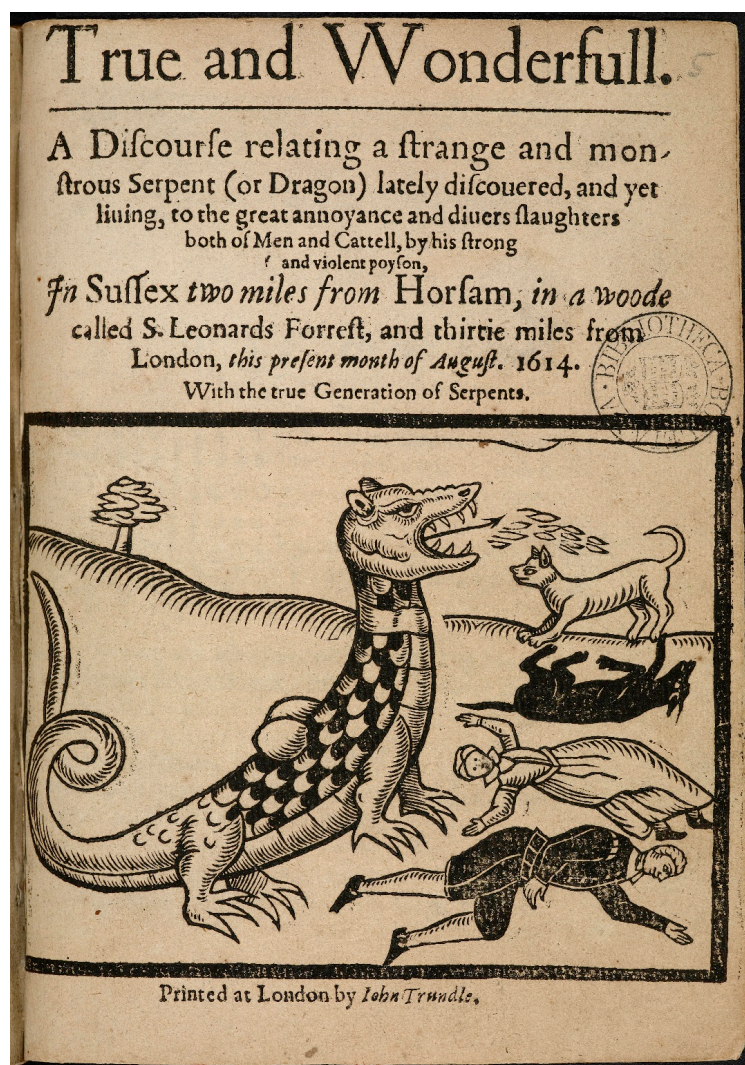


## EARLY ENGLISH BOOKS ONLINE



### A TRAINING WORKSHOP

Centre for Digital Scholarship

27 January 2017



Bodleian Libraries  
UNIVERSITY OF OXFORD

An earlier version of this document was produced by the Oxford EEBO-TCP digital editors to assist Bodleian Libraries' reading room staff. For more information, see the Oxford EEBO-TCP website, or email

[eebotcp@bodleian.ox.ac.uk](mailto:eebotcp@bodleian.ox.ac.uk).

# Table of Contents

Early English Books Online Text Creation Partnership	1
What is EEBO?	1
What is in EEBO?	1
What is EEBO-TCP?	1
Viewing the texts in EEBO	1
Bibliographic records	1
Based on MARC records	1
Tips	1
Document images	1
Illustrations	1
Tips	2
Thumbnails	2
Full text	2
Accessing the content	2
There are two methods to access the contents of the database: search or browse. Browsing is by author, which can be particularly useful if a name is spelt in various forms.	2
Basic searching	2
Boolean searching	2
Proximity searching	2
Keywords	2
Author keyword	2
Title keyword	2
Subject keyword	3
Bibliographic number	3
Variant searching	3
Tips	3
Advanced searching	3
Other bibliographic fields: illustration search	3
Other bibliographic fields: UMI collection	3
Using the material	3
What else can you do with the content?	4
Advanced EEBO: different interfaces for searching the full text	4
Using EEBO-TCP texts for your own research	4

# Early English Books Online Text Creation Partnership

## What is EEBO?

[Early English Books Online](#) (EEBO) is a database owned by ProQuest (which now owns Chadwyck Healey). EEBO is available to subscribing institutions.

