Beyond the Facsimile: Early English Books Online (EEBO) and the Text Creation Partnership

Welzenbach, Rebecca

http://hdl.handle.net/2027.42/90981
Beyond the Facsimile: Early English Books Online (EEBO) and the Text Creation Partnership

Enriching Scholarship

May 7, 2012
Goals

• Understand the relationship between Early English Books Online and Text Creation Partnership
• Develop ability to “read” an interface: to think about, ask questions of, and make sense of what is being revealed or hidden by websites for digitized resources
• Understand what the Text Creation Partnership produces and some of the ways it can be used—as well as what it cannot do
• Share ideas for using these resources in research and teaching
Agenda

• Introductions (5 min)
• Background: What is the Text Creation Partnership? What is its relationship to EEBO? (10 min)
• Walk through/demo different ways to search the texts (25 min)
  • ProQuest’s EEBO
  • UM’s EEBO-TCP
  • Differences, similarities, preferences?
• Behind the scenes (25 min):
  • What does the TCP really produce?
  • What are the challenges and goals of this work? What purpose does it serve?
  • document analysis exercise
• Ideas/examples for teaching and research (20 min)
• Critical thinking: gaps and limitations of these resources (10 min)
• Questions/discussion (10 min)
• Evaluation (5 min)
Introductions

Who am I?
Rebecca Welzenbach, Text Creation Partnership Project Outreach Librarian

Who are you?

• Name
• Affiliation
• Do you already use EEBO or TCP? In what capacity?
• Any specific questions, concerns, or things you hope will be covered today?
Background: EEBO

- Contains more than 125,000 titles listed in Pollard & Redgrave’s Short-Title Catalogue (1475-1640) and Wing’s Short-Title Catalogue (1641-1700) and their revised editions, as well as the Thomason Tracts (1640-1661) and the Early English Books Tract Supplement

- Commercial product developed by ProQuest

- Purchased or licensed by libraries

- Grew out of Early English Books microfilm series

- Consists of scanned microfilm and (now) direct scans from books

- EEBO alone lets you search the MARC catalog records for each book and view the pages of the book. It does not provide full-text search capabilities

- Page images + searchable metadata

For more information: http://eebo.chadwyck.com.proxy.lib.umich.edu/about/about.htm
Background: EEBO-TCP

• More than 40,000 texts done so far. Goal is to complete ~70,000, or one edition of each unique work in the EEBO catalog

• Each page is keyed by hand (OCR doesn’t work; double/triple-keying) and mark-up added

• Product is a large set of searchable encoded text files that may be matched up with page images in EEBO for search and display, or put to other uses entirely

• Not-for-profit, library-based consortial project

• Libraries pay to support this work; gain immediate access to the output

• Text files belong to partner library institutions and will ultimately be made freely available to everyone

• Model has been replicated with ECCO- and Evans-TCP

For more information: http://textcreationpartnership.org
Demo: ProQuest’s EEBO Interface

http://eebo.chadwyck.com.proxy.lib.umich.edu/

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Demo: ProQuest’s EEBO Interface

http://eebo.chadwyck.com.proxy.lib.umich.edu/search

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Demo: ProQuest’s EEBO Interface

How does the Variant Spellings feature work?

The Variant spellings box appears on the Basic Search, Advanced Search and Periodicals Search screens. It is checked by default.

If you type a search term in the Keyword(s) box and the Variant spellings box is checked when you submit your search, you will automatically retrieve all instances of your search term and its early modern variant forms in EEBO. For example, if the box for Variant spellings is checked and you type the word murder in the Keyword(s) field, when you submit your search you will retrieve all occurrences of the word murder and its early modern variants murther, mordre, mordr and mordr.

This will also work if you check the box and type search terms in other fields like Title keyword(s) and Imprint.

If you type a phrase in the Keyword(s) search field (for example Keyword(s): “so foul and fair a day”) and the Variant spellings box is checked when you submit your search, your results will include instances of the phrase where the spelling varies (such as so foule and faire a day). Similarly, if you type a series of terms connected by Boolean or proximity operators in this field (for example jealous and green-eyed), your search will include all available spelling and typographic variants of each term (such as jealous, green-eyed and green eyed).

