

PART II—APPENDIX B: COMPARING PRE-1923 DIGITIZATIONS IN ACLS HUMANITIES E-BOOK WITH DIGITIZATIONS OF THE SAME TEXTS IN GOOGLE BOOK SEARCH AND MICROSOFT LIVE SEARCH BOOKS

Appendix B compares digitizations of 10 pre-1923 volume from ACLS Humanities E-Book (HEB) with digitizations of the same volume from either Google Book Search (GBS), Microsoft Live Search Books (MLSB), or both.

1) NUMBERS OF TITLES AND VOLUMES SEARCHED AND SELECTED

Number of volumes reviewed

Total: 25

GBS: 8

[entries 204A, 205A, 206A, 207A, 208A, 210A, 211A, 212A]

MLSB: 7

[entries 204B, 205B, 206B, 207B, 208B, 209A, 213A]

HEB: 10

[entries 204C, 205C, 206C, 207C, 208C, 209B, 210B, 211B, 212B, 213B]

2) ACCESS

OPAC records

Number of volumes for which OPAC did not contain links for retrieval of records

GBS: 3 (37.5%)

[entries 207A, 210A, 211A]

MLSB: 7 (100%)

[entries 204B, 205B, 206B, 207B, 208B, 209A, 213A]

HEB: none

3) PHYSICAL CONDITION OF VOLUMES REVIEWED

All three projects contained volumes in which the physical condition affected OCR reading significantly.

GBS: 1 (12.5%)

[entry 206A]

MLSB: 2 (28.6%)

[entries 206B, 209A]

HEB: 4 (40%)

[entries 206C, 207C, 209B, 212B]

4) METADATA

Number of pages

Number of volumes with listings for number of pages different than that given in authoritative sources (LC, Harvard Library's Hollis Catalog, last page of text)

GBS: none

MLSB: 7 (100%)

[entries 204B, 205B, 206B, 207B, 208B, 209A, 213A]

HEB: none

Page indicators

Number of volumes with discrepancies between numbers listed in the page indicator and actual page numbers appearing in the text

GBS: 1 (12.5%)

[entry 210A]

MLSB: none.

HEB: 2 (20%)

[entries 204C, 205C]

Subject headings

Number of volumes that used LC main subject headings but not the subheadings provided by LC

GBS: 6 (75%)

[entries 204A, 205A, 207A, 208A, 210A, 211A]

MLSB: none.

HEB: none.

Number of volumes that either omitted some or all LC headings; used headings not derived from LC; used no subject headings even though LC provided some; or disjointed LC subject headings so that subdivisions became separate headings

GBS: 2 (25%)

[entries 206A, 210A]

MLSB: 2 (28.6%)

[entries 204B, 208B]

HEB: none.

Contents links

Number of volumes with bad contents links

GBS: 2 (25%)
[entries 207A, 208A]
MLSB: 3 (42.9%)
[entries 205B, 206B, 209A]
HEB: none

5) QUALITY OF SCANS

Number of volumes rated “poor”

GBS: 2 (25%)
[entries 210A, 211A]
MLSB: 2 (28.6%)
[entries 205B, 209A]
HEB: 2 (20%)
[entries 204C, 205C]

Number of volumes in which OCR reading was affected adversely on 3 or more pages

GBS: 6 (75%)
[entries 206A, 207A, 208A, 210A, 211A, 212A]
MLSB: 7 (100%)
[entries 204B, 205B, 206B, 207B, 208B, 209A, 213A]
HEB: 7 (70%)
[entries 205C, 206C, 207C, 208C, 210B, 212B, 213B]

Missing pages

Number of volumes with missing pages

GBS: none
MLSB: 1 (14.3%)
[entry 205B]
HEB: none

Illegible pages

Number of volumes with one or more illegible pages

GBS: none
MLSB: 1 (14.3%)
[entry 209A]
HEB: none

Pages not read by OCR

Number of volumes with one or more pages not read by OCR

GBS: none
MLSB: 1 (14.3%)
 [entry 211A]
HEB: none

Pages with portions cut off or obscured

Number of volumes with portions of one or more page cut off or obscured

GBS: none
MLSB: 2 (28.6%)
 [entries 204B, 208B]
HEB: none

6) ACCURACY OF WORD SEARCHING WITHIN TEXTS

OCR misreadings due to footnote indicators

Number of volumes in which searching was inaccurate because footnote indicator were read by OCR as parts of words

GBS: none
MLSB: none
HEB: 3 (30%)
 [entries 207C, 208C, 213B]

OCR misreadings in indexes, footnotes, and bibliographies

Number of volumes in which searching was inaccurate because OCR did not read words of text correctly in indexes, footnotes, or bibliographies

GBS: 4 (50%)
 [entries 205A, 207A, 210A, 212A]
MLSB: 6 (85.7%)
 [entries 204B, 205B, 206B, 207B, 209A, 213A]
HEB: none.