## What is in EEBO?

Works covered by the Short-Title Catalogues (STC), essentially all works printed in England or in English from 1473 to 1700. It covers a vast range of different subjects in many languages.

## What is EEBO-TCP?

[Early English Books Online Text Creation Partnership](#) (EEBO-TCP) is a partnership between ProQuest and more than 150 universities and libraries around the world, led by the [University of Oxford](#) and the [University of Michigan](#). It is a not-for-profit partnership which makes highly accurate, fully searchable, XML-encoded digital editions of works on the EEBO corpus.

## Viewing the texts in EEBO

There are four ways to view the books in EEBO: through the bibliographic records, the document images, the full digital text, and thumbnail images.

## Bibliographic records

There is a bibliographic record for every edition recorded in Pollard and Redgrave's Short-Title Catalogue (1475–1640) and Wing's Short-Title Catalogue (1641–1700). One copy of nearly all of these editions (127,074 out of 128,070) is available in digital form. ProQuest continues to scan the remaining books and update the database.

## Based on MARC records

- Uses [Library of Congress Subject Headings](#) – basis of subject search
- The source library from which the microfilm was made is supplied
- If microfilm is available, it gives the reel number for this particular copy

## Tips

The catalogue records on EEBO are based on the STC, and have been updated throughout EEBO's existence. Where errors such as misattributions or misdatings have been corrected, there may be discrepancies with the STC or [English Short-Title Catalogue](#) (ESTC).

## Document images

Most images are scanned from microfilm, and are converted from greyscale to bitonal when digitized. This can affect their clarity. From 2012 onwards, new scans have been created in greyscale. ProQuest has no current plans to rescan existing images in greyscale.

There are some mistakes in the microfilming and digitization process. Pages are occasionally missing, or captured twice. ProQuest encourages EEBO-users to report mistakes by emailing their [Webmaster](#).

## Illustrations

When the microfilm is digitized, a manual process, a record is made of any page-image that contained an illustration, and a brief category is recorded for it. It is therefore possible to search for particular types of illustration, and to view only the pages with illustrations in a selected text.

## Tips

The categorization of images was done by non-experts and is not exhaustive or consistent. It is a helpful starting point, but it should not be assumed that all charts, for example, have been categorized as such.

## Thumbnails

These allow users to scan the contents of a whole book in thumbnail-sized images. This can be particularly useful when looking for illustrations, chapter headings, or typographical features.

## Full text

For some (currently c. 43,000) EEBO texts, a searchable version has been created by the Text Creation Partnership. These full texts have been transcribed using the digitized images, so they only show text that is legible in the images. They have also been created by non-expert staff, so they should not be used as authoritative editions in themselves.

- Links back to page images
- Representation of certain features (decorative initials) but not page layout
- Abbreviations recorded and expanded
- Digital table of contents generated for ease of navigation

## Accessing the content

There are two methods to access the contents of the database: search or browse. Browsing is by author, which can be particularly useful if a name is spelt in various forms.

Searching is the most popular way in to EEBO.

### Basic searching

Note “limit to” option (all records, all image sets, and all full texts).

### Boolean searching

Boolean operators (AND, OR, NOT) can be used. You can use parentheses to specify the order of searching, with bracketed terms being searched first. E.g. “cat AND dog”; “cat NOT dog”. N.B. to search for the phrase “cat and dog”, use double inverted commas.

### Proximity searching

Searches can be refined by using proximity operators (NEAR, FBY: followed by). You can restrict the limits by number of words. E.g. “cat NEAR.5 dog” searches for “dog” appearing within five words of “cat”.

### Keywords

In this context, a keyword is any word contained in the record or the full text. Clicking on “select from a list” shows you all the words in the database. Using this list can be helpful to see possible variant spellings.

### Author keyword

Author keyword uses name authority files, as you will see if you click “select from a list”. Selecting from the list can help identify the spelling and form of the author’s name, as only a slight variation can result in no hits being found.

### Title keyword

Title keyword will look in the title field in the bibliographic record and will return any title that contains that word.

