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 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


Now available – EEBO Interactions: A social network for Early English Books Online. (sign in, register as a new Interactions user, or find out more)

The EEBO Introductions Series, providing concise and informative commentaries on some of the less frequently discussed texts in EEBO

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, murdre, murther 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 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 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

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 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 privie Counsell.

Right Honourable, and my very good Lord, after I had long since published in print many Navigation and Discoveries of Strangers in divers languages, as well here at London, as in the state of Paris, during my late voyages abroad in France, with the worthy Knight Sir Edward Stafford your brother in law, her Maiesties most prudent and careful Ambassador with the French King; and had resided on several years in the sweet study of the history of Discoveries, I began to find, and to remark, that with all the diligence, care, and pains that might be gotten and perhaps, that might be gotten and perhaps, might commend our nation for their high courage and singular activity in the Search and Discovery of the most unknown quarters of the world. Henceforward, seeing no man to stop forth to undertake the recording of so many memorable actions, and every man to follow his private affairs, the fruit of all my country's labours all difficulties, and as it were with a sharp glance provided me and thrust me forward into this most troublesome and painful action. And after great changes and infinite care, after many waterfalls, toils, and travaill, and warning out of my weak body, at length I have collected three yeares 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, so as that this or nothing worth of matters have been performed in former ages, but not so much by the care of any hundred years, have brought to light many very rare and worthy monuments, which long have lain miserably scattered in musty corners, and rottenly hidden in musty darknesse, and were very like for the greatest part to have been buried in perpetual oblivion. The first volume of these works I have thus for the present brought to light, reserving the other two until the next spring, when by God's grace when they shall come to the press. In this season, beholding my safety of some mean estate and subsistence, I called to mind your honourable Ladyship, who is both in regard of my particular obligation, and also in regard of the subject and matter, might possibly desire the publication thereof. For first, I remembered that much I was bound, and that deeply indebted for my younger brother Edmund Hakluyt, to whom for the space of four whole yeares your Lordship committed the government and instruction of that honorable young noble man, your sonne & here apparent, the Lord William

View document image (2)

Howard, of whose high spirit and wonderful totherliness full many a time hath he boasted unto me. Secondly, the bounteous grace which I owe to your most desirous wife the lady Stafford, my singular good lady & honourable 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 Moscou company, that my noble mother being then long high Admiral of England, was one of the first navigators and furthurers, with her pursuance and continuance, of the strange and wonderful Discoveries of Russia, the chief centre of this present Volume; then I remembrace the said saying of sweet Isobazon. That since our country ought not only to be inheritors of their fathers substance, but also of their commendable virtue and honours. But what speak I of your ancestors heroes (which to say the truth, are very great, and such as our Chronicles have notably praised) when as your owne Heroicall actions from time to time have shewed themselves so admirable, as no antiquite hath afforded greater, and the future times will not in haste [I think] perform the like. To come to some particulars, when the Emperors sister, the spouse of Spino, with a fleet of one hundred, stoutly and proudly passed the narrow Seas, your Lordship accompanied with ten ships only of her Maiesties Navis Ro balance, enrobed their Fleet in most strange and wonderful sort, enforced them to sail, and to visit their interests for the Queen of England, and made them perfectly to understand that alone speach of the prince of Poets,


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

Early English Books Online

Text Creation Partnership

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

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

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, priludges, discourses, and observations incident to the same.

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

Lib. 9. cap. 10.

ANno Christi, 517. Arthurus, secundo regni sui anno, subiugatis totius Hyberniae partibus, classem suam direxit in Islandiam, quamque debellato populo subiugavit. Exin diutiusque per caeteras insulas rumore, quod e nullo Provincia resistere poterat, Doldauius rex Gotlandiae, & Gunfacius rex Orklandiae, vibro venerunt, promissisque vertiguli subjectionem fecerunt. Emensa deinde hyeme, reuerus est in Britannian, statimque regni in firmam pacem renouant, moenum duodecim annis ibidem fecit.

The same in English.

IN the yere of Christ, 517. king Arthur in the second yeere of his reigne, 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 throrowout all the other Islands, that no countrey was able to withstand him, Doldauies the king of Godland, and Gunfaxies the king of Orkney, came voluntarily vnto him, and yeelded him their 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 Logatis, insitantur tam ex Gallijis quam ex collatetinibus Insulis Oceani, qui ad curiam venire deberen, &c. Et paulo post: Ex collatetinibus autem Insulis, Guillaumeus rex Hyberniae, Malauius rex Islandiae, Doldauies rex Gotlandiae, Gunnasius rex Orkhandum, Lot rex Noruegiae, Aschilius 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

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">
<PB REF="1">
<HEAD>THE EPISTLE OF PAUL TO PHILEMON</HEAD>
<ARGUMENT>
<P>1 Paul, a prisoner of Jesus Christ, and 
</P>
<HB>our</HB> brother Timotheus, write Philemon our dear friend, and fellow helper.</P>
<P>2 And to 
</P>
<HB>our</HB> desire 
</P>
<HB>inter</HB> and 
</P>
<HB> Apphia, and Archippus our fellow soldiers, and to the Church that is in thine house.</P>
<P>3 Grace be to 
</P>
<HB>you</HB> with you, and peace from God our Father, and 
</P>
<HB>from</HB> the Lord Jesus Christ.</P>

<NOTE PLACE="marg1">
</NOTE>give thanks to my God, 
</NOTE>

<NOTE PLACE="marg2">
1. 5. 
</NOTE>

<NOTE PLACE="marg3">
4. 6. 7. 
</NOTE>

<NOTE PLACE="marg4">
2. 4. 5. 
</NOTE>

<NOTE PLACE="marg5">
3. 
</NOTE>

<NOTE PLACE="marg6">
That by the fellowship of faith, he meaneth those duties of charity which are bestowed upon the Saints, and those faith of an effectual faith. 
</NOTE>

<NOTE PLACE="marg7">
That by these means all men may perceive how rich you are in Christ, to wit, in faith, charity, and all bountifulness. 
</NOTE>

<NOTE PLACE="marg8">
For were there great joy and consolation in thy love, because by these, brother, the Saints 
</NOTE>

<NOTE PLACE="marg9">
Because these are as the refreshing of the water, that is to say, when they received these words of charity, they conceived a greater charity, which they communicated to others, and did refresh the whole Church, and did comfort and strengthen them.</NOTE>

<NOTE PLACE="marg10">
8 Wherefore, though I be very bold in Christ to command thee, that which is conformable. 
</NOTE>

<NOTE PLACE="marg11">
Yet I beseech thee for love sake. 
</NOTE>

<NOTE PLACE="marg12">
Colos. 4. 9. 
</NOTE>

<NOTE PLACE="marg13">
Onasimus, whom I have begotten in my bonds, 
</NOTE>

<NOTE PLACE="marg14">
11 Which in times past was to thee unprofitable, but now profitable both to thee and to me. 
</NOTE>

<NOTE PLACE="marg15">
As mine own son, and as if I had begotten him of mine own body. 
</NOTE>

<NOTE PLACE="marg16">
13 Whom I would have retained with me, that in thy stead he might have ministered unto me in the bonds of the Gospel. 
</NOTE>
</DIV>
</BODY>
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, shortcomingds 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?

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