OCR misreading of dates

Number of volumes in which searching was inaccurate because OCR misread the numeral “1” as “I” in dates

GBS: none
MLSB: none
HEB: 1 (10%)
[entry 205C]

OCR misreading due to margin headings

Number of volumes in which searching was inaccurate because OCR read margin headings mixed in with lines from the main body of text

GBS: 3 (37.5%)
[entries 207A, 208A, 210A]
MLSB: 2 (28.6%)
[entries 207B, 208B]
HEB: 3 (30%)
[entries 207C, 208C, 210B]

OCR misreadings due to causes other than those surveyed above

Number of volumes in which searching was inaccurate because OCR misread words for reasons other than those surveyed above

GBS: 2 (25%)
[entries 207A, 208A]
MLSB: none
HEB: none

APPENDIX A: RESULTS OF AN EXAMINATION OF DIGITIZATIONS OF THREE MULTI-LANGUAGE REFERENCE WORKS

201) Bibliographic citation: Riddle, Joseph Esmond, and Thomas Kerchever Arnold. *A Copious and Critical English-Latin Lexicon Founded on the German-Latin Dictionary of Dr. Charles Ernest Georges*. First American Edition, rev. and with a dictionary of proper names by Charles Anthon. New York: Harper & Brothers, 1849.

Project: Google Book Search.

Source of physical book: Harvard University, Widener Library.

Provenance: By exchange of Duplicates with the Rev. [] Perry, Class of 1854.

Pages viewable: full text.

Pages viewed: 4-23, 188-207, 306-325, 596-615, 718-737.

Physical condition: Ink splotches appeared on four pages obscuring words (727-730). Light areas appeared on many pages, but text was readable.

Ease of finding: easy.

Metadata: Page indicators: No page numbers were listed in the page indicator for pages after p. 11. Because of this, retrieving specific pages after p. 11 was possible only by trial and error. Subjects: “English language” and “Latin language”; LC lists “English language--Dictionaries—Latin” and “Latin language--Dictionaries--English.” Key words and phrases listed “Viä,” an OCR misreading of “Vid.”; “Vi<1,” “ineit” (no occurrences); among many other irrelevant combinations of letters that Google presents as words.

MARC records: not available.

OPAC records: no.

Quality of scans: Poor (see above and below). Letters were not sharp, although pages were legible. Following p. 17, pages 8 through 17 were reproduced a second time.

Accuracy of keyword searches: Word searches were impossible. For most pages checked, there were no OCR readings. For the few pages with limited OCR readings, all OCR readings resulted in incomprehensible symbols that were not words.

202) Bibliographic citation: Woodhouse, S. C., comp. *English-Greek Dictionary: A Vocabulary of the Attic Language*. London and New York: Routledge, 1932.

Project: Google Book Search.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: preview.

Pages viewed: 43-47, 50-54, 57-61, 64-68, 71-72; 149-150, 153-157, 160, 163-167, 170-174, 177-179; 325, 328-332, 335-339, 342-346, 348-349, 352-353, 356; 584-587, 591-592, 595, 597, 599, 602-603, 605-607, 610, 612, 614-615; 972-974, 976, 978, 980, 982, 984-985, 989-991, 993, 995-996, 1000-1001, 1003.

Physical condition: good.

Ease of finding: easy.

Metadata: Pages listed: 1038; LC lists “viii, 1029.” Page indicator: 10 pages were mislabeled in the page indicator (149, 150, 153, 154, 155, 348, 349, 1000, 1001, 1003). One page labeled 349 in the page indicator displayed a gloved hand and was completely

unreadable. **Subjects:** “Dictionaries,” “English language,” “Foreign Language Study / Greek (Modern),” and “Foreign Language Study / Ancient Languages”; LC lists “English language—Dictionaries—Greek.”

MARC records: not available.

OPAC records: yes.

Quality of scans: Poor. 4 pages were omitted (347, 348, 350, 351). In addition, the right sides of 6 pages were blurred and/or cut off so as to make one word per line unreadable (974, 976, 978, 980, 984, 1000).

Accuracy of keyword searches: In addition to the problems noted above, words searches achieved perfect results for English language words, with the exception of the last line of p. 58 and words hyphenated between lines.

203) Bibliographic citation: Meusel, H. *Lexicon Caesarianum*. Volume II. Berlin: W. Weber, 1893.

