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

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
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, murther, murder and murther.

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 far 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, greene-eyed and green eyd).

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
Demo: ProQuest's EEBO Interface

TO THE RIGHT HONORABLE MY SINGULAR GOOD LORD, THE LORD CHARLES HOWARD, Earl 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 proue Counsell.

Right Honorable and my very good Lord, after I had long since published in print many Navigations and Discoveries of Strangers in divers languages, as well here at London, as in the city of Paris, during my late sojourn abroad in France, with the best wishes Knight Sir Edward Stafford your brother in law, her Maiesties most prudent and careful Ambassador with the French King, who had voiced on old matter and farther in the sweet study of the history of Discoveries, I began at last to conceive, that with diligent association, some things might be gathered which might commend our nation for their high courage and singular activity in the Search and Discoveries of the most unknown quarters of the world. Hence, seeing no man to stop forth to undertake the recording of so many memorable actions, and many men to follow his private affaires, the honest base of my country despaired all difficulties, and as it were with a sharp good pommelled me and thrust me forward into this most troublesome and painful action. And after great charges and infinite cares, after many wantsovers, tents, and tarps, and warning out of my woak body, at length I have collected these several Volumes of the English Navigations, Travels, and Discoveries, to strange, remote, and far distant countries. Which works of mine I have not included within the compass of things only done in these latter days, as though like no other works of memory have been performed in former ages. But recording only by the space of many hundred years, have brought to light many very rare and worthy monuments, which long have lien miserably scattered in muske corners, and rashly hidden in moist darkness, and were very like for the greatest part to have been buried in perpetual oblivion. The first Volume of this work I have thus for the present brought to light, reserving the other two until the next Spring, when by God's grace they shall come to the Press. In the meantime season beholding my safe of some new fluent and sonfulfull Perusals, I called to mind your honorabile Lordship, who both in regard of my particular obligation, and also in respect of the subject and matter, might justly challenge the patronage thereof. For first I considered how much I was bound, and how deeply indebted for my younger brother Edmund Hacklby, to whom for the space of four whole years your Lordship committed the governement and instruction of that honorable young noblem man, your came & here present, the lord William

Howard, of whose high spirit and wonderful towardliness full many a time hath he boasted unto me. Secondly, the boundless fortune which I owe to your most desirous the lady Stafford, my singular good lady & honorable mistress, admonished me to be mindful of the renowned family of the Howards.

Thirdly, when I found in the first Patent granted by Queen Marie to the Moscourt company, that my lord your father being then lord high Admiral of England, was one of the first Travelers and Purchasers, with his pursure and courtoise, of the strange and wonderful Discoveries of Russia, the chief centers of this present Volume, then I remembe regarding the sage saying of sweet Joannes. That storms ought not only to be inhbiters of their fathers substance, but also of their commendable vertues and honours. But what specially 1 of your ancestors heroes (which to say the truth), are very great, and such as our Chronicles have notably bewazent when as your own Heroicall actions from time to time have shewed themselves so admirable, as no antiquitie hath afforded greater, and the future times will not in haste (I think) performe the like. To come to some particulars, when the Emperors sister, the spouse of Spino, with a Fleece of 1 L30, sailed, stoutly and proue passed the narrow Seas, your Lordship accompanied with ten ships anly of her Maiesties Navio Roi, enronned their Fleet in most strenghe and warlike sorte, enforced them to sail, and to waile their bones for the Queen of England, and made them perfectly to understand that empery speach of the prince of Peas;

Hen II imperium regii saecuum (que) tridentem, et timi seriee datum.


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

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

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

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
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?

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
Different search results

Rebecca Welzenbach | tcp-info@umich.edu | textcreationpartnership.org | @TCPStream
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.
<BODY>
<DIV TYPE="book">

P% REF="1">

<HEAD>THE EPISTLE OF PAUL TO PHILEMON.</HEAD>

<ARGUMENT>

P% 1 Paul handling a base and small matter, yet according to his manner mouthed aloft unto God, 8 Sending again to Philemon his vagabound and theeish servant, he entreateth pardon for him, and very gravely preacheth of Christian equity. P% <ARGUMENT>

P% 1 P Paul a prisoner of Jesus Christ, and

HB> our & HB> brother Timotheus, vinto Philemon our dear friend, and fellow helper.</HB>

P% 2 & 2 And to

HB> our & HB> dear

HB> brethren & HB> Apphia, and to Archippus our fellow soilder, and to the Church that is in thine house.</HB>

P% 3 & 3 Grace

HB> be & HB> with you, and peace from God our Father, and

HB> from & HB> the Lord Jesus Christ.</HB>

P% 4 & 4

<NOTE PLACE="marg1" N%.""

