CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/20/08

http://en.wikipedia.org/wiki/Diego_de_Torres_y_Villaroel

Diego de Torres Villarroel (1693 - 19 June 1770) was a Spanish writer, poet, dramatist, doctor, mathematician, priest and professor of the University of Salamanca. His most famous work is his autobiography, Vida, ascendencia, nacimiento, crianza y aventuras del Doctor Don Diego de Torres Villarroel (first published 1743).

Book title: Vida, Ascendencia, Nacimiento, Crianza Y Aventuras Del Dr. D. Diego de Villarroel.
Author: Torres de Villaroel, Diego
Year of Publication: 1792
Place of Publication: Madrid
Publisher: Imprenta de Gonzalez
ISBN number?: OCLC 43064011
Content/Disciplinary area: Spain – Literature; 18th Century Spanish Literature

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Below average. PDF. The first several pages are missing! Unclear if they were torn out or if they were simply not scanned. Metadata does not mention the missing pages.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata available on Google Books missing place of publication. Surprising that it is the only edition available when the version scanned is missing the first few pages.
MARC record available?	Yes. Through WorldCat..
Can the book be searched through an OPAC? If so, which one?	Yes WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	The book has section divisions but no table of table of contents; the metadata does not reflect the section divisions but instead randomly generates "chapters" based on certain keywords.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
Inner Cover	Poor.	Fingers	N/A	N/A
Title page	Above average	Large black mark – does not affect legibility.	N/A	N/A
1-3	Not available	Pages MISSING	N/A	N/A
4	Poor	Large black mark – does affect legibility.	N/A	N/A
10, 20	Above average	N/A	N/A	N/A
30, 40	Above average	N/A	N/A	N/A
50, 60	Above average	N/A	N/A	N/A
70, 80	Average.	Page 80 blurry – unclear	N/A	N/A

		if it's a printing error (ink running through back of page) or scanning error – legibility slightly affected.		
90, 100	Average.	90 and 100 slightly blurry – legibility not affected.	N/A	N/A
110, 120	Average.	110 blurry – legibility not affected.	N/A	N/A
130, 140	Average.	130 blurry – legibility not affected.	N/A	N/A
150, 160	Above average.	N/A	N/A	N/A
170, 180	Above average	N/A	N/A	N/A
190, 200	Below average.	200 blurry. Legibility affected.	N/A	N/A
210, 220	Average.	210 blurry. Legibility not affected.	N/A	N/A
230, 240	Below average.	240 blurry – legibility affected.	Page edge visible.	N/A
250, 260	Below average.	250 blurry – legibility affected.	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N – **ONLINE version only, not PDF.**
Villaroel	Poor. Found one instance of the author's last name and remaining results not indicated but still highlighted.
Madrid	Excellent. (Note – there was a technical error and it took me a couple of times to have the search mechanism highlight the result on the title page. The highlighted word was instead from the previous search)
Vida	Excellent.
MDCCXCII	Excellent.
él (w accent)	Excellent (Note: "él" is a different word from "el" and search worked great)
el (w/o accent)	Excellent.
año	Excellent.
ñ	No results. Individual letters cannot be searched unless they are words themselves. "ñ" is not a word.
á	Excellent.
Torres	Excellent.
Introducción	Excellent.
Lucas	Poor. **Blurred pages don't produce results.**

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	The covers and inner flaps are in color – not the remainder of the book.
Is the image scan an OCR or a PDF? If	PDF with no recognition available.

OCR, how successful is the recognition?	
Is the scan on a gray-scale or color scale?	In the PDF the covers and inner flaps are in color – not the remainder of the book.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Inner flap marbling could be useful to scholars.
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	None.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Below average.
Ease of access to full text	Easy but first few pages missing so I still had to find another digitized edition. Only one available from the Instituto Cervantes and it is a text only at http://www.intratext.com/IXT/ESL0030/_P1.HTM (not a scan).
Ability to copy, quote, cite?	None.
Ability to search?	Yes but online version only – not pdf.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Online version yes. PDF no.
How successful is OCR recognition?	Poor. OCR works fine for legible pages but not for numerous non-legible pages (which is basically the 2nd half of the book).
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good	No.

enough that this book could be reproduced into analog form?	
Issues with this book that would affect usefulness for scholars	Quality may be poor of the original text scanned (examples: pages missing, blurred pages).
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	Pages missing!
Other notes	Discrepancies between pdf and online version are all too apparent in this edition.