Project: Google Book Search.

Source of physical book: Stanford University Libraries.

Provenance: Leland Stanford Junior University.

Pages viewable: full text.

Pages viewed: 5-44, 150-189, 311-350, 599-638, 999-1038. (Each page contains two columns, each of which is paginated separately.)

Physical condition: good.

Ease of finding: easy.

Metadata: Contents: 24 sections were listed, but these did not correspond to sections of the text. Sections 2-6, for example, linked to pages v, vi, vii, ix, and x for no understandable reason.

MARC records: not available.

OPAC records: no.

Quality of scans: reference.

Accuracy of keyword searches: All pages contained numerous OCR misreadings of Latin words, although many words were read correctly by OCR. Often, OCR read a Latin word mistakenly as a similarly-spelled English-language word. For example, the word “multitudo” was read as “multitude”; “documento” as “document.” In addition, the letter “s” often was misread as “e”: “suis” as “suie,” “sibi” as “eibi,” “ipsi” as “ipei,” “essel” as “eseet,” etc.

**APPENDIX B: COMPARING PRE-1923 DIGITIZATIONS IN ACLS
HUMANITIES E-BOOK WITH DIGITIZATIONS OF THE SAME TEXTS IN
GOOGLE BOOK SEARCH AND MICROSOFT LIVE SEARCH BOOKS**

204A) Bibliographic citation: Barbour, Violet. *Henry Bennet, Earl of Arlington, Secretary of State to Charles II*. Washington: American Historical Association, 1914

Project: Google Book Search.

Source of physical book: Harvard College Library.

Provenance: from the bequest of Charles Sumner.

Pages viewable: full text.

Pages viewed: 5-24, 70-89, 137-156, 200-219, 263-282.

Physical condition: good.

Ease of finding: easy.

Metadata: Subjects: “Great Britain”; LC lists “Arlington, Henry Bennet, Earl of, 1618-1685” and “Great Britain—History—Charles II, 1660-1685.”

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition. NOTE: An odd problem occurred near the end of the volume when clicking on the next page button. The last page before the index is p. 275. A blank page follows, and the index begins on p. 277. When using the double-page view that displays the left and right pages of the book on the same screen, a click on the next page button of the page that displayed 274-275 led to pp. 278-279. In the single-page view, however, a click on the next page button of p. 275 led to p. 277. After a subsequent shift to the double-page view, p. 275 then appeared both as the right-side page of the 274-275 display and the left-side page of the next page display that offered p. 277 on the right-side.

Accuracy of keyword searches: The OCR problems with reading lines on the index pages for the Microsoft digitization of this book (see 204B) did not occur in this Google digitization. In addition, the last footnote on p. 5 was read correctly by OCR in the Google digitization (unlike in 204B), although an italicized word in the footnote, “Oxonianses,” was read incorrectly by OCR as “Ozonianses.”

204B) Bibliographic citation: Barbour, Violet. *Henry Bennet, Earl of Arlington, Secretary of State to Charles II*. Washington: American Historical Association, 1914

Project: Microsoft Live Search Books.

Source of physical book: University of Toronto Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 5-24, 70-89, 137-156, 200-219, 263-282.

Physical condition: Markings appeared on only one page (154) with no affect on OCR.

Ease of finding: easy.

Metadata: Pages listed: 330; LC lists “xii, 303 p.” Subjects: “Arlington, Henry Bennet, Earl of, 1618-1685”; LC lists “Arlington, Henry Bennet, Earl of, 1618-1685” and “Great Britain—History—Charles II, 1660-1685.”

MARC records: not available.

OPAC records: no.

Quality of scans: recognition. Page 271 was reproduced too close to the left margin, so the word “State” was read as “tate” by OCR.

Accuracy of keyword searches: Of the 6 index pages viewed (277-282), 4 had multiple lines that were not read by OCR. In addition, on p. 5, the last footnote was not read by OCR; the italicized word “Négotiations” in a footnote on p. 89 was read as “Negotiations” by OCR; and in the footnote on p. 217, the word “prejudiciall” [sic] was read by OCR as “prejudicial!”

204C) Bibliographic citation: Barbour, Violet. *Henry Bennet, Earl of Arlington, Secretary of State to Charles II*. Washington: American Historical Association, 1914

Project: ACLS Humanities E-Book.

Source of physical book: unknown..

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 5-24, 70-89, 137-156, 200-219, 263-282.

Physical condition: good.

Ease of finding: easy.

Metadata: Accurate, comprehensive, derived from MARC, with the exception of labeling of page numbers in the Page indicator (see below).