P% 1 & 1 <HB>That,</HB> 12. 2. <HB>thee,</HB> 13. <NOTE>

P% 5 & 5 (When I heart of thy love and faith, which thou hast toward the Lord Jesus, and toward all Saints.)

P% 6 & 6 That the <NOTE PLACE="marg2" N%.""

P% 2 & 2 By fellowship of faith, he meaneth those duties of charity which are bestowed upon the Saints, and those of them of an effectual faith. <NOTE> Fellowship of thy faith may be made effectual, and that whatsoever good thing is in you through Christ Jesus, may be</NOTE>

<NOTE PLACE="marg2" N%.""

P% 4 & 4 That by this means all men may perceive how rich you are in Christ, to wit, in faith, charity, and all bountifulness. <NOTE> Known.</NOTE>

P% 7 & 7 For we have great joy and consolation in thy love, because by thee, brother, the Saints</NOTE>

<NOTE PLACE="marg2" N%.""

P% 4 & 4 Because they did not so doffly and cheerfully refresh the Saints, that they conceived of the word as of a precious joy; for by this word (Bowels) is meant not only the inward feeling of wights and miseries that men have one of another's state, but also that joy and comfort which entereth into the very bowels, as though the heart were refreshed and comforted. <NOTE> Bowels are comforted.</NOTE>

P% 8 & 8 Whencefore, though I be very bold in Christ to command thee, that which is convenient</NOTE>

P% 9 & 9

<NOTE PLACE="marg2" N%.""

P% 1 & 1 An example of a Christian exercise & & comktion for an other man. <NOTE>

HB> Yet HB> for thou saile I rather beseech thee, though I be as I am, even Paul aged, and even now a prisoner for Jesus Christ.</HB>

P% 10 & 10 I beseech thee for my sake</NOTE>

<NOTE PLACE="marg2" N%.""

HB> Colos.<HB> 4.9. <NOTE>

P% 11 & 11 Which in times past was to thee unprofitable, but now profitable both to thee and to me.</NOTE>

P% 12 & 12 Whom have sent against thee therefore receive him, that is mine own</NOTE>

<NOTE PLACE="marg2" N%.""

P% 13 & 13 Whom I would have retained with me, that in thy stead he might have ministered unto me in the bonds of the Gospel.</P>

</ARGUMENT>

<&auloy. To the Hebrews. Christs excellencel.

THE EPISTLE OF PAUL TO PHILEMON.

1. Paul, a servant of Jesus Christ, and a brother Timotheus, vunto Philemon our deere friend, and fellow helper,

2. 2 And to our deere brethren and brethren Apphia, and to Archippus our trustye and loving brother, and to the Church that is in thy house.

3. 3 Grace be with you, and peace from God our Father, and from the Lord Jesus Christ.

4. 4 These things I write unto thee, that thou readest and considerest them, and observe them, and follow after them, for they be thy desire.

5. 5 For we have great joy and consolation in thy love, because by thee, brother, the Saints

6. 6 Because they did not so doffly and cheerfully refresh the Saints, that they conceived of the word as of a precious joy; for by this word (Bowels) is meant not only the inward feeling of wights and miseries that men have one of another's state, but also that joy and comfort which entereth into the very bowels, as though the heart were refreshed and comforted. <NOTE> Bowels are comforted.</NOTE>

7. 7 Whencefore, though I be very bold in Christ to command thee, that which is convenient

8. 8 An example of a Christian exercise & & comktion for an other man. <NOTE>

9. 9 Yet for thou saile I rather beseech thee, though I be as I am, even Paul aged, and even now a prisoner for Jesus Christ.

10. 10 I beseech thee for my sake

11. Onasimus, whom I have begotten in my bonds,

12. Which in times past was to thee unprofitable, but now profitable both to thee and to me.

13. Whom have sent against thee therefore receive him, that is mine own

14. As mine own son, and as if I had begotten him of mine owne body.

15. Whom I would have retained with me, that in thy stead he might have ministered unto me in the bonds of the Gospel.


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?
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