

<http://emop.tamu.edu>

eMOP so far

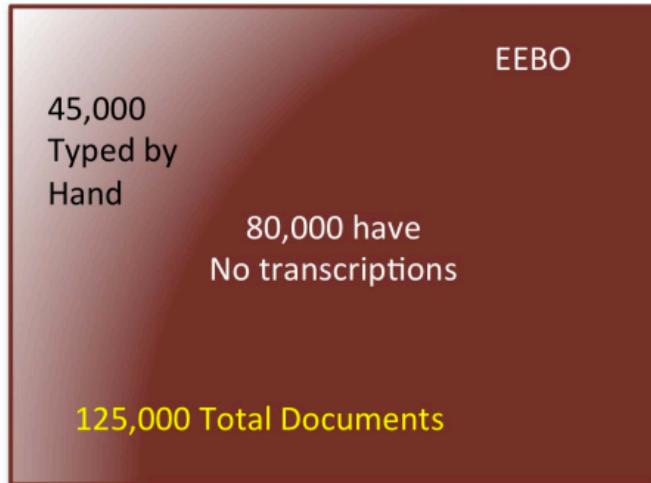
Chicago DHCS Colloquium
23 October 2014
Laura Mandell

The screenshot shows the homepage of the Early Modern OCR Project (eMOP). The header features the project's logo, "eMOP" in large gold letters with a blue swoosh, and the subtitle "EARLY MODERN OCR PROJECT". Below the header is a search bar. The main content area displays two columns of medieval lyrics. The left column contains lyrics from "The Merchant's Tale" and "The Wife of Bath's Prologue". The right column contains lyrics from "The Miller's Tale" and "The Wife of Bath's Prologue". Below these lyrics are three sidebar boxes: "Why eMOP Matters", "Recent News", and "@IDHMC_Nexus".

The Early Modern Data Set

Time

Owned by ProQuest



1500 -
Fonts Imported

1640 -
Spelling regularizes
Fonts still Imported

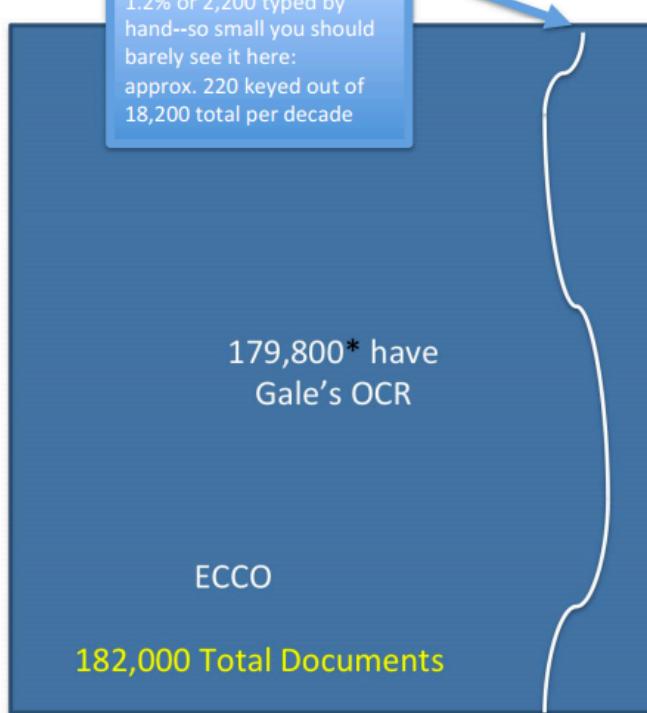
1700
Fonts still Imported

1720 -
Primarily Caslon and Caslon-like Fonts

1755 -
Spelling Standardized

1790 -
Long S
1800

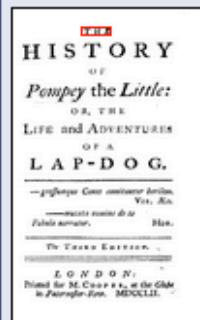
Owned by Gale Cengage-Learning



Tools

"The history of Pompey the littl..."

delete edits

[abc Insert Above](#)[abc Insert Below](#)OCR Source: **gale**Document Home [first](#) [prev](#) Page **2** [next](#) [last](#)

THE
HISTORY

[report this page](#)
[resize red box](#)


1

2 HISTORY

Document Home [first](#) [prev](#) Page **2** [next](#) [last](#)
Keyboard Shortcuts

Key	Action
Ctrl+Delete	Delete the line
Ctrl+Enter	Assert the line is correct
Ctrl+Y	Undo/Redo
Ctrl+Shift+I	Insert line above
Ctrl+Shift+H	Find
Ctrl+I	Insert line below
Down Arrow	Go down one line
End	Move cursor to end of line
Enter	Go down one line
Home	Move cursor to beginning of line
Page Down	Go down three lines
Page Up	Go up three lines
Up Arrow	Go up one line

In this document the original spelling should be retained.

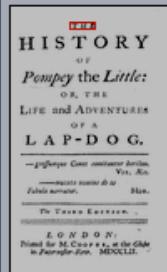
Instructions:

Use arrow buttons above-right or click on the thumbnail to jump around the page. You can add and delete lines; even if some lines are incompletely outlined in red, type every word on the line in which the red box appears.

- If a word or portion of a word is illegible, type "@" in its place; please do not make any guesses about what a word might be.
- Copy original spelling and punctuation, typing what you see on the page, except in the case of the long 's': use 's' and not 'f' when 's' is called for.
- Include end-of-line hyphens, preserving the syllables as they occur on each line.

[Feedback](#)[Help](#)[FAQs](#)[Optional Markup Guide](#)

"The history of Pompey the Little; or, the life and adventures of a lap-dog."

[abc Insert Above](#)[abc Insert Below](#)OCR Source: **gale**[report this page](#)[resize red box](#)

TH
H I S T



1 THE

2 HISTORY

Keyboard Shortcuts

Key	Action
Ctrl+Delete	Delete the line
Ctrl+Enter	Assert the line is correct
Ctrl+Y	Undo/Redo
Ctrl+Shift+I	Insert line above
Ctrl+Shift+H	Find
Ctrl+I	Insert line below
Down Arrow	Go down one line
End	Move cursor to end of line
Enter	Go down one line
Home	Move cursor to beginning of line
Page Down	Go down three lines
Page Up	Go up three lines
Up Arrow	Go up one line

In this document the original text is:

Instructions:

Use arrow buttons above-right of the text area to move the cursor. You can add and delete lines; even entire pages. Click on the line in which the red box appears.

- If a word or portion of a word is underlined, click on it to guess what the word is. You can then type over your guess.
- Copy original spelling and punctuation. In this case of the long 's': use 's' and 'f' instead of 'ſ' and 'ſ'.
- Include end-of-line hyphens and punctuation.

[Feedback](#)[Help](#)[markup](#)

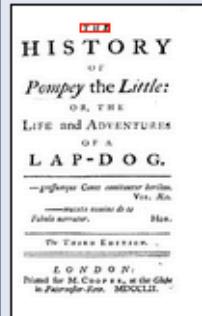
For more details on the Markups that TypeWright recognizes, visit the [TypeWright Markup Page](#)

DEBUGGING INFO: Word statistics

	Odd punctuation	1 occurrence
This page:	11: -grejfumfue (1) ↴	38: A ↴
First 100 pages:	358: ! (?) ↴	3492: "Marivaux" ↴

[Close](#)

"The history of Pompey the littl..."

[delete edits](#)[abc Insert Above](#)[abc Insert Below](#)OCR Source: **gale**Document Home [first](#) [prev](#) Page **2** [next](#) [last](#)[report this page](#)[resize red box](#)**THE****HISTORY**

1 <hi rend="smallcaps">THE</hi>



2 HISTORY

Document Home [first](#) [prev](#) Page **2** [next](#) [last](#)**Keyboard Shortcuts**

Key	Action
Ctrl+Delete	Delete the line
Ctrl+Enter	Assert the line is correct
Ctrl+Y	Undo/Redo
Ctrl+Shift+I	Insert line above
Ctrl+Shift+H	Find
Ctrl+I	Insert line below
Down Arrow	Go down one line
End	Move cursor to end of line
Enter	Go down one line
Home	Move cursor to beginning of line
Page Down	Go down three lines
Page Up	Go up three lines
Up Arrow	Go up one line

In this document the original spelling should be retained.