MARC records: not available.

OPAC records: yes.

Quality of scans: Poor (see below).

Accuracy of keyword searches: Labeling of page numbers in the page indicator was incorrect throughout the book. This led to problems with the retrieval of pages in which search terms occurred. Page 1 was labeled p. IX. Page 5 was labeled p. 1. Each subsequent page was labeled with a number 4 less than the number of the page. Word searches retrieved hits for pages in which the searched-for words occurred, indicating page numbers that were correct; however, clicking on links for these pages led to the wrong page—the number as listed in the page indicator. For example, a search for “prejudiciall” retrieved a hit for p. 217 (the correct page), but clicking on the link for 217 led to p. 221, where the word did not appear.

205A) Bibliographic citation: Brown, Louise Fargo. *The Political Activities of the Baptists and Fifth Monarchy Men in England during the Interregnum*. Washington: American Historical Association, 1913.

Project: Google Book Search.

Source of physical book: Library of the University of Michigan.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 60-79, 110-129, 160-179, 200-219.

Physical condition: A splotch on p. 18 interfered with one word. Underling appeared on one page (5), but OCR was not affected.

Ease of finding: easy.

Metadata: Subjects: “Baptists”; “Fifth Monarchy Men”; and “Church and state.” LC lists “Baptists—England—History—17th century”; “Fifth Monarchy Men”; and “Great Britain—Politics and government—1649-1660.”

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: OCR misreadings occurred with italicized words in footnotes on 2 pages (72, 176).

205B) Bibliographic citation: Brown, Louise Fargo. *The Political Activities of the Baptists and Fifth Monarchy Men in England during the Interregnum*. New York: Burt Franklin, 1911.

Project: Microsoft Live Search Books.

Source of physical book: Erindale College.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 60-79, 110-129, 160-179, 200-219.

Physical condition: good.

Ease of finding: easy.

Metadata: Pages listed: 280; last page is 258. Contents: the link to Chapter 1, p. 1, retrieved a page before the title page.

MARC records: not available.

OPAC records: no.

Quality of scans: poor. Missing pp. 202-203. Page 204-205 were labeled 202-203 then repeated.

Accuracy of keyword searches: OCR did not read lines of footnotes on 2 pages (64, 129) and lines of the Bibliography on 3 pages (215, 216, 219). In addition, light or broken letters in the footnotes on p. 14 and 72 resulted in OCR misreadings. (This also occurred with the ACLS digitization (205C) of these pages.)

205C) Bibliographic citation: Brown, Louise Fargo. *The Political Activities of the Baptists and Fifth Monarchy Men in England during the Interregnum*. New York: Burt Franklin, 1911.

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 60-79, 110-129, 160-179, 200-219.

Physical condition: good.

Ease of finding: easy.

Metadata: Accurate, comprehensive, derived from MARC, with the exception of labeling of page numbers in the Page indicator (see below).

MARC records: not available.

OPAC records: yes.

Quality of scans: Poor (see below).

Accuracy of keyword searches: Labeling of page numbers in the page indicator was incorrect throughout the book. This led to problems with the retrieval of pages in which search terms occurred. Page 1 was labeled “XI – Table of Contents.” Page 3 was labeled p. 1. Each subsequent page was labeled with a number 2 less than the number of the page. Word searches retrieved hits for pages in which the searched-for words occurred, indicating page numbers that were correct; however, clicking on links for these pages led to incorrect pages. For example, a search for “Inner Life” retrieved p. 3 (the correct page), but clicking on the link for 3 led to p. 5, where the word did not appear. OCR misreadings occurred on 3 pages (3, 14, 72). OCR read 558 items, including dates, beginning with “16” as “I6”; it read 410 items beginning with “16” correctly.

206A) Bibliographic citation: Droysen, Johann Gustav. *Outline of the Principles of History (Grundriss Der Historik)*. Translated by E. Benjamin Andrews. Boston: Ginn, 1893.

Project: Google Book Search.

Source of physical book: Harvard University.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: xvi-xxxv, 4-23, 34-53, 64-83, 94-113.

Physical condition: Markings appeared on 24 pages (xviii, xxi, xxiii-xxxiii, 4-6, 8-9, 11-12, 15, 17, 20, 22), with OCR significantly affected on 9 of these pages (xviii, xxiii-xxv, xxvii, xxxi, 5, 6, 11).

Ease of finding: easy.

Metadata: Subjects: “History.” LC lists “Buckle, Henry Thomas, 1821-1862. History of civilization in England” and “History.” Key words and phrases included “little x” (2 occurrences), “diadochi” (2 occurrences), and “definite series” (3 occurrences).