When typing a search expression that includes Truncation and wildcard operators (e.g. Keyword(s): je?lo?s*), you should uncheck the Variant spellings and Variant forms boxes.
Demo: ProQuest’s EEBO Interface

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Demo: ProQuest’s EEBO Interface

Catalog record only

Catalog record, images, illustrations, full text from TCP

Catalog record, images, illustrations

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Demo: ProQuest’s EEBO Interface

TO THE RIGHT HONORABLE MY SINGULAR GOOD LORD, THE LORD CHARLES HOUWARD, Earle of Nottingham, Baron of Effingham, Knight of the noble Order of the Garter, Lord High Admiral of England, Ireland, and Wales, &c. one of her Maiesties most honourable priuse Counsell.

Right Honourable and my very good Lord, after I had long since published in Print many Navigationes and Discoveries of Strangeres in divers languages, as well here at London, as in the citie of Paris, during my few yeares abode in France, with the worthye Knight Sir Edward Stafford your brother in lawe, her Maiesties most prudent and carefull Ambassador with the French King, and had waited on severale other noble persons, who had bringt to light many very rare and worthy memorials, which long have lien miserably scattered in musick corners, & rottenly hidden in mistie darkness, and were very like for the greatest part to have beene buried in permanent oblivion. The first Volume of this workes I have found for the present brought to light, receiving the other two which the next Spring, when by Gods grace they shall come to the Prasse. In the meantime season beholding my hope of some manfull and bountifull Patron, I came to mind your honorable lordship, who both in regard of my particular obligaition, and also in respect of the subject and matter, might lustily embrace the paternous care. For first I conceived how much I was bound, and had deeply invidted for my younger brother Edmund Hackley, to whom for the space of foure whole yeares your Lordship commited the government and instruction of that honorable yong noble man, your sonne & here aparent, the Lord William.

View document image (2)

Houward, of whose high spirit and wonderful tovallershe full many a time hath he boasted unto me. Secondly the bountie which liste I owe to your most deare sister the lady Steelford, my singular good lady & honorable mistresse, admonished me to be mindful of the renowned familye of the Houwardes. Thirdly, when I found in the first Patent granted by Queene Marie to the Moscoue companie, that my leaues ther being then lord high Admiral of England, was one of the first fueaters and furnitures, with his purport and counterpart, of the strange and wonderfull Discoverie of Russia, the chief centres of this present Voiurne, then I remembre the same saying of sweet Testiscnes. That stories ought not only to be inuironers of their fathers substance, but also of their commendable vertues and honours. But what speckle of your ancestors hones (which to say the truthe) are very great, and such as our Chronicles have notably bleezoned) when as your owne Henricall actions from time to time have shewen themselues so admirable, as no antiquite hath afforded greater, and the future times will not in haste (I thinke) performe the like. To come to some particulars, when the Emperors sister, the spouse of Sporne, with a Fleece of 130. solies, stoutly and proudly pass the narrow Seas, your Lordship accompanied with ten ships anely of her Maiesties nauie Rolali, enuironed their Fleece in most strange and warlike sort, enforced them to sallie, and to vaile their bonnes for the Queene of England, and made them perfectly to understand that else speachers of the prince of Poland.

Non illi imperium beliici saeculum(que) tridentem, 

et tibi sito olim datum.


Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Demo: University of Michigan’s EEBO-TCP Interface

Basic Search
Boolean Search
Proximity Search
Bibliographic Search
Word Index (Phase I)
Word Index (Phase II)
Browse (Phase I)
Browse (Phase II)

About EEBO-TCP
Demo EEBO-TCP
comments to: UMDL Help
Demo: University of Michigan’s EEBO-TCP Interface

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
The FIRST VOLUME OF THE principal Navigations, Voyages, Traffiques, and Discoveries of the English nation, made to the North and Northeast quarters of the World, with the directions, letters, privilegges, discourses, and observations incident to the same.

Certeine testimonies concerning K. Arthur and his conquests of the North regions, taken out of the historie of the Kings of Britaine, written by Guifridus Monemutensis, and newly printed at Heidelberg, Anno 1587.