## Subject keyword

Subject keyword uses the Library of Congress subject headings in the bibliographic record. It is important to use the correct word to produce successful search results, so, again, referring to the list may be helpful.

## Bibliographic number

If you know the STC number or the microfilm reel number, searching by bibliographic number can be the fastest way to find a book.

## Variant searching

There are two options to include variants in a search: spelling and form.

Some variant spellings are found by automatic checking of common patterns (vv and w); some have been found as part of the CIC CLI Virtual Modernisation Project (led by the Center for Library Initiatives (CLI) of the Committee on Institutional Cooperation (CIC)) that grew from a [Northwestern University](#) project [Wordhoard](#).

Variant forms searches for different forms of the same word (a search for “murder” will also return inflected forms: “murdered”, “murdering”, “murderer”, “murders” etc.). Variant-form searching can only be used in conjunction with the variant spellings option, not on its own.

## Tips

As the default, the variant-spelling search is used and the variant-form search is not. Truncation and proximity searching are also available (see [help page](#) for details).

## Advanced searching

Advanced searching introduces limits to full text searching, e.g. to specific parts of a work.

To search the text of a particular work, first search for the work, then for the text string, and then use “Search history” to combine the searches. There is another interface which makes this easier (details below).

## Other bibliographic fields: illustration search

Illustration search shows the full list of categories established during the digitization phase. Note that this list should be treated with caution (see above).

## Other bibliographic fields: UMI collection

University Microfilm International (UMI) was the commercial company that started the microfilm project in the 1930s. The different collections are UMI’s in-house way of referring to the groups of digitized files, and are a further way to limit the search.

## Using the material

Once a record is identified, a viewing option can be selected. The various options (bibliographic record, images, thumbnails, full text) are described above.

Records can be downloaded or printed from the record screen.

The current image from the document images screen can be downloaded, the view adjusted, and the current image printed.

The full text can be printed from the full text screen. It can also be downloaded by choosing file and save, as in your browser window. The table of contents can be used to navigate the text. At any point in the full text, you can click a link to view the document image. You can also use the ‘find’ option in your browser to search within the text. [Periodicals](#) and the [Thomason Tracts](#) can be used in the same way.

### What else can you do with the content?

Records can be added to a Marked List by checking the box to the left of a bibliographic record. By clicking through from the Marked List page (in the black banner at the top of the screen), bibliographic records can be printed or downloaded.

Downloaded records can be exported directly to [RefWorks](#) or [EndNote](#).

Image sets for whole texts can be downloaded as PDFs from the Marked List. This provides the flexibility to resize and print the images.

### Advanced EEBO: different interfaces for searching the full text

TCP texts are co-owned by partner institutions, and members of those universities and libraries are free to use the texts within the partnership. Five years after the end of each phase of the project, the embargo is lifted, and texts created in that phase are available for free distribution. On 1 January 2015, more than 25,000 texts were made available under an open license. A further c.40,000 texts will be made available in 2020.

An [auxiliary interface](#) has been created by the Text Creation Partnership. It allows advanced searching within a particular text or group of texts, including Boolean searching and Proximity searching. [JISC Historical Texts](#), another subscription resource, offers access to EEBO-TCP texts alongside Eighteenth Century Collections Online and the British Library's nineteenth-century, Microsoft-digitized, OCR'd texts.

Another interface, [Philologic](#), has been created by the [University of Chicago](#), and has more options for searching the full text. It is currently available only to EEBO subscribers in North America.

### Using EEBO-TCP texts for your own research

Many projects are based on EEBO-TCP texts. If you would like advice on using some or all of the TCP texts in your own research, please contact [eebotcp@bodleian.ox.ac.uk](mailto:eebotcp@bodleian.ox.ac.uk).

EEBO-TCP Phase I (& II) texts (Phase II only available to subscribers until 2020) available as HTML, ePub, and TEI P5 XML via the [Oxford Text Archive](#).

You can also find EEBO-TCP Phase I texts on [GitHub](#).

### Title page image

A.R., *True and wonderfull...* (1614). 4° R 21(5) Art. Seld., Title page: <http://solo.bodleian.ox.ac.uk/OXVU1:oxfaleph014672096>. Image: Bodleian Libraries.