Book URL:
http://books.google.com/books?id=Yq8GAAAAQAAJ&printsec=frontcover&dq=Villarroel&as_brr=1&ei=ogaDSKOfCae6jgH688y3DQ
Date accessed: 7/20/008

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/13/08

http://www.amazon.com/Iracema-Classicos-da-Literatura-Brasileira/dp/0850515009

Although a recognized classic of Brazilian literature, Iracema has received scant attention in the English-speaking world. First published in 1865, the second in a cycle of three Indian-themed novels authored by Alencar, this passionate tale of love and loss was deliberately designed to reflect and celebrate the diversity of the Brazilian people as they struggled to forge an independent national identity during the nineteenth century. Set in the sixteenth century, the mesmerizing narrative recounts the twists and turns of the ill-fated love affair between Iracema, a beautiful Tabajara Indian maiden, and Martim, a compassionate Portuguese soldier. Although their romance necessarily ends in tragedy with the death of Iracema, the birth of their son symbolizes the genesis of the new Brazilian nation. Another historically and culturally significant addition to the Library of Latin America series. Margaret Flanagan --This text refers to an out of print or unavailable edition of this title.

Book title: Iracéma: the honey-lips, a legend of Brazil
Author: Alencar, José Martiniano de, 1829-1877
Year of Publication: 1886
Place of Publication: London
Publisher: Bickers
ISBN number?: OCLC 3450736
Content/Disciplinary area: Romance; South America; Victorian Literature

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes through WorldCat
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata very poor. I actually discovered that there was another text (different author, same publisher) bound together with this one that is not mentioned in OAT. I am just going to look at the first, Iracema.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) The PDF is corrupt and numerous pages came up blank – my analysis here is just of the online version.

Page(s):	readability/ quality of scan	anomalies?	other?	other?
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	Ink inscription – does not affect legibility.	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100	Excellent	N/A	N/A	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N
Iracema	Average. The spelling for the title is old Portuguese orthography – modern day spelling is Iracema for what this edition has as Iraçéma. Iracema is partially searchable while Iraçéma cannot be searched.
á, é, í, ó, ú	Unreadable.
London	Excellent.
Alencar	Excellent.
III	Excellent.
75	Unreadable.
Preface	Excellent.
Chapter	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Online version. Average recognition.
Is the scan on a gray-scale or color scale?	Color.
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming great.
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Not available. PDF is corrupt.
Ability to search?	Yes.
Ability to download?	Yes but pdf is corrupt.
Ability to print?	Yes but pdf is corrupt.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Can't tell because the pdf is corrupt. If they fixed the file, then yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Can't tell because the pdf is corrupt. If they fixed the file, then yes.
Issues with this book that would affect usefulness for scholars	Can only look at it online can't download the whole thing.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Can't tell because the pdf is corrupt. If they fixed the file, then yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/iracmahoneylip00alenuoft
Date accessed: 7/13/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/7/08

Book title: The Golden Ass of Apuleius
Author: Apuleius, Madaurensis; Adlington, William.
Year of Publication: 1893
Place of Publication: London
Publisher: David Nutt
ISBN number?: OCLC 1984559
Content/Disciplinary area: Latin Novel – 2nd century.

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata adequate.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	No table of contents included in metadata.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	Other?	Other?
i, x, xx, xxx	Excellent	N/A	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A
200, 210	Excellent	N/A	N/A	N/A
220, 230	Excellent	N/A	N/A	N/A
240	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
(Is OCR search permitted) **Y** / N	

Apuleius	Excellent
Ass	Excellent
Latin	Excellent
222	Excellent
XLVI	Excellent
Chapter	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition
Is the scan on a gray-scale or color scale?	Color scale
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming very nice.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	No table of contents included in metadata.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	No table of contents included in metadata.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/goldenassofapule00apuliala
Date accessed: 9/7/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/03/08

http://en.wikipedia.org/wiki/The_Book_of_Good_Love

The Book of Good Love (*El Libro de Buen Amor*), considered to be one of the masterpieces of Spanish poetry, is a semi-biographical account of romantic adventures by Juan Ruiz, the Archpriest of Hita, dating from 1330.