Instructions:

Use arrow buttons above-right or click on the thumbnail to jump around the page. You can add and delete lines; even if some lines are incompletely outlined in red, type every word on the line in which the red box appears.

- If a word or portion of a word is illegible, type "@" in its place; please do not make any guesses about what a word might be.
- Copy original spelling and punctuation, typing what you see on the page, except in the case of the long 's': use 's' and not 'f' when 's' is called for.
- Include end-of-line hyphens, preserving the syllables as they occur on each line.

[Feedback](#)[Help](#)[FAQs](#)[Optional Markup Guide](#)

"The history of Pompey the littl..."

[abc Insert Above](#)[abc Insert Below](#)OCR Source: **gale**Document Home [first](#) [prev](#) Page 3 [next](#) [last](#)

[iii]

[report this page](#)
[resize red box](#)

-- top of page --

1 [iii



2 TO

Document Home [first](#) [prev](#) Page 3 [next](#) [last](#)[report this page](#)
[resize red box](#)

[iii]

[iii]

"The history of Pompey the littl..."

delete edits

abc Insert Above

abc Insert Below

OCR Source: gale

Document Home first prev Page 3 last



[iii]

[report this page](#)
[resize red box](#)

-- top of page --



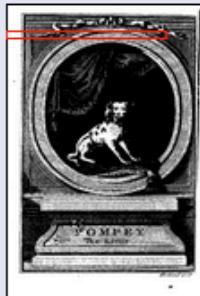
1 [iii]



2 TO

Document Home first prev Page 3 last

"The history of Pompey the littl..."

[delete edits](#)[abc Insert Above](#)[abc Insert Below](#)OCR Source: **gale**[Document Home](#)Page **1**[next](#) [last](#)[report this page](#)[resize red box](#)

- 0.125 qweqweweqwe
- 0.25 another
- 0.5 qweqwe

[Document Home](#)Page **1**[next](#) [last](#)

Keyboard Shortcuts

Key	Action
Ctrl+Delete	Delete the line
Ctrl+Enter	Assert the line is correct
Ctrl+Y	Undo/Redo
Ctrl+Shift+I	Insert line above
Ctrl+Shift+H	Find
Ctrl+I	Insert line below
Down Arrow	Go down one line
End	Move cursor to end of line
Enter	Go down one line
Home	Move cursor to beginning of line
Page Down	Go down three lines
Page Up	Go up three lines
Up Arrow	Go up one line

Live Edit Status

The following people are currently editing this page:

mandelic (0 0:0:27)

In this document the original spelling should be retained.

Instructions:

Use arrow buttons above-right or click on the thumbnail to jump around the page. You can add and delete lines; even if some lines are incompletely outlined in red, type every word on the line in which the red box appears.

- If a word or portion of a word is illegible, type "@" in its place; please do not make any guesses about what a word might be.
- Copy original spelling and punctuation, typing what you see on the page, except in the



OCR Source: gale

Document Home

Page 1

next last

abc Insert

[report this page](#)

[resize red box](#)



[report this page](#)

[resize red box](#)



3 *.k; ...

4 "*****T*****liiiijiiihllllliilli



5 UserB updated

Document Home

Page 1

next last

ZOOM

+
-
✖
↔

Erscheint jeden Mittwoch und Samstag.

Nr. 3.	Samstag, den 9. Januar 1904.	XVI Jahrgang.
Schriften oder Besuchs erblitten an das Redaktionsbureau Bannstr. 51.	Öffnungsstunde 12—2 Uhr. Telefon 1517.	Abonnement laut bez. reich illustrierten Romanenhefte vierstellige 1. Kl. Kunstgalerie Nr. 5 Preis. Interesse mit 10-pfennige Beiträge 25 ♂, im Stellamenten 80 ♂.
Inhalt: Wochen-Rundschau. — Karneval. — Der Verkehrsbeamte und Offizier a. D. als — Penitentiant — Eine schöne Frau von einer „feinen Dame“ u. s. w.		
Wochen-Rundschau.		
<p>Die vom Verkehrsministerium haben sich gesetzlich gelegentlich der Inhaberung dieses Amtes in der verkehrlichen Weise angehängt. Das Ehe, was von diesen Ministerium zu vermelden ist, doch war die Freundschaft der Beamten mit einem geschwollen Rind mit einem „Eben“ zu versehen anzuschaut. Hingegen bedeutungsvoll fügt die Sache schon gleich an, das muss man sagen.</p> <p>Schon wieder ein Eher. Wenn schon die Millionäse zu Verbrechen werden in unserem Reich, wie sollte es mit den hunderttausenden armen Deuteln, die vor der Hand in den Mund leben, aber von Tag zu Tag sich um die elende Existenz abquälen und abvögeln müssen, noch werden? Ja, die Unzufriedenheit, der noch immer kein Regel vorgeschoben wird, die bringt überall Angst und Elend. Heute haben wir es mit einem Millionär, einem von denen, welchen der Teufel zu Gold geworden ist, zu tun. Ebenso, auch einer von jenen in der Gestaltung unserer Stadt, welche die gütige Vorstellung zu reichen Leuten gemacht hat, hat sich eine kleine Meineidserziehung gefestigt und ist wohl kaum etwas verschafft worden. — Ein gewisser Ledner (Sendlinger Brückenstraße) müsste sich bei Gericht offenbaren, dass er ebenfalls zu den Proletarien (Leuten, die nichts haben) gehört. Er offenbarte sich offiziell und dazu hat ihm der Millionär Eisenbahn etwas versprochen, indem er seinen Spezi, den Ledner, riet, er soll ein dem letzteren gehöriges Grundstück in Werte von 60.000 ♂ ihm, dem Millionär, zurückholen lassen und nicht angeben. Das kam zur Kenntnis des Gerichtes und jetzt ist der habschere und edle Millionär wegen genannten Verbrechens etwas interessiert worden und kann der Dinge, die da kommen werden. Wir immer hübsch ehrlich stehen und wenn es auch schwer fällt.</p> <p>Die Bedeutung vor dem neuen Volks-Theater an der St. Paulikirche ist wirklich einzig. Die</p>		

CONTENT

□
^
▼

CHANGES

█
↻
?

CREATE REGION

Page Header

Page Identifier

Title

Paragraph

Paratext

Printer's Mark

Other Text

Image

Decoration

Separator

Other

Insert Footer Content Here © 2014 Null posuere lectus at dolor ullamcorper suscipit. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Cras quis felis varius, eleifend eros a, interdum augue.

Live Edit Status

The following people are currently editing this page:

mandellc (0 0:4:32)
mandellc (0 0:3:13)
mandellc (0 0:1:35)

The following people are currently editing other pages in this document:

mandellc (0 0:4:47) (page: 2)

EXPLORE[Browse Books](#)
[Search for Titles](#)
[Add a Book](#)**ABOUT**[Introduction](#)
[Contact Info](#)
[Logout](#)**ROLES**[Viewer](#)
[Authenticated](#)
[Stuff](#)
[Superuser](#)
[Transcriber](#)**Books**

- [+40094 Alkinou eis ta tou Platonos eisagoge Alcinoi in Platonicam philosophiam introductio.](#) by Albinus. (1667) published by Oxoniae : Typis Lichfieldianis ..., 1667.
- [+40104 The Ages of sin, or, Sinnes birth & growtheth with the stepps and degrees of sin, from thoug](#) by Anon. (1655) published by [London : Thomas Jenner, ca. 1655]
- [+40138 Sir Henry Vanes politicks, or, His cases of conscience lately found in his cabinet at Arabie,](#) by Anon. (1661) published by London : [s.n.], Printed in the year, 1661.
- [+Alkinou eis ta tou Platonos eisagoge Alcinoi in Platonicam philosophiam introductio.](#) by Albinus. (1667) published by Oxoniae : Typis Lichfieldianis ..., 1667.
- [+Of religious melancholy a sermon preach'd before the Queen at White-hall](#) by Moore, John, 1646-1714 (None)
- [+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)
- [+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)
- [+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)
- [+Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)
- [+Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)
- [+Of the immortality of the soul a sermon preached before the King and Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

EXPLORE

[Browse Books](#)
[Search for Titles](#)
[Add a Book](#)

ABOUT

[Introduction](#)
[Contact Info](#)
[Logout](#)

ROLES

[Viewer](#)
[Authenticated](#)
[Stuff](#)
[Superuser](#)
[Transcriber](#)

**OF RELIGIOUS MELANCHOLY
A SERMON PREACH'D BEFORE
THE QUEEN AT WHITE-HALL**

[Book Overview](#)
[Start Reading View](#)
[Detailed View](#)
[View in Repository](#)
[Compiled PDFs](#)
[Add to compare](#)

OF RELIGIOUS ... ADMIN

[Edit Internal Pages](#)
[Edit External Views](#)
[Edit References](#)
[Bitstream Metadata File](#)

Of religious melancholy a sermon preach'd before the Queen at White-hall

Author: Moore, John, 1646-1714

Published: None

Collection: John Moore 3

Repository: Texas A&M University Repository

Issued: March 12, 2014, 11:17 a.m.