MARC records: not available.

OPAC records: yes.

Quality of scans: reference.

Accuracy of keyword searches: see “Physical condition” above.

206B) Bibliographic citation: Droysen, Johann Gustav. *Outline of the Principles of History*. Boston: Ginn, 1893.

Project: Microsoft Live Search Books.

Source of physical book: University of Toronto Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: xvi-xxxv, 4-23, 34-53, 64-83, 94-113.

Physical condition: Markings appeared on 65 pages, with OCR significantly affected on 6 pages (7, 9-11, 14, 95).

Ease of finding: easy.

Metadata: Pages listed: 166; LC lists “xxxv, 122 p.” Contents: Link in to p. 16 retrieved p. 10; link to p. 36 retrieved p. 30.

MARC records: not available.

OPAC records: no.

Quality of scans: reference.

Accuracy of keyword searches: In addition to problem noted above, OCR misreadings occurred on 3 pages (xxxi, 4, 10), and the footnote on p. 18 was not read by OCR.

206C) Bibliographic citation: Droysen, Johann Gustav. *Outline of the Principles of History*. Boston: Ginn, 1893.

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: xvi-xxxv, 4-23, 34-53, 64-83, 94-113.

Physical condition: Splotches appeared on 8 pages (34, 66, 75, 77, 95-98), with OCR affected on 4 of these pages (77, 95-97).

Ease of finding: easy.

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: see above.

207A) Bibliographic citation: Wriothesley, Charles. *A Chronicle of England during the Reigns of the Tudors, from A.D. 1485-1559*. Edited, from a ms. in the possession of Lieut.-General Lord Henry H. M. Percy, by William Douglas Hamilton. 2 vols. Westminster: Printed for the Camden Society, 1875-77. [Volume I]

Project: Google Book Search.

Source of physical book: Harvard College Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: ii-xxi, 34-53, 80-99, 150-169, 206-225.

Physical condition: good.

Ease of finding: A title search retrieved only vol. II. A click on the “More editions” link then retrieved vol. I.

Metadata: Key terms and phrases included Old English words (“daie,” “payd,” “shott”) and Latin words (“sciri” and “armiger”) that have no reason to be termed key terms.

Contents: the 14 “Sections” did not correspond to any divisions in the text; for example “Section 1” retrieved a page before the introduction that listed names of persons who were involved with the Camden Society, which published the book, while “Section 2” retrieved a page in the midst of the introduction. Subjects: “Great Britain”; Hollis Catalog lists “Great Britain—History—Tudors, 1485-1603.”

MARC records: not available.

OPAC records: no [Hollis Catalog includes a link only to vol. II.]

Quality of scans: reference.

Accuracy of keyword searches: At least one footnote on each of 3 pages was not searchable, although all were viewable in the plain text mode (xii, 37, 43). The name “Henry” was read by OCR as “Heury” 27 times; the name “Richard” as “Kichard” 2 times. In addition, OCR misreadings were noted on 2 pages (xiv, 47). Most pages contained headings in the margins that were read by OCR mixed in with the sentence that

appeared on the same line. For example, on p. 161, the first sentence of the first full paragraph reads in part, “clocke in the morninge” with the word “moringe” appearing on the next line. In the right margin at the end of these lines, the heading “St. Giles church burnt” appears. OCR read “clocke in the St. Giles morninge,” combining the text and the first line of the heading.

207B) Bibliographic citation: Wriothesley, Charles. *A Chronicle of England during the Reigns of the Tudors, from A.D. 1485-1559*. Edited, from a ms. in the possession of Lieut.-General Lord Henry H. M. Percy, by William Douglas Hamilton. 2 vols. Westminster: Printed for the Camden society, 1875-77. [Volume I]

Project: Microsoft Live Search Books.

Source of physical book: University of St. Michael’s College. John M. Kelly Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: i-xx, 35-54, 80-99, 150-169, 207-226.

Physical condition: good.

Ease of finding: easy.

Metadata: Pages listed: 290; last page is 226.

MARC records: not available.

OPAC records: no.

Quality of scans: reference.

Accuracy of keyword searches: Footnotes and lines of text were not read by OCR on 25 pages. On 15 of these pages (207-217, 219-223), no words at all were read by OCR presumably because all of the text was in Latin. Most pages contained headings in the margins that were read by OCR mixed in with the sentence that appeared on the same line. For example, on p. 161, the first sentence of the first full paragraph reads in part, “clocke in the morninge” with the word “moringe” appearing on the next line. In the right margin at the end of these lines, the heading “St. Giles church burnt” appears. OCR read “clocke in the St. Giles morninge,” combining the text and the first line of the heading. Misreadings by OCR of Old English words were not uncommon. In addition, the name “Richard” was read as “Kichard” six times.

