such obstacle, loading is quick enough if the image has been requested previously, but it is currently common to receive the message:

Congratulations! You're the first user in the world to request this page image. Please wait while the system moves it from CD-ROM to the server for you to view (this may take up to five minutes.)

Naturally, the more collective use is made of EEBO the less this notice will need be issued. In the meantime, delays may be expected.

8. To receive the digital images of a book's pages requires software which the average user is unlikely to possess: the DejaVu plug-in for Netscape Navigator and Internet Explorer, or Adobe Acrobat Reader. The former allows the images to be seen within one's web-browser window, while the latter allows selected images (or even an entire early printed book) to be downloaded in the Portable

Document Format (PDF), which is rapidly gaining ground in the printing industry, ready to be viewed off-line. The DejaVu plug-in and Adobe Acrobat Reader are available for free download over the Internet, but corporate users such as a libraries may well have to make hard decisions about the computers on which they choose to install such non-standard software and who has responsibility for keeping the software up-to-date. (It is increasingly common for free 'reader' software, although not actually time-limited by design, to 'expire' when it receives content which is significantly newer than the reader itself and then attempt to load a more recent copy of itself from the manufacturer's website. Such attempts generally fail on public-access machines which have been set-up to prevent installation of new software by public users.)

9. Compared to winding a microfilm, using the DejaVu plug-in makes the archive slow and cumbersome to navigate. Locating a particular page in a long book can be a hit-and-miss affair as one can only 'Go To' a particular image number, the numbers being simply a serial count of the frames in the original microfilm. A system of cross-referencing to page numbers and signatures in the printed book itself would hugely improve the utility. By contrast to the DejaVu plug-in, one can move rapidly through a downloaded PDF file read using Adobe Acrobat Reader because the data is loaded directly from one's hard drive rather than travelling from EEBO's webserver screen-at-a-time. During the period of review the DejaVu software was revised by its authors and greatly improved. We found that early versions of the plug-in were unable to print images, producing a Windows 'Exception Violation' after an extended and unproductive delay while it considered the task. The current version, 3.2, prints reliably and can scale the image to fit the paper, a necessary expedient when printing to standard A4 (or 8.5" x 11") office paper. If one selects 'Printer Properties' from within the DejaVu print dialogue box one can set the paper orientation to 'Landscape', which suits the 'short and wide' shape of most verso-recto book images. Adobe Acrobat Reader also printed reliably, although even after changing the application's Page Setup 'Orientation' setting to 'Landscape', and changing to 'Landscape' mode within the Windows printer driver, the images were always produced in 'Portrait' mode on our office Lexmark Optra R+ laser printer, thus making poor use of the paper and crowding the images unnecessarily.
10. Educators seeking to enliven a presentation might well want to 'cut & paste' images into another document such as a PowerPoint presentation. Early versions of the DejaVu plug-in provided no way of doing this, but version 3.2 has an 'Export to File' option on its right-click menu. The only file format DejaVu can export to is Windows bitmap (.BMP), typically producing a 30-40 megabyte file which was too large to edit on an office 333MHz Pentium PC with 128 megabytes of RAM. The version of Adobe Acrobat Reader we used, 4.0, offers no export facility and makes inoperable the 'Copy' and 'Paste' items on the application's 'Edit' menu. Experienced users of either DejaVu or Acrobat Reader can still maximize the image window as large as their monitors allow and then use the little-known 'ALT+PrintScreen' facility to dump the screen image to the Windows clipboard for further manipulation within a graphics package (such as removal of the window border to leave just the microfilm image within), but it seems mean to deny the average user this facility by omission of a more intuitive method.