Books

COMPARE BOOKS

[+40094 Alkinou eis ta tou Platonos eisagoge Alcinoi in Platonicam philosophiam introductio.](#) by Albinus.

(1667) published by Oxoniae : Typis Lichfieldianis ..., 1667.

[+40104 The Ages of sin, or, Sinnes birth & growtheth with the stepps and degrees of sin, from thoug](#) by Anon.

(1655) published by [London : Thomas Jenner, ca. 1655]

[+40138 Sir Henry Vanes politicks, or, His cases of conscience lately found in his cabinet at Arabie.](#) by

Anon. (1661) published by London : [s.n.], Printed in the year, 1661.

[+Alkinou eis ta tou Platonos eisagoge Alcinoi in Platonicam philosophiam introductio.](#) by Albinus. (1667)

published by Oxoniae : Typis Lichfieldianis ..., 1667.

[+Of religious melancholy a sermon preach'd before the Queen at White-hall](#) by Moore, John, 1646-1714

(None)

[+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714

(None)

[+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714

Add to comparison list

[+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714

(None)

[+Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714

(None)

[+Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714

(None)

[+Of the immortality of the soul a sermon preached before the King and Queen at White-Hall](#) by Moore, John,

1646-1714 (None)

- Of religious melancholy a sermon preach'd before the Queen at White-hall

- Of religious melancholy a sermon preach'd before the Queen at White-Hall

COMPARE

RESET

Books

[+40094 Alkinoou eis ta tou Platonos eisagoge Alcinoi in Platonicam philosophiam introductio.](#) by Albinus. (1667) published by Oxoniae : Typis Lichfieldianis ..., 1667.

[+40104 The Ages of sin, or, Sinnes birth & growtheth with the stepps and degrees of sin, from thoug](#) by Anon. (1655) published by [London : Thomas Jenner, ca. 1655]

[+40138 Sir Henry Vanes politicks, or, His cases of conscience lately found in his cabinet at Arabie.](#) by Anon. (1661) published by London : [s.n.], Printed in the year, 1661.

[+Alkinoou eis ta tou Platonos eisagoge Alcinoi in Platonicam philosophiam introductio.](#) by Albinus. (1667) published by Oxoniae : Typis Lichfieldianis ..., 1667.

[+Of religious melancholy a sermon preach'd before the Queen at White-hall](#) by Moore, John, 1646-1714 (None)

[+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

[+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

[+Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

[+Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

[+Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

[+Of the immortality of the soul a sermon preached before the King and Queen at White-Hall](#) by Moore, John, 1646-1714 (None)

COMPARE BOOKS

- Of religious melancholy a sermon preach'd before the Queen at White-hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy A sermon preach'd before the Queen at White-Hall
- Of religious melancholy A sermon preach'd before the Queen at White-Hall
- Of religious melancholy A sermon preach'd before the Queen at White-Hall
- Of the immortality of the soul a sermon preached before the King and Queen at White-Hall

COMPARE

RESET

Digital Books at Texas A&M

[Of religious melancholy a sermon preach'd before the Queen at White-hall](#) by Moore, John, 1646-1714
(None)

[Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714
(None)

[Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714
(None)

[Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714
(None)

[Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714
(None)

[Of religious melancholy A sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714
(None)

[Of the immortality of the soul a sermon preached before the King and Queen at White-Hall](#) by Moore,
John, 1646-1714 (None)

[Of religious melancholy a sermon preach'd before the Queen at White-hall](#) by Moore, John, 1646-1714
(None)

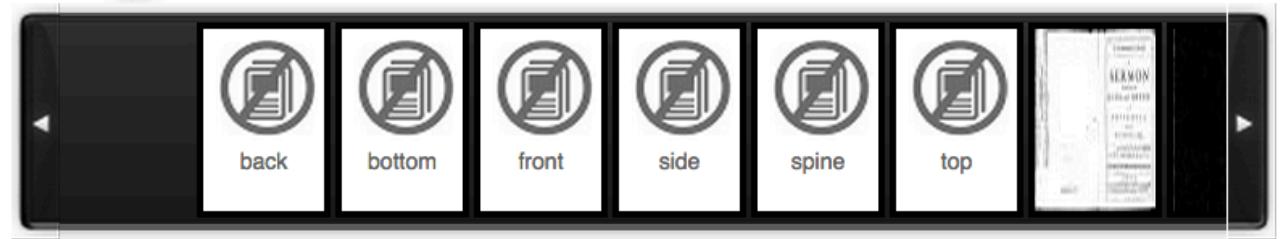


COMPARE BOOKS

- Of religious melancholy a sermon preach'd before the Queen at White-hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy a sermon preach'd before the Queen at White-Hall
- Of religious melancholy A sermon preach'd before the Queen at White-Hall
- Of religious melancholy A sermon preach'd before the Queen at White-Hall
- Of religious melancholy A sermon preach'd before the Queen at White-Hall
- Of the immortality of the soul a sermon preached before the King and Queen at White-Hall

COMPARE

RESET



Of religious melancholy a sermon preach'd before the Queen at White-Hall by Moore, John, 1646-1714
(None)

Open Access Repository of the University of Texas at Tyler 1646-1714



1x ▲ Page 1-9 of 18



1x ▲ Page 2-10 of 18



1x ▲ Page 7-15 of 26



[Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714
(None)

[Of religious melancholy a sermon preach'd before the Queen at White-Hall](#) by Moore, John, 1646-1714



Transcriptions

(-) OF Religious Melancholy. A SERMON Preach'd before the QUEEN AT WHITE-HALL, March the 6th, 1691/2. By the Right Reverend Father in God, John, Lord Bishop of Norwich. Published by Her Maiesty's special Command. The Second Edition. LONDON: Printed for William Rogers at the Sun over aigainst St. Dunstan's Church in Fleetstreet. 1692.



Transcriptions

A- A A+



Transcriptions

Processes



tesseract-ocr

An OCR Engine that was developed at HP Labs between 1985 and 1995... and now at Google.

 Search projects[Project Home](#)[Downloads](#)[Wiki](#)[Issues](#)[Source](#)[Summary](#)[People](#)

Project Information

Starred by 4023 users
[Project feeds](#)

Code license
[Apache License 2.0](#)

Labels
OCR, Utility, CPlusPlus, Google

Members

[theraysm...@gmail.com](#),
[david.e...@gmail.com](#), [tmb...@gmail.com](#),
[breidenb...@gmail.com](#)
12 committers

Featured

Downloads

[tesseract-3.02.02-win32-lib-include-dirs.zip](#)
[tesseract-ocr-3.02-vs2008.zip](#)
[tesseract-ocr-3.02-win32-portable.zip](#)
[tesseract-ocr-3.02.02-doc-html.tar.gz](#)
[tesseract-ocr-3.02.02.tar.gz](#)
[tesseract-ocr-API-Example-vs2008.zip](#)
[tesseract-ocr-setup-3.02.02.exe](#)

[Show all »](#)

Wiki pages

[3rdParty](#)
[AddOns](#)
[FAQ](#)
[ImproveQuality](#)
[ReadMe](#)
[TrainingTesseract3](#)
[Show all »](#)

Tesseract is probably the most accurate open source OCR engine available. Combined with the [Leptonica Image Processing Library](#) it can read a wide variety of image formats and convert them to text in over 60 languages. It was one of the top 3 engines in the 1995 UNLV Accuracy test. Between 1995 and 2006 it had little work done on it, but since then it has been improved extensively by Google. It is released under the [Apache License 2.0](#).