207C) Bibliographic citation: Wriothesley, Charles. *A Chronicle of England during the Reigns of the Tudors, from A.D. 1485-1559*. Edited, from a ms. in the possession of Lieut.-General Lord Henry H. M. Percy, by William Douglas Hamilton. 2 vols. Westminster: Printed for the Camden society, 1875-77. [Volume I]

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: ii-xxi, 34-53, 80-99, 150-169, 206-225.

Physical condition: Markings appeared on 4 pages (82, 88, 162, 166), though OCR was affected on only 1 of these pages (166). A splotch on p. 36 obscuring two words resulted in OCR misreadings of these words.

Ease of finding: easy.

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: On each of 3 pages, a footnote indicator was read by OCR as part of words (ix, 43, 46). Most pages contained headings in the margins that were read by OCR mixed in with the sentence that appeared on the same line. For example, on p. 161, the first sentence of the first full paragraph reads in part, “clocke in the morninge” with the word “moringe” appearing on the next line. In the right margin at the end of these lines, the heading “St. Giles churche burnt” appears. OCR read “the clocke in the St. Giles chnrche burnt. moringe,” combining the text and the two lines of the heading.

208A) Bibliographic citation: Wriothesley, Charles. *A Chronicle of England during the Reigns of the Tudors, from A.D. 1485-1559*. Edited, from a ms. in the possession of Lieut.-General Lord Henry H. M. Percy, by William Douglas Hamilton. 2 vols. Westminster: Printed for the Camden society, 1875-77. [Volume II]

Project: Google Book Search.

Source of physical book: Harvard University Library.

Provenance: Bought with the income from the bequest of Thomas Hollis.

Pages viewable: full text.

Pages viewed: 1-20, 40-59, 85-104, 120-139, 151-170.

Physical condition: good.

Ease of finding: easy.

Metadata: Key terms and phrases included Old English words (“streates,” “daie,” “payd,” “sayd”) and Latin words (“sciri” and “armiger”) that have no reason to be termed key terms. Contents: the 4 “Sections” did not correspond to any divisions in the text.

Subjects: “Great Britain”; Hollis Catalog lists “Great Britain—History—Tudors, 1485-1603.”

MARC records: not available.

OPAC records: yes.

Quality of scans: reference.

Accuracy of keyword searches: Most pages contained headings in the margins that were read by OCR mixed in with the sentence that appeared on the same line. (See 207A-C).

The name “Henry” was read by OCR as “Heury” 30 times; the word “King” as “Kiug” 3 times.

208B) Bibliographic citation: Wriothesley, Charles. *A Chronicle of England during the Reigns of the Tudors, from A.D. 1485-1559*. Edited, from a ms. in the possession of Lieut.-General Lord Henry H. M. Percy, by William Douglas Hamilton. 2 vols. Westminster: Printed for the Camden society, 1875-77. [Volume II]

Project: Microsoft Live Search Books.

Source of physical book: University of St. Michael’s College.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 1-20, 40-59, 85-104, 120-139, 151-170.

Physical condition: good.

Ease of finding: easy.

Metadata: Pages listed: 184; last page is 170.

MARC records: not available.

OPAC records: no.

Quality of scans: reference. The left side of 2 pages was cut off so that the first words of headings in the margins were obscured.

Accuracy of keyword searches: Most pages contained headings in the margins that were read by OCR mixed in with the sentence that appeared on the same line. (See 207A-C).

208C) Bibliographic citation: Wriothesley, Charles. *A Chronicle of England during the Reigns of the Tudors, from A.D. 1485-1559*. Edited, from a ms. in the possession of Lieut.-General Lord Henry H. M. Percy, by William Douglas Hamilton. 2 vols. Westminster: Printed for the Camden society, 1875-77. [Volume II]

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 1-20, 40-59, 85-104, 120-139, 151-170.

Physical condition: Markings occurred on one page (90).

Ease of finding: easy

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: On each of 2 pages, a footnote indicator was read by OCR as part of words (126, 134). An OCR misreading occurred on p. 3 (“Whitehall” read as “Whitchall.”) Most pages contained headings in the margins that were read by OCR mixed in with the sentence that appeared on the same line. (See 207A-C).

209A) Bibliographic citation: Cowan, James. *The New Zealand Wars: A History of the Maori Campaigns and the Pioneering Period*. 2 vols. Wellington, New Zealand: W. A. G. Skinner, R.E. Owen, government printer, 1922-23. [Volume I]

Project: Microsoft Live Search Books.