Lib. 9. cap. 10.

ANno Christi, 517. Arthurus, secundo regni sui anno, subiugatis totius Hyberniae partibus, cladem suam directo in Islandiam, atque debellato populo subiugauit. Exin diutius per caeteras insulas rumore, quod e- nulo Provinzia resistere poterat, Doldaius rex Godlandiae, & Gunfanus rex Orkudam, vibro venerunt, promisoque vertigii subjectiorem fecerunt. Emensa deinde hyeme, reuerentiam in Britanniam, statimque regni in firmam pacem renounam, moenam duodecim annis ibidem fecit.

The same in English.

IN the yere of Christ, 517, king Arthur in the second yeere of his regne, hauing subdued all parts of Ireland, sailed with his fleet into Island, and brought it and the people thereof under his subjection. The rumour afterwards being spread thoroughout all the other Islands, that no countrey was able to withstand him, Doldaius the king of Godland, and Gunfanus the king of Orkney, came voluntarily vnto him, and yeelded him that obedience, promising to pay him tribute. The Winter being spent, he returned into Britaine, and establishing his kingdome in perfect peace, he continued there for the space of twelues yeres.

Lib. 9 cap. 12.

Missis deinde in diversa regna Legatis, inuitatur tam ex Gallijis, quam ex collateneribus Insulis Oceani, qui ad curiam venire deberent, &c. Et tunc post: Ex collateneribus autem Insulis, Guillaumei rex Hyberniae, Maluasia rex Islandiae, Doldaius rex Godlandiae, Gunnasius rex Orkudam, Lot rex Noruegiae, Aschilus rex Danorum.

The same in English.
Which meets your needs? Try it...

In both interfaces...
  • Try a keyword/full text search for “cookery”
  • Try searching within verse for “love”

To make things more interesting, add these additional variables:
  • Try it with and without spelling/form variants allowed in EEBO
  • Try changing the limits in EEBO (all items, full text only, etc.)

What do you find?
  • Questions?
  • Frustrations?
  • Inconsistencies?
  • Do you find you have a clear preference for one over the other?
Different search results

1. A. W. (G)

A book of cookery: very necessary for all such as delight therein. Gathered by A. W., at London: printed by Edward Alle, 1587.

Date: 1587
Bib name / number: STC (2nd ed.) / 24986
Copy from: New York Public Library (New York, N.Y.)

2. A. W. (G)


Date: 1591
Bib name / number: STC (2nd ed.) / 24987
Physical description: [1], 20 leaves
Copy from: Bodleian Library

3. Addis, Lancotol, 1632-1703. (C)

West Barbary, or, A short narrative of the revolutions of the kingdom of Fez and Morocco with an account of the present customs, sacred, civil, and domestic / by Lancotol Addis...; [London, printed by John Whiting...]

Date: 1671
Bib name / number: Wing / A8402
Copy from: Henry E. Huntington Library and Art Gallery

4. Aelian, 3rd cen (G)

Which one do I use?

ProQuest’s EEBO Interface

• search more than 120,000 books, whether or not they have full text
• Use fuzzy spelling or variant searching
• Easily download/print sections of a work as a PDF

University of Michigan’s TCP Interface

• Only search works with keyed full text
• Get most up-to-date batch of texts
• Some more precise searching options (e.g. proximity, sorting by frequency)

Both can: move between text and image view, display table of contents, restrict search to certain areas or kinds of text (e.g. letters, drama, verse, etc.)

Big picture: many of the same capabilities, some differences. Use what you’re comfortable with and what meets your needs. If something seems missing from one, try the other. Be aware that both—and any interface—will cause things to look and behave slightly differently, even with the same data underneath
What’s going on behind the scenes?