- [ReadMe](#) - Installation and usage information.
- [Compiling](#) - How to build Tesseract on a variety of platforms.
- [FAQ](#) - Common questions and problems. Please check before [filing a bug](#) or [consulting the forum](#).
- [Too many errors?](#) - See the guidance on getting the best out of Tesseract.

Supported Platforms

Tesseract works on Linux, Windows (with VC++ Express or CygWin) and Mac OSX. See the [ReadMe](#) for more details and install instructions. It can also be compiled for other platforms, including Android and the iPhone, though these are not as well tested platforms. See also the [AddOns](#) page for other projects using Tesseract on various platforms.

If you're interested in supporting other platforms or languages, please get in touch with Ray Smith or the [Developers](#).

A Note about Downloads

With the discontinuation of downloads at code.google.com, new source downloads will be posted to [GoogleDrive](#). Other download folders will be setup as new files are uploaded, and the original Downloads page will go away. During the transition, other downloads can still be found at the [Old Downloads](#) page.



IDHMC
TEXAS A&M UNIVERSITY

Initiative for Digital Humanities, Media, and Culture

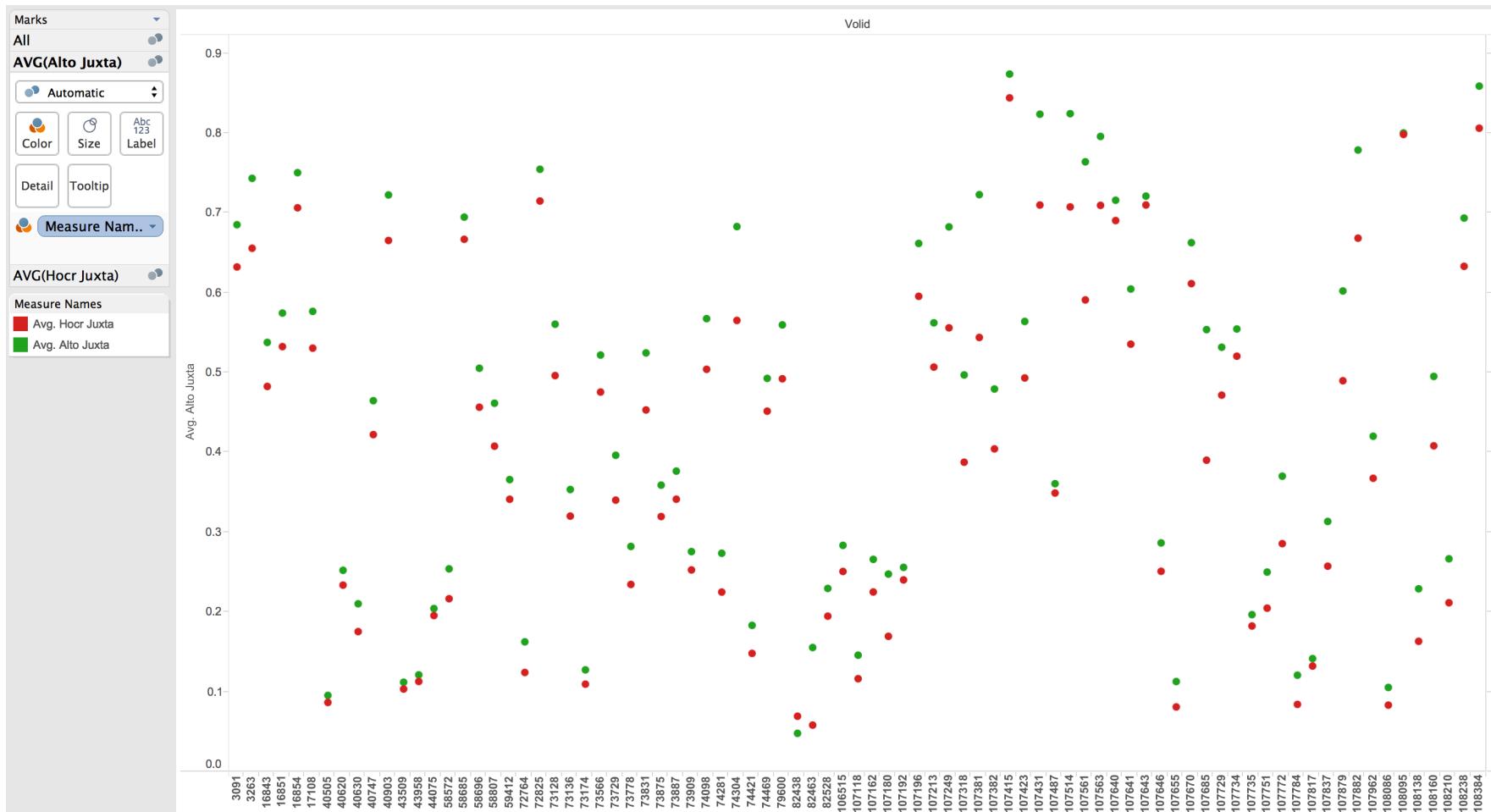


Franken+

About

Franken+ was developed by Bryan Tarpley at the Initiative for the Digital Humanities, Media, and Culture at Texas A&M University. It is a specialized tool designed to allow users to perform OCR on historic fonts. Franken+ ingests output from [PRIMALab's Aletheia](#) (currently, only output from Aletheia version 2.1 is supported, which is still available for download on their site under "previous version"). Aletheia allows the user to open a scanned text document, binarize it, and draw boundaries around each individual character (among many other things). Franken+ currently only works with Aletheia projects wherein the characters have been outlined using polygons (not boxes). Once each character has been bounded appropriately, Aletheia saves the project as an XML file which adheres to PRIMA's [PAGE XML format](#). Provided the binarized image and the resulting XML file generated with Aletheia, Franken+ extracts individual .tif images for each letter blocked-out using Aletheia, giving the user the opportunity to hand-pick the best instances of each letter (thus producing a "font" consisting of only hand-picked images). Using this font, Franken+ can then create synthetic TIF images of text "printed" using this font, with corresponding [BOX files](#), which are then used to train [Google's open-source Tesseract OCR engine](#) in order to OCR images of documents printed with the relevant historic font. Using these synthetic images and their corresponding BOX files, Franken+ then automates the Tesseract font training process and allows a user to test this font.