11. The images themselves derive from the master copies of the EEB microfilms, and although film can, albeit to a limited extent, register greyness, the digitization process used 1 bit-per-pixel (0 for a white dot, 1 for a black dot) so all shades of grey on the film are lost. The scanning resolution was 400 dots-per-inch which produces reasonably smooth edges to letters: jaggedness appears only when the images are greatly magnified. On a standard 15 inch office monitor the zoom facility is usually necessary as the text is difficult or impossible to read at the standard image size. Whether using DejaVu or Adobe Acrobat Reader the images occasionally appear with heavy horizontal striations which can be made to disappear by moving the image up and down slightly; only experience revealed that these were not, as we first thought, scratches on the original microfilm.
12. UMI is now part of Bell & Howell, who also recently acquired Chadwyck-Healey, makers of the Literature Online (LION) database [2] which contains electronic texts of every work of poetry, drama, and prose fiction published before 1910. Although the items in the Pollard and Redgrave and Wing STCs are, of course, not limited to literary works (their selection criteria is 'whatever gets published' not 'genre'), many of their records are for literary works and there are potentially enormous scholarly benefits to the linking of LION and EEBO. With a full-text search, one could locate any work of literature in the period by its contents (LION indexes down to individual words in the body text) and if it was printed before 1700 there is sure to be a digital image of the work's early printing in EEBO. This opportunity has not escaped Bell & Howell. As this review was being completed, the 'What's New' section of Literature Online (updated 15 December 2000) announced that records in LION are being linked to their corresponding page images in EEBO and that over 4500 potential correspondences have been identified. Although users "will still have to log into the service [EEBO] as usual" a hypertext link would simplify movement from one database to the other. In January 2001 we could not get this feature of LION to work. For a number of LION records concerning Thomas Middleton (including the plays *A Game at Chess* and *A Mad World My Masters*) we found a promising looking EEBO icon after the bibliographical data in the 'Summary of results' listing for a 'Find authors' enquiry. The icon was also present in the listings for 'Find works' and 'Browse authors and works' enquiries, but in all cases the icon was merely decorative and not a hypertext link to the images in EEBO. When reaching the same two plays by another route into the LION database — the 'Search text' feature — the EEBO icon did not appear at all. It seems that the linking of LION and EEBO has only just begun and cannot yet be used effectively.
13. We have dwelt on mostly minor glitches. With a little perseverance EEBO is an invaluable research tool. One of its great benefits is that it allows immediate transition from searching for information to examining the relevant texts. In other words, it combines the facilities of ESTC and the UMI microfilm series in one resource. Whether this facility will justify purchase by institutions that already hold the microfilms is doubtful, but libraries who do not have the microfilms should certainly be urged to obtain access to this resource. The potential for generating new research in early modern studies is considerable indeed. If proper integration with LION can be achieved, the potential usefulness will be greater than the sum of these two databases. The relationship between

'literature' as an idealized conceptual category and its manifestation in particular textualizations is currently a hot topic in theoretical work on the early modern period, and electronic products such as EEBO and LION enable new forms of scholarly study which were not possible using paper and film technologies.

Notes

1. The English Short Title Catalogue has been reviewed in this journal. See E. Thomson Shields. "Review of the English Short Title Catalogue." *Interactive Early Modern Literary Studies* (May, 2000) 1-10: <URL: <http://purl.oclc.org/emls/iemls/reviews/shieldsestc.htm>>.

2. Various parts of Literature Online, some of which are also available in CD-ROM format, have been reviewed in this journal. See Hardy M. Cook "Review of the Chadwyck Healey Editions and Adaptations of Shakespeare" *Interactive Early Modern Literary Studies* (January, 2001) 1-12: <URL: <http://purl.oclc.org/emls/iemls/reviews/eas/cookeas.htm>>; Hardy M. Cook "Review of Chadwyck-Healey The Bible in English Internet and CD-ROM databases" *Interactive Early Modern Literary Studies* (June, 2000) 1-10: <URL: <http://purl.oclc.org/emls/iemls/reviews/cooklion.htm>>; and Catherine Alexander "Review of Chadwyck-Healey *English Poetry, Early English Prose Fiction, and English Verse Drama* CD-ROM databases." *Interactive Early Modern Literary Studies* (October, 1999) 1-16: <URL: <http://purl.oclc.org/emls/iemls/reviews/alexlion.htm>>.

Works Cited

- Anon. *The Owles Almanacke*. London: Edward Griffin for Laurence Lisle, 1618
- Cervantes Saavedra, Miguel de. *Exemplarie Novells*. London: John Dawson for Ralph Mabbe, 1640.
- "ESTC—English Short Title Catalogue." *English Short Title Catalog*. Web Page. URL: <http://www.ohio.rlg.org/estc.html> 20 November 2000.
- Lodge, Thomas. *Euphues Golden Legacie*. London: for John Smethwick, 1623.
- Middleton, Thomas. *A Game at Chess*. London?: n. pub., 1615?
- Middleton, Thomas *A Mad World My Masters*. London: H[enry] B[allard] for Walter Burre, 1608.
- Montaigne, Michel de. *The Essays*. London: M. Bradwood for Edward Blount, 1613.
- Pollard, A. W., and G. R. Redgrave. *A Short-Title Catalogue of Books Printed in England, Scotland, and Ireland and of English Books Printed Abroad, 1475-1640*. London: The Bibliographical Society, 1926.
- Wing, Donald Goddard. *Short-Title Catalogue of Books Printed in England, Scotland, Ireland, Wales, and British America, and of English Books Printed in Other Countries, 1641-1700*. 3 vols. New York: Index Society, 1945-51.