Two key components:
• Accurately transcribed text
• Structural markup: XML (more precisely, SGML)

That is: tags that explicitly describe the structure of a document, allowing a computer to “read” and understand the relationship of parts of the text to one another, just as we have learned to do when looking at a page.
SGML Markup

<BODY>
<DIV TYPE="book">
<PE REF="1">
<HEAD>THE EPISTLE OF PAUL TO PHILEMON</HEAD>
<ARGUMENT>
<p>1 Paul writing a simple and small matter, yet according to his manner mouthed aloft unto God. 2 Sending again to Philemon his vagabond and the indifferent servant, he wrote thee with his own hand, and very gravely precepts of Christian equity.</p>
<ARGUMENT>
<p>N. P. 1 Paul a prisoner of Jesus Christ, and
 Hib our brother Timotheus, unto Philemon our dear friend, and fellow helper.</p>
<p>N. P. 2 And to</p>
<p>N. P. 3 Grace be unto you, and peace from God our Father, and</p>
<p>N. P. 4 From the Lord Jesus Christ.</p>
<p>N. P. 5 Peace of grace and faith, which thou hast toward the Lord Jesus, and toward all Saints.</p>
<p>N. P. 6 That the</p>
<p>N. P. 7 By fellowship of faith, be meanest these duties of charitie which are bestowed upon the Saints, and those forth of an effectual faith. Fellowship of thy faith may be made effectual, and that whatsoever good thing is in you through Christ Jesus, may be</p>
<p>N. P. 8 That by this means all men may perceive how rich you are in Christ, to wit, in faith, charitie, and all bountiffuless.</p>
<p>N. P. 9 For we have great joy and consolation in thy love, because thee, brother, the Saints</p>
<p>N. P. 10 Because though it be small, and cheerfulness refresh the Saints, that they contented with the word of God, and</p>
<p>N. P. 11 They refresh the heart of thine enemies, and comfort them, and</p>
<p>N. P. 12 Wherefore, though I be very bold in Christ to command thee, that which is convenient.</p>
<p>N. P. 13 An example of a Christian exercise, &amp; commendation for an other man.</p>
<p>N. P. 14 For lothes sake I would rather beseech thee, though I be as I am, even Paul aged, and even now a prisoner for Jesus Christ.</p>
<p>N. P. 15 But I beseech thee for my sake, even</p>
<p>N. P. 16 Onasimus, whom I have begot in mine bonds.</p>
<p>N. P. 17 Which in times past was to thee unprofitable, but now profitable both to thee and</p>
<p>N. P. 18 Whom I have sent against, thou therefore receive him, that is mine own</p>
<p>N. P. 19 As mine own son, and as if I had begot him of mine own body.</p>
<p>N. P. 20 Whom I would have retained with me, that in thy stead he might have ministered unto me in the bonds</p>
</ARGUMENT>
</DIV>

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Document analysis exercise

• Pick a printout or several
• Work in small groups
• Identify elements of the document’s structure: what can you tell by looking at the page that a computer wouldn’t know? What information would you want to have pulled out?
• Do you agree about what is important to tag, and how to categorize it?
What is the point of all of this?

• An interface/program can only work with what information you give it in the first place

• Our goal is to produce a lot of high-quality data that can more or less be generally agreed upon. This gives a search engine something useful to grab onto, but also leaves room for further enhancements

• We turn wheat into flour, and hope you’ll make bread
Examples of new projects built on TCP data

• Specialized scholarly editions: [The Holinshed Project](#)
• Alternative search/display interfaces: [PhiloLogic @NU](#)
• Additional analysis: [Monk Workbench](#)
• Thematic groupings/websites: [Witches in Early Modern England](#)
• For readable, scholarly comments on text analysis: [Ted Underwood’s blog](#)
EEBO-TCP in the classroom

• Sample assignments: http://www.lib.umich.edu/tcp/eebo/scholarship.html

• 2012 Spring Term course on history of taste: mining EEBO- and ECCO-TCP for recipes

• Document analysis?

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Critical Thinking: Reading the Database

What problems, gaps, shortcomings do you see? (in the interfaces we looked at or in the data underneath)

What would you/do you warn your students about when encouraging them to use EEBO or TCP?
Further questions/discussion

• Please get in touch if we can be of help with your research, or with an assignment in your classroom (happy to give targeted overviews/presentations)
• Thanks and good luck!
• Please fill out the evaluation: http://www.umich.edu/~teachtec/es.html