PageCorrection Tool



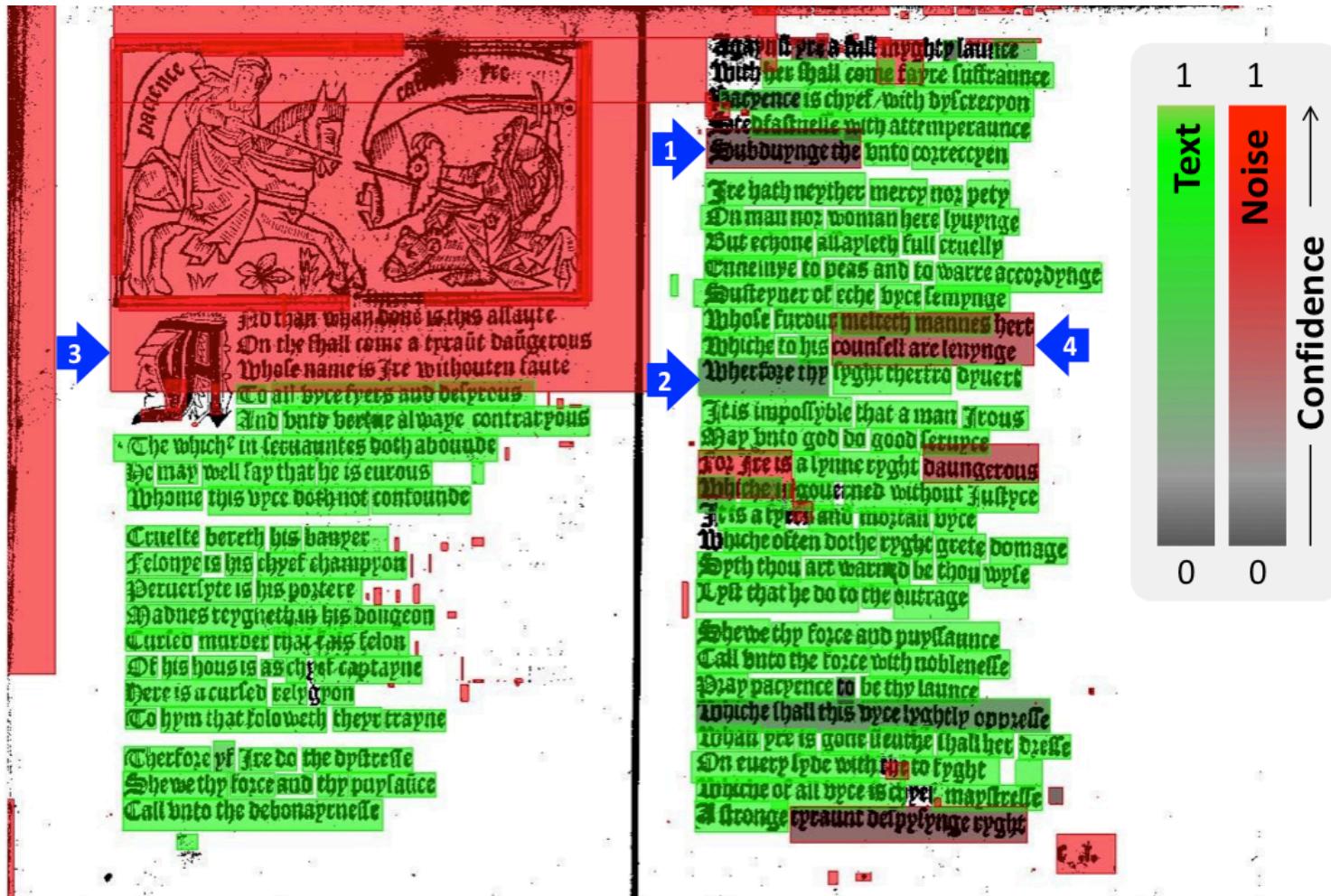


Fig. 6. Iterative relabeling results for a test image. Color denotes MLP confidence: the more saturated, the higher the confidence. Red: noise; green: text

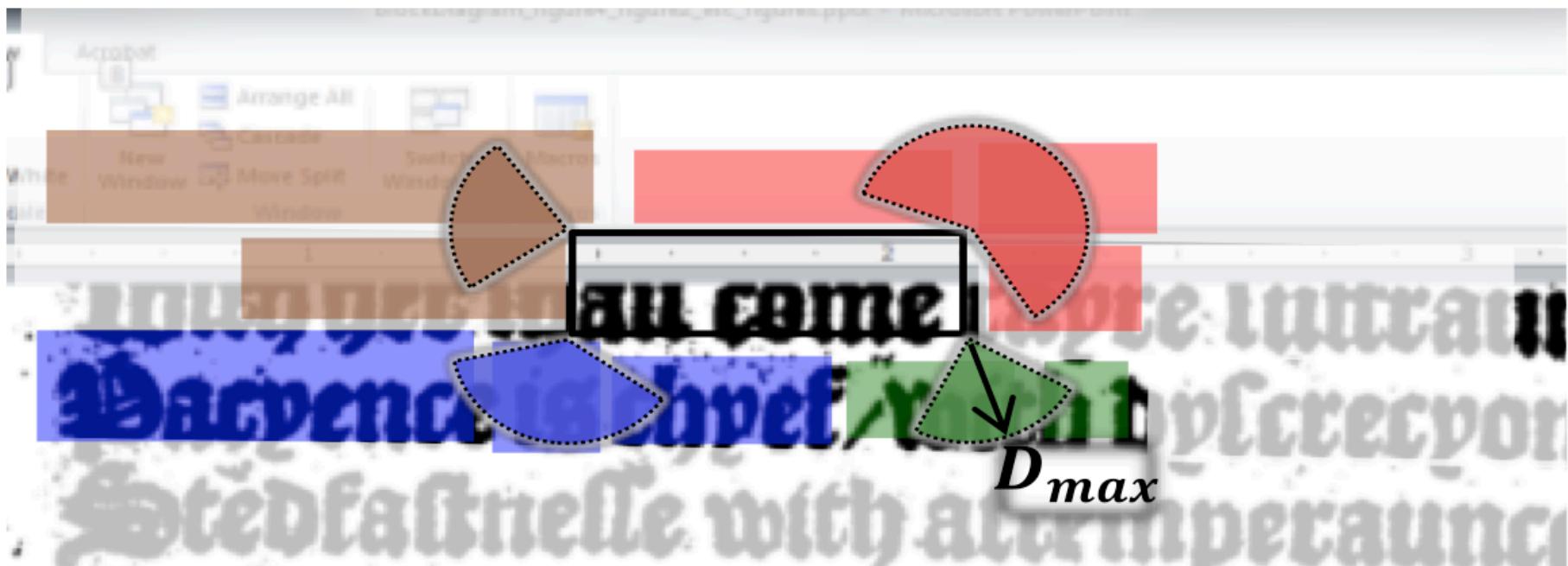


Fig. 3. Finding nearest neighbors. Only those within D_{max} from the corners of the target BB (outlined) are considered. Colors indicate the corner to which neighbors are assigned

Automatic assessment of OCR quality in historical documents

Anshul Gupta¹, Ricardo Gutierrez-Osuna¹, Matthew Christy², Boris Capitanu³,
Loretta Auvil³, Liz Grumbach², Richard Furuta¹, and Laura Mandell²

¹Department of Computer Science and Engineering, Texas A&M University

²Initiative for Digital Humanities, Media, and Culture, Texas A&M University

³Illinois Informatics institute, University of Illinois at Urbana-Champaign

{anshulg,rgutier,mchristy,egrumbac,furuta,mandell}@tamu.edu, {capitanu,lauvil}@illinois.edu

Abstract

Mass digitization of historical documents is a challenging problem for optical character recognition (OCR) tools. Issues include noisy backgrounds and faded text due to aging, border/marginal noise, bleed-through, skewing, warping, as well as irregular fonts and page layouts. As a result, OCR tools often produce a large number of spurious bounding boxes (BBs) in addition to those that correspond to words in the document. This paper presents an iterative classification algorithm to automatically label BBs (i.e., as text or noise) based on their spatial distribution and geometry. The approach uses a rule-base classifier to generate initial text/noise labels for each BB, followed by an iterative classifier that refines the initial labels by incorporating local information to each BB, its spatial location, shape and size. When evaluated on a dataset containing over 72,000 manually-labeled BBs from 159 historical documents, the algorithm can classify BBs with 0.95 precision and 0.96 recall. Further evaluation on a collection of 6,775 documents with ground-truth transcriptions shows that the algorithm can also be used to predict document quality (0.7 correlation) and improve OCR transcriptions in 85% of the cases.

from microfilm converted from photographs –four decades and three generations away from the originals.

Motivated by these issues, in 2013 we started the Early Modern OCR Project (eMOP; <http://emop.tamu.edu>) with funding from the Andrew W. Mellon Foundation. eMOP is a two-year mass digitization project that seeks to improve OCR for some 45 million pages from the Eighteenth Century Collections Online (ECCO) and Early English Books Online (EEBO) proprietary database products. Beyond producing accurate transcriptions for these collections, eMOP aims to create tools (dictionaries, workflows, and databases) to support scholarly research at libraries and museums. Much like our team, these organizations lack the resources to manually transcribe their collections or contract with commercial OCR services (e.g., Prime Recognition Corp.) As such, and as required by Mellon, all tools used and produced by eMOP must remain free or open-source.