Source of physical book: University of Toronto Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 6-25, 100-119, 226-245, 337-356, 435-454.

Physical condition: Beginning at p. 233 and continuing until the end of the book (p. 459), a vertical crease near the left margin traverses each odd-numbered page. OCR was affected on many of these pages. On 233, for example, the word “Maori” was read by OCR as “M lori”; “original” as “oiginal.” On 235, “authority” was read as “uthority.”

Ease of finding: easy.

Metadata: Pages listed: 488; last page is 459. Subjects: “New Zealand History”; LC lists “Maori (New Zealand people)” and “New Zealand—History.” Contents: the link to Chapter I, p. 1 retrieved a page prior to the title page.

MARC records: not available.

OPAC records: no.

Quality of scans: poor. In addition to the problem noted above, p. 340 was blurry, but readable, though no word on the page was word searchable. Page 341 was blurry, but readable, and many words on the page were not retrievable through word searches, including all text in the first 9 lines.

Accuracy of keyword searches: In addition to the problems noted above, on index p. 454, a number of lines were not read by OCR. OCR misreadings were found on 2 additional pages (13, 108).

209B) Bibliographic citation: Cowan, James. *The New Zealand Wars: A History of the Maori Campaigns and the Pioneering Period*. 2 vols. Wellington, New Zealand: W. A. G. Skinner, R.E. Owen, government printer, 1922-23. [Volume I] Note: This was a different edition than that digitized by Microsoft in 209A. Page 13 in the Microsoft version appears on p. 14 in the ACLS version; p. 108 as p. 111; p. 454 as 467.

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 6-25, 100-119, 226-245, 337-356, 435-454.

Physical condition: 3 pages were smudgy enough for OCR to be affected (111, 235, 355).

Ease of finding: easy

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: Other than the problem noted above, words searches retrieved perfectly accurate results.

210A) Bibliographic citation: Edward VI, King of England. *Literary Remains of King Edward the Sixth*. Ed. from his autograph manuscripts, with historical notes, and a biographical memoir by John Gough Nichols. 2 vols. London: J. B. Nichols and sons, 1857. [Vol. I]

Project: Google Book Search.

Source of physical book: Oxford University Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: xcv-cxiii, ccl-cclxix, cccxxxii-ccli, 50-69, 170-189.

Physical condition: good.

Ease of finding: easy.

Metadata: Subjects: "Great Britain." LC lists "Edward VI, King of England, 1537-1553" and "Great Britain—History—Edward VI, 1547-1553." Key words and phrases included "maue" (an OCR misreading, occurring on 19 pages, of "Ma^{tie}", an abbreviation for "Majestie," as in "the Kinges Ma^{tie}"); the Latin words "tuam" and "nihil"; and the Old English word "dyvers."

MARC records: not available.

OPAC records: no.

Quality of scans: poor. **Page indicator:** The first 310 pages of the volume are paginated in roman numerals. Google's page indicator listings were incorrect for approximately 87 of these pages. For one run of pages beginning with xxxvi, Google's page indicator listings began with page no. 1538, presumably because the date "A.D. 1538" appears at the right corner of the page, even though the page number "xxxvi" is at the left corner.

Accuracy of keyword searches: Many pages contained headings in the margins. On some pages, these were read by OCR mixed in with the sentence that appeared on the same line (p. cii); on other pages, headings did not interfere (p. cclii). On 3 pages, searches for words in footnotes or near the bottom of the page did not retrieve results, even though the words appeared in the "plain text" version of the pages (pp. cclii, cccxl, 56).

210B) Bibliographic citation: Edward VI, King of England. *Literary Remains of King Edward the Sixth*. Ed. from his autograph manuscripts, with historical notes, and a biographical memoir by John Gough Nichols. 2 vols. London: J. B. Nichols and sons, 1857. [Vol. I]

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: xcv-cxiii, ccl-cclxix, cccxxxii-cccli, 50-69, 170-189.

Physical condition: good.

Ease of finding: easy

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: Many pages contained headings in the margins. On these pages the headings were read by OCR mixed in with the sentence that appeared on the same line (pp. cii, cclii). On one page, a name in a footnote was misread by OCR (175). On one page, lettering in a Gothic font resulted in words misread by OCR (179).

211A) Bibliographic citation: Forbes, Anna. *Insulinde: Experiences of a Naturalist's Wife in the Eastern Archipelago*. Edinburgh: W. Blackwood and sons, 1887.

Project: Google Book Search.

Source of physical book: New York Public Libraries. Research Libraries.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 72-91, 142-161, 212-231, 282-301.