As a step towards this goal, this paper describes an approach to assess the quality of historical documents that

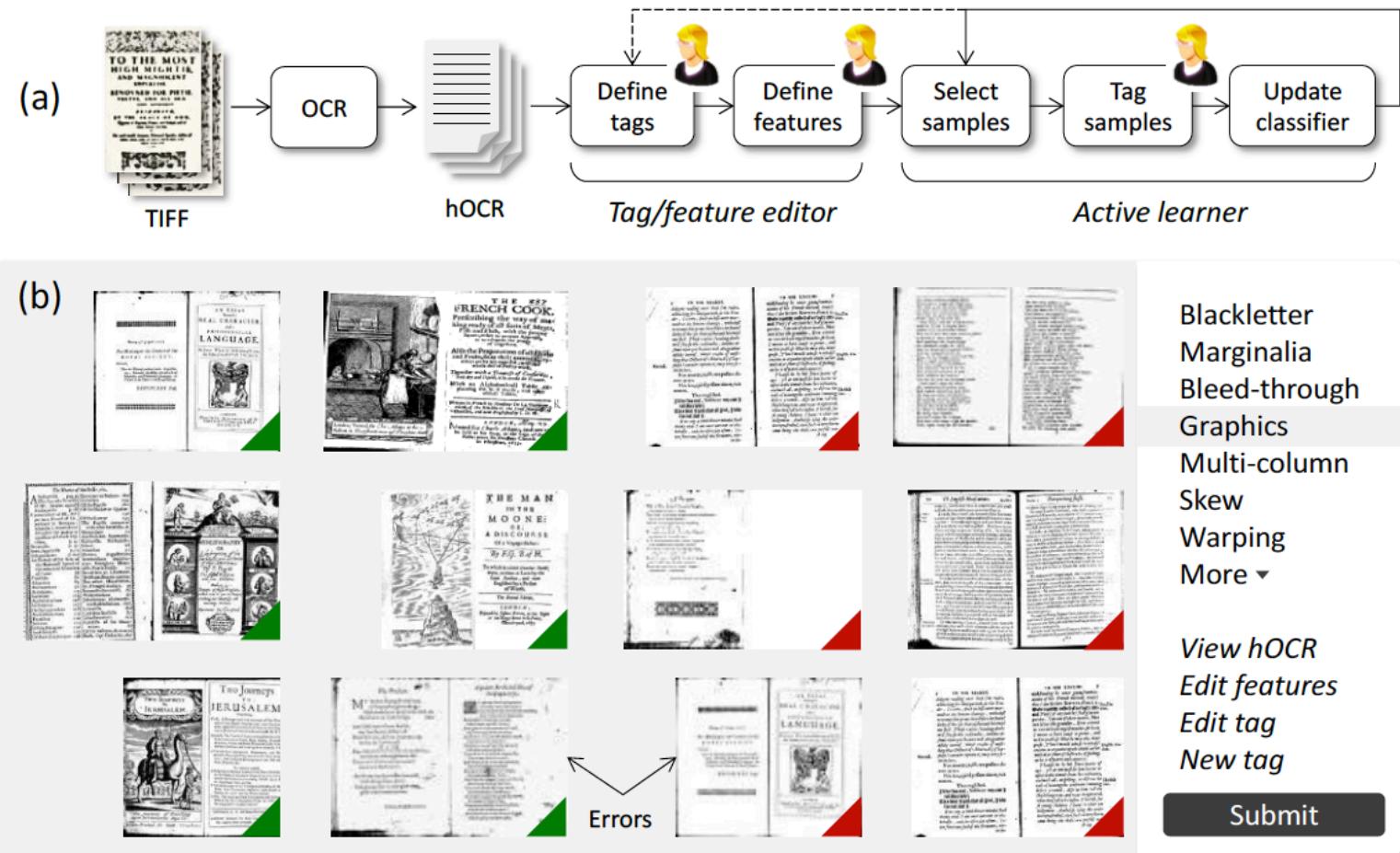


Fig. 1. (a) Stages in the interactive machine learning pipeline. (b) Mockup of the tagML interface

Results

EEBO
Sample

240

4

Theſe things being obſerved it muſt be conſidered, that the Facts now in effect charg'd againſt the Dutchels, are ſuggeſted to be done many years ſince, and were debated and conſidered before the Peers, before the rejeſtion of the two firſt Bills, and long before the Agree‐ment for putting an end to all Controversies and Debates. Now 'tis very hard to put the *Dutchels* to account again, for thoſe very Facts, conſidering that after eight or nine years, many People are dead, others diſperſed, and not in a little time to be found out, and circumſtances of times and places (which in the nature of all these Cafes is almoſt all that is left to discover a *Dutchel*) forgo.

The Dutches is also under the unavoidable necessity of proving a Negative against down right Swearing, and that without any matter Introductory, and also by two mean Servants, the one turned out of her Place for keeping company with Foot Soldiers, and the other leaving his Place in disgust, because he could not have the Play Money.

Matters are already too much in the power of their Servants, and if they charge their Matters with Adultery, Felony and even Treason, it is not easily in the power of the Master to defend himself against down right Swearing; Servants having those opportunities of the knowledge of times and places and Company, which cannot be denied or avoided, and which others have not, whereupon they may frame and build ill Evidence, and many times are ill Principles and desperate Fortunes, and of tempers very Revengeful, so that whenever turns away a Servant, he is in his power for his Estate, Honour and even Life it self. And therefore for the safery and freedom of Families in other Countries, they are not permitted to be Evidence against their Masters, in any matter criminal whatsoever.

Note also the Dutches thinks, She as an English Subject has by *Magna Carta*, the same right to a Tryal in the Legaland ordinary way of Justice, as the rest of the Kings Subjects, especially considering that there never yet was in England one Precedent, of a Bill exhibited in Parliament to dissolve a Marriage at the first instance, without any previous proceedings in the Ecclesiastical Courts; (which in these Cases is the Law of the Land) and not above Five or Six Bills, in above 600 years ever passed to dissolve Marriages, or make Second Marriages good, even after there had been a divorce in the Spiritual Courts, and thofe Bills too in Cafes generally where the Husbands were without any recummittion.

No Impediment appears why the Duke should not have endeavoured a Divorce at Law, before he had attempted a Bill to dissolve the Marriage. For any Application to the Legislature, for Tryal of matters of Private right, is improper, because there are proper Judicatures, that give that remedy the Law allows.

To ask any thing of the *Legislature*, in private Causes between Party and Party, beyond what the known Laws in force give, seems to be improper, because

1. It is to make a Law in one persons Cafe, which is not the Law in another.

2. It is to retrospect actions, and after the Fact to Augment or alter the Penalty. Our Laws are certain, and known, that persons may conform their actions to them.

3. The Courts and Forms of Law are equally the Subjects right, with the Law, and the application to the Legislature, takes away all that right of Form and Charging, whereas a certainty in time to Answer, and exceptions to Witnesses, and other Legal defences of the Fact, cannot be denied the meanest Subject.

The consequence that short and summary ways of proceedings, may have on the Settlements of Estates and Families, may be very dangerous, and though the Duke of Norfolk's Bill passed the House of Peers, yet a great number of the Peers both Spiritual and Temporal entred their Protest against its Passing, and their reasons for it.

It ought to be no prejudice to the Dutches, that the next Heir presumptive to the Duke, doth nor as yet appear to be a Protestant, because when a Bill which hath lately passed both Houses is become a Law, it is not likely that Noble Family will be without a Protestant at the head of it.

If any of the Witneses formerly examined on the Bill in 1691, be now again produced, it is humbly apprehended, that as the validity of their testimony they then gave was totally overthrown, by a great number of Witneses, so it will be again tho' under the disadvantage of the great distance of time.

F I N I S

Work: The case of Mary dutchess of Norfolk

Batch: 738: EEBO Doc for Larua - corrected -SC...

Print Font: Training Test Suite

[<< Back To Dashboard](#)

[Select All](#)

[Schedule Selected](#)

[Set Print Font for Work](#)

						◆	Page Number	◆	Juxta	◆	RETAS	◆
<input type="checkbox"/>	●	◀	▶	📄	↔	🔗	1		0.867		0.779	
<input type="checkbox"/>	●	◀	▶	📄	↔	🔗	2		0.824		0.73	
<input type="checkbox"/>	●	◀	▶	📄	↔	🔗	3		0.828		0.716	

[<< Back To Dashboard](#)

1- It is to make a Law in one persons Case, which is not the Law in ano-
thers. 2. It is to retrospect actions, and after the Fact to Augment or alter tho Pe-
nalty. Our Laws are certain' and known, that persons may conform their
actions to them-

' The Coutts and Forms of Law are equally the Subjects right with tho Law,
and the application to the Legislature' takes away all that tiglic of Foism and
Charging, whereas a certainty in time to Answer, and exceptions to Witnesses'
and other Legal defences of the Fact' cannot be denied the meanest
Subject.

The consequence that short and summary ways of proceedings, may have ori
the Settlements of Estates and Families, may be very dangerous' and though
the Duke of Nors?)lk's Bill nailed the House of peerffi yor great number Of the
years both spiritual and Temporal entred their Protest against its Passing' and
their reasons for it.

It ought to be no prejudice to the Dutchess, that the next Heir presumptive
to tho Duke, doth not as yet appear to be a Protestant, because when a Bill
which hath lately palled both Houses is become a Law, it is not likely that
Noble Fatuity will be without a Protestant at the head of it.

-, be now a ain

produced' it is humbly apprehended' that as the validity of their testimony they
then gave was totally overthrown, by a great number of Witnesses, so it will
be again tho' under the disadvantage othe great distance of time

(II)

Thro' Error's maze a Leader she would prove,
And light the Lamp of pure celestial Love?

On Thule's heathy heights see VIRTUE stand,
With ling'ring looks, prepar'd to leave the Land!
To waste in northern night the lonely hour,
And mourn 'mid Hecla's caves her ravish'd pow'r:
See fond Affection warn her e'er she flies,
And turn to you her supplicating eyes;
Where Hope still struggles thro' each chrystal Tear,
That watry shines, and streaks the gloom of Fear!

ECCO
Sample

Recall her with your smiles, confirm her reign,
The trait'rous Rival of her empire slain!
For yours, for Albion's sake retard her flight;
When Virtue sets, 'tis Beauty's blackest night.
In vain you'll mourn the absence of her ray,
To ruthless Lust an unresisting Prey!
Find civil Order to Confusion hurl'd,
And mental Darkness overwhelm the world!

D

Again

Work: On illicit love. Written among the ruins of Godstow nunnery, near Oxford. By John Brand, ...

Batch: 735: ECCO Doc for Laura - SC-wCAS-D2-Rev

Print Font: Test Training Suite-ECCO

[<< Back To Dashboard](#)

[Select All](#) [Schedule Selected](#) [Set Print Font for Work](#)

						Page Number	Juxta	RETAS
						1	0.078	0
						2	0.74	0.594
						3	0.866	0.783
						4	0.784	0.576
						5	0.955	0.75
						6	0.935	0.778
						7	0.922	0.762
						8	0.84	0.81
						9	0.846	0.764
						10	0.929	0.752
						11	0.963	0.785
						12	0.96	0.772
						13	0.94	0.752
						14	0.935	0.814
						15	0.902	0.691
						16	0.943	0.802
						17	0.963	0.843
						18	0.955	0.836
						19	0.89	0.684
						20	0.97	0.789
						21	0.561	0.315

[<< Back To Dashboard](#)



(11)

Thro' Error's maze a Leader she would prove'
And light the Lamp of pure celestial Love?

On Thule's heathy heights fee VIRTUE stand,
Wich ling'ring looks' prepar'd to leave the Land!
To 4waste in northern night tho lonely hour'

And mourn 'mid Hecla's caves her revih'd pow'r:
See fond Affection warn her e'er ihe flies'

And turn to you her supplicating eyes;

Where Hope still struggles thro' each chrystal Tear'
That watry ihines' and streaks the gloom of Fear!

Recall her with your smilos' confirm her reign'
The trait'rous Rival of her empire ilain!
For yours' for Albion's fake retard her flight;
When Vir'tue sets' 'tis Beauty's blackest night-
In vain you'll mourn the absence of her ray'
To ruthless Lust an unrelisting Prey!
Find civil Order to Confufion hurl'd,

And mental Darkness overwhelm tho world!'
D

[3]

IT may be Obje^cted, That the Burgeesses who returned the Sitting-Members had not the Precept with them to produce to the Sheriffs when they made the Return; and that the Petitioners had the Precept, and produced it with their Return, which was tendered to the Sheriff, and refused.

IT is Answered, First, That this Fact is not insisted upon in the Petition, nor made any Part of the Petitioners Complaint; and it is apprehended, that no Objection can be taken in the House but what is contained in the Body of the Petition.

Secondly, BUT if this Fact be admitted, it is insisted, that by the 23 H. 6. ch. 15. it is directed, that the Precept be returned by Indenture; but it is not necessary the Precept should be annex'd to the Indenture, or delivered to the Sheriff with it, and the Practice is generally otherwise; for in this present Parliament, there are near Three Hundred Members who have no Precepts annexed to their Returns. An Indenture between the Sheriff and the Returning Officer, mentioning the Election to have been made by Virtue of a Precept to him directed, is certainly good, which shews the Precept need not to be annexed to the Indenture.

IF it be Objected, That tho' the Precepts are not always returned by the Sheriffs to the Clerk of the Crown, yet they must be delivered to the Sheriff by the Returning-Officers; it may be Answered, That it being not necessary for the Sheriff to return the Precept with the Indenture to the Clerk of the Crown; neither can it be necessary to be return'd to the Sheriffs by the Parties with their Indenture; for this being a Matter which concerns the Sheriff's personal Satisfaction only, it remains absolutely in his Breast, whether he will call for it or not; and it can be no Error if he accepts a Return without it; and in common Cases of Proofs, the Officer that hath the Sheriff's Warrant, doth not after Execution of it, deliver it back to the Sheriff.

WE apprehend, that the Precept being by this Petition admitted to be directed to the Burgeesses and Electors only, and the Return which was tendered by the Petitioners, being also thereby admitted to be a Return of the Constables and Burgeesses; this Return was not made by the Persons to whom the Precept was directed; and therefore was not proper to be received by the Sheriff, nor such as the House can now accept; for the Precept ought to be returned by the same Persons to whom it is directed.

IT may be Insisted, that these are Points in Law proper to be argued by Council; and therefore fit to be referred to a Committee, and the Merits of the Election to be Heard at the same Time.

AS to the Sheriff's proceeding in this Matter, it was with the greatest Fairness; for the Return of the Sitting-Members, was brought to him about Seven or Eight in the Evening, of the Day of the Election, and he did not forward it for London till the next Day.

SINCE the Printing what is above, we have certain Information, That the Precept was not produced to the Sheriff, with the Return of the Petitioners; neither was there ever any Precept Return'd, with any of the Indentures, for former Elections in Minehead; or any Precept Return'd, with any of the Indentures, for any Election for this Parliament, in any Borough of the County of Somerset, of which Minehead is one.

Using Prime Recognition

A N at all Times' during the Memory of Man'' Whenever any E-
.lectiōn' Was Controverted (till the vlyast Parliament of Queen Anne) 't'lie
Constables never so much as pretended to prefide as 'Judges' of che :Poll ;
but the Candidates of each Side, Chose out othe Body othe Burgessses,
such Persons as they Agreed upoil to direct and take tho Poll.
A T tho last Election, the Precept' according to the Ancient Usage,
was directed to the Burgolses and Electors in General ;v who in pursuance
of the Precept' Elected che Sitting-Members, 'Mr''. Edwin' and Mr. Gage,
by a fair and undoubted Majority, and by their Indenture duly Exe-
cuted, Returned them to the Sheriff accordingly-
The Petition.
BUT now a Petition is prefered in the Lathe of constables'
and several Voters' in behalf of themselves and Other Voters of the-
baid Borough, against this Return; Suggesting' That the Constables are
the Presiding and Returning ('fficers' and that the Precepts are directed
Gage' Esquires.
livered to him as a Burgess only. fai'-g;iz:
and deliver'd to them; and that the 'now-Sh'eriff directed his -Pre
Burgensem keys Electrinus Burgi de Minhead ; but deliver
the Constables, and took his Receipt for it.
TH AT the Election being upon tho 11th of -April last, the Candi-
dates wore Sir John Trevelyc'm, James Milner-' Samuel Edwin' and Thomas
copt
'd lt to Offer ('f
T 1-1 A T the Petitioners, the Constables, Declared Sir John Trevegyazi,
and Mr. Milner, duly Elected ;j and thereupon Indentures were duly
Executed by che Constables' and Majority othe Burgeffes and Electors,'
and tendered to the Sharest by order of the Constables : But the Sheriff
refused to receive it; but has Return'd Mr. Edwin and' Mr. Gage,. by
Virtue- of other Indentures brought to the Sheriff' not by the proper Re-
turning slicers of tho said Borough-, but by other Persons who Signed
it,- toth prejudice Of the Petitioners'; who therefore Pray' that Return'
may be taken off, and the Return of the Constables put upon the File.'.-
UPGN this Petition it may be Observ'd;
T H E Preamble of the Petition sets forth' That John Jones and George
James are Constables of the Borough of Minehead, in rder to make
good their Claim of being Returning.-Officers.
BUT in the First' Place it is Answered' that they are not Consta-
bles of the said Borough' Which consists of several Manors and Tythins,
Which have their proper officers' and Wherein-the Constables have no
Jurisdiction.
THE Sheriff could not direct his Precept to the Constables of the
Borough of Minehead, because there are no such Officers; and a Pre-
cept to them directed, Would be a Void and improper Directim; and
any Return made by them must be consequently so too-
T HAT the Merit Of the Petitioners Case is founded solely on the
Constables being the Returning-officers of this Borough; and there-
fore' in che First Place, they ought to be put upon the Proof of this
Allegation, That they are tho Ret'ning-officers, which for 'the
Reasons afore-given' we think impossible to 'be maintained-
T may be objected' That the Sheriff, by delivering-the Precept to
one Of the Constables, hath taken Notice that 'they are che Returning
officers; to Which it is answered, That the Person to whom the'Pre-
cept happen'd to be deliver'd, Was a Burgess as well as Constable;
and the Precept Was directed to the Burgeffes and Electors' and de-
ITV.