Physical condition: good.

Ease of finding: easy.

Metadata: **Subjects:** "Malay Archipelago" and "Indonesia"; LC lists "Malay Archipelago—Description and travel." **Key words and phrases** included "Fcap," "Crown 8vo," and "LL.D." (terms used repeatedly in the publisher's list of books included at the end of the volume); "verandah" (part of the house the author mentions at times, but

certainly not a “key term”); “trassi” (2 occurrences—a name for mashed fish used in stews; again, not a “key term”).

MARC records: not available.

OPAC records: no.

Quality of scans: poor. 9 pages (15, 144, 153, 160, 215, 218, 221, 231, 288) were not read by OCR, so word searches for those pages were not available.

Accuracy of keyword searches: see above.

211B) Bibliographic citation: Forbes, Anna. *Insulinde: Experiences of a Naturalist's Wife in the Eastern Archipelago*. Edinburgh: W. Blackwood and sons, 1887.

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 72-91, 142-161, 212-231, 282-301.

Physical condition: good.

Ease of finding: easy

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: archival.

Accuracy of keyword searches: All searches retrieved accurate results.

212A) Bibliographic citation: Sabbadini, Remigio. *Storia del ciceronianismo e di altre questioni letterarie nell' età della rinascenza*. Torino: Ermanno Loescher, 1885.

Project: Google Book Search.

Source of physical book: Oxford University.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 30-49, 58-77, 86-105, 114-133.

Physical condition: good.

Ease of finding: easy.

Metadata: minimal.

MARC records: not available.

OPAC records: yes.

Quality of scans: recognition.

Accuracy of keyword searches: OCR misreadings of italicized words occurred on 3 pages (2, 4, 117). Word searches for last footnote on the page were not accurate on 7 pages (10, 37, 39, 58, 93, 95, 115), even though the footnote appeared in the “plain text” mode.

212B) Bibliographic citation: Sabbadini, Remigio. *Storia del ciceronianismo e di altre questioni letterarie nell' età della rinascenza*. Torino: Ermanno Loescher, 1885.

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 2-21, 30-49, 58-77, 86-105, 114-133.

Physical condition: 14 pages were light with broken letters that resulted in OCR misreadings (2, 5, 10, 11, 13, 14, 17, 19, 32, 33, 36, 77, 123, 132).

Ease of finding: easy

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: reference.

Accuracy of keyword searches: see above.

213A) Bibliographic citation: Wallace, Malcolm William. *The Life of Sir Philip Sidney*. Cambridge: University Press, 1915.

Project: Microsoft Live Search Books.

Source of physical book: University of Toronto Library.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 1-20, 88-107, 220-239, 303-322, 380-399.

Physical condition: Minor markings occurred on p. 238, but these did not affect OCR.

Ease of finding: easy.

Metadata: Pages listed: 446; LC lists "vii, 428 p."

MARC records: not available.

OPAC records: no.

Quality of scans: reference. Ghosting of letters occurred on 2 pages (101, 105), but all words were readable, and OCR was not affected.

Accuracy of keyword searches: On 18 pages (2, 5-7, 10, 13, 17, 19, 88, 97, 222, 224, 233, 304, 308, 381, 387, 397), no words from one or more footnotes were retrieved through word searches. On one page (390), no words from 14 lines of text in French were retrieved through word searches. OCR misreadings of italicized words in footnotes occurred on 2 pages (16, 318).

213B) Bibliographic citation: Wallace, Malcolm William. *The Life of Sir Philip Sidney*. Cambridge: University Press, 1915.

Project: ACLS Humanities E-Book.

Source of physical book: unknown.

Provenance: unknown.

Pages viewable: full text.

Pages viewed: 1-20, 88-107, 220-239, 303-322, 380-399.

Physical condition: good.

Ease of finding: easy

Metadata: Accurate, comprehensive, derived from MARC.

MARC records: not available.

OPAC records: yes.

Quality of scans: reference.

Accuracy of keyword searches: Because footnote indicators in the text occurred before periods at the end of sentences rather than after, OCR read the footnote indicator as part

of the word, so that a search for the word alone did not retrieve the page desired. For example, on p. 221, the footnote indicator “2” appeared after the word “Vienna.” A search for “Vienna” did not retrieve p. 221, while searches for “Vienna2” and “Vienna*” did retrieve the page. This problem occurred on every page checked, with the exception, on occasion, when the footnote indicator “1” was read as an apostrophe (232, 319). This problem did not occur with the Microsoft digitization of the book (213A). In addition, OCR misread an italicized word in a footnote on p. 2 and an accented word in the text on p. 390.