Completely Open Source (and not yet done)

stans
s-' J
j IT may be Objected, That the Burgesses Who returned the Sitting.;
Members had not the Precept with them to produce to the Sheriff;
when they made the .Return; and that the Petitioners had the Precept,
and-produced lcwlth their Return, which was tendered to the Sheriff
and refused.
I T'--i's Answered, First, That this Fact is not insisted upon in the Pe-
tlthn, nor made any Part- of the Petitioners Complaint; and ic
is apprehendeth that no ('bje'cti'on can be taken anthi House but
What is contained in the Body ofthe Petition.
Secondly, B U T if this Fact be admitted,' it is insisted, that bythe
2; H. 6. ch. 15- iris directed' that tho' Precept be returned by In-
denture; but it is not necessary the Precept should be annex'd to
the Indenture' or delivered to the Sheriff with it' and the Practice is
generally otherwise; for in this present P'arliament' there are near
Three Hundred Members who have nojPrecepts annexed to their
Returns. An Indenture between the Sheriff and the Returning
officer, mentioning the Election to have been made by Virtue of a
Precept to him directed, is certainly good, which lhews'the Pre-
cept need not to be annexed to che Indenture.
I F it be objected., That tho' the Precepts are not always returned by
the Sheriffs to the Clerk of the Crown, yet they must be delivered to the
Sheriff bythe Returning-Officers; it may be Answered, That it being
not necessary for the Sheriff to return che Precept With the Indenture to
the Clerk of the Crown ; neither can it be-necefl'ary to be return'd to the
'Sheriffs by- the Parties ''with their Indenture; for this being Matter
which concerns the sheriff's personalisatisfaffion only, it remains dabs.-
lately in his Breast, whether he will call for it ornot; and it can be no
Error if he accepts a Return Without it ; and 'in common Cases of Process,
the 'officer that hath- the Sheriff's Warrant,- doth not after Execution. of
it, deliver it back to the Sheriff.
WE apprehend, that the Precept being by this Petition admitted
to be directed to tho Burgessos and Electors only, and the Return which
was tendered by the Petitioners, being also thereby admitted to be a
Return of tho Constables and Burgesses; this Return was not made by tho
Persons to whom the Precept was directed ;- and therefore was not pro-
perty be received by the Sheriff' nor such as the House can now accept;
for tho Precept ought to be returned by the fame Persons to whom it is
directed.
I T may be Insisted' that there are Points in Law proper to be argued
by Council; and therefore fit to be referred to a Committee' and tht7 Me-
rits of tho EloctiOn to be Heard at the fame Time-,-
A S to the Sheriff's proceeding in this Matter' it Was With the great-
est Fairness; -for the Return of the Sitting-Members, was brought to
him about Seven or Eight in che Evening, of the Day of the Election,
and he did not forward it for Landon till the next Day.
S I N C E the Printing What is above'. we have certain Information'-
That the Precept Was'no-t produced to the Sheriff' with the Return of
che Petitioners; neither was there ever any Precept Return'd, with any
of the Indentures, for former Elections in Mimhead ; or any Precvept Res-
turn'd,- with any of che Indentures, for any Election for this P'a'rliamenc'
in any Borough ofthe County of Somerset,- of which Winehead is one-

1549100500_3-Gale

Change



denture; but it is not necessary the Precept~ shoud be annex'd to the Indeature, or delivered to the Sheriff with it, and the Pra~tice is generally otherwise ; for in this present Parliaarent, there are near Thkree Hundred Members who have no Precepts annexed to their i ~Returns. An Indenture between the Sheriff and the Returning i Officer, mentioning the Eleaion to have been made by Virtue of st Precept to him dire~ted, is certainly good, which shews the Precept need not to be annexed to the Indenture.

IF it: be Objc~ted, That tho' the Precepts are not always returned by the Sheriffs to the Clerk of the Crown, yet: they must be delivered to the Sheriff by the Returning-Officers; it may be: Answered, That it being not necessary for the Sheriff to return the Precept with the Indentur~e to the Clerk of the Crown ; neither can it be neceffary to be return'd to the Sheriffs by the Parties with their Indenture; for this being a Matter which concerns the Sheriff~s personal Satisf~tion only, it: remains abfo~ately in his Breast, vihether he will call for it or not; and it can be no Error if he accepts a Return without: it ; and in common Cafes of Procefis, the Officer that hath: the Sheriff's Warrant doth not after Execution: of it, deliver it back to the Sheriff.

WE apprehend, that the Precept; being by this Petition admitted to be direted to the Burgesses and Elet~ors only, and the Return which was tendered by the Petitioners, being also thereby admitted to be a Return of thle Conflables and Burgesses ; this Returh was not made by the Persons to whom the Precept was dir~eaed; and therefore was not proper to be received by the Sheriff, nor s~uch as the Hous~e can now accepti orthe Precept ought to be returned by the famne Persons to whom it is direc~ed. IT may be Insisted, that there are Points in Law proper to be argued by Coulnci ; and therefore fit to be referred to a Corarninee, and the. Merits of the Ele~tion to be Heard at the f'amre Time.

A S to the Sherif~s proceeding in this Matter, it was with the grea~eft Fairness -for the Return of the Sitting-Mdetmbers, was brought td him about Seven or Eight in the Evening, of the Day of the El Sion, and he did not forward it for London till the next Day.

438323_3-ECCO-corrected

Change

the Indenture' or delivered to the Sheriff with it' and the Practice is generally otherwise; for in this present P~arliament' there are near Three Hundred Members who have no Precepts annexed to their Returns. An Indenture between the Sheriff and the Returning officer, mentioning the Election to have been made by Virtue of a Precept to him directed, is certainly good, which lhews'the Precept need not to be annexed to the Indenture.

I F it be objected., That tho" the Precepts are not always returned by the Sheriffs to the Clerk of the Crown, yet they must be delivered to the Sheriff bythe Returning-Officers; it may be Answered, That it being not necessary for the Sheriff to return the Precept With the Indenture to the Clerk of the Crown ; neither can it be neceffl'ary to be return'd to the 'Sheriffs by- the Parties "with their Indenture; for this being Matter which concerns the sheriff's personalisatisaffion only, it remains dabs'o~ately in his Breast, whether he will call for it ornot; and it can be no Error if he accepts a Return Without it ; and in common Cases of Process, the 'officer that hath- the Sheriff's Warrant,- doth not after Execution, of it, deliver it back to the Sheriff.

WE apprehend, that the Precept being by this Petition admitted to be directed to the Burgessos and Electors only, and the Return which was tendered by the Petitioners, being also thereby admitted to be a Return of the Constables and Burgesses; this Return was not made by the Persons to whom the Precept was directed ; and therefore was not property be received bythe Sheriff nor such as the House can now accept; for tho the Precept ought to be returned by the fame Persons to whom it is directed.

I T may be Insisted' that there are Points in Law proper to be argued by Council; and therefore fit to be referred to a Committee' and tht7 Merits of the Elocti0n to be Heard at the fame Time,-,

A S to the Sheriff's proceeding in this Matter' it Was With the greatest Fairness; -for the Return of the Sitting-Members, was brought to him about Seven or Eight in the Evening, of the Day of the Election, and he did not forward it for Landon till the next Day.

S I N C E the Printing What is above'. we have certain Informration'

Gale

eMOP

Not yet implemented:

1. All the dictionaries and n-gram work performed by SEASR at the University of Illinois have not yet been implemented due to problems on our end;
2. Printer's Database: specific font training;
3. Image Detection and Optimization;
4. Optimizing the OCR engine itself.