The work is considered as the best piece in the medieval genre known as Mester de Clerecía.

The Book begins with prayers and a guide as to how to read the work, followed by stories each containing a moral and often comical tale.

Book title: Libro de Buen Amor (vol. 1)
Author: Ruiz, Juan, Arcipreste de Hita [Editor: Julio Cejador y Frauca (1864-1927)]
Year of Publication: 1931
Place of Publication: Madrid (Spain)
Publisher: Espasa-Calpa, S.A.
ISBN number?: OCLC 40851780
Content/Disciplinary area: Spanish poetry -- To 1500

AVAILABILITY SEARCH

Sources of scanned book:	1. Project Gutenberg
Comments on condition/provenance of book:	U/A because it is a searchable website online.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata pretty good and the online version has both volumes available.
MARC record available?	Yes. Available through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Excellent

SCAN SEARCH

340	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	I am not going to do any word searches with this book because it is not a book search but a websearch at this point.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	N/A
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Average search capabilities. Cannot search for accented words.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	N/A
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes because it is in analog form
Issues with this book that would affect usefulness for scholars	Not a book but instead a webpage.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No
Other notes	I have now run into this issue twice and as a result, have now learned the difference between a book (perhaps typed out?) - versus a scanned book. I also checked the same book at the Instituto Cervantes (http://www.cervantesvirtual.com/servlet/SirveObras/24661685545133385754491/p0000001.htm#74) with similar problems – it is a plain text version and not a scanned book.

Book URL: http://www.gutenberg.org/files/16625/16625-h/16625-h.htm
Date accessed: 8/03/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/15/08

http://en.wikipedia.org/wiki/Auto_de_fe

The phrase auto de fe refers to the ritual of public penance of condemned heretics and apostates that took place when the Spanish Inquisition or the Portuguese Inquisition had decided their punishment (that is, after the trial). Auto de fé in medieval Spanish means "act of faith". The phrase also commonly occurs in English in its Portuguese form auto-da-fé.

In the popular imagination "auto-da-fé" has come to refer to burning at the stake for heresy.

Title: The Inquisition and Judaism. A sermon addressed to Jewish martyrs, on the occasion of an Auto da fe at Lisbon, 1705 by the Archbishop of Cranganor. Also a reply to the sermon by Carlos Vero.
Author: Diego Annunciazaro Justiniano, abp. of Cranganor; Carlos Vero (trans. Moses Mocatta)
Year of Publication: 1845
Place of Publication: London
Publisher: J. Wertheimer and Co.
ISBN number?: OCLC 8601644
Content/Disciplinary area: Jews – Persecutions; Inquisition – Portugal

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Average. No information on the table of contents provided.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)	All of the pages are marked up by someone's notes, underlining, bracketing, etc. Why then choose this edition to scan?

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Excellent	N/A	N/A	N/A
x	Excellent	pencil markings – legibility unaffected	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A

140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N (Search permitted in online version, but not pdf)
Inquisition	Excellent
1845	Excellent
Fé	Unreadable (only unaccented word "fe" is searchable)
Introduction	Above average.
Cranganor	Average – on upper margin only odd numbered pages return results – not even numbered.
Jehovah	Excellent
XVIII	Excellent
Israel	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Above average recognition
Is the scan on a gray-scale or color scale?	Color
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming excellent.
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent
Ease of access to full text	Easy
Ability to copy, quote, cite?	Yes
Ability to search?	Yes
Ability to download?	Yes
Ability to print?	Yes
Scanned book allow for consulting of contents?	Yes
How successful is OCR recognition?	Above average
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes
Issues with this book that would affect usefulness for scholars	Search recognition above average
Can this book be shared with other scholars/students?	Yes
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/inquisitionjudai00annurich
Date accessed: 8/15/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/7/08

Book title: The Poems and Prose Poems of Charles Baudelaire
Author: Charles Baudelaire; James Huneker
Year of Publication: 1919
Place of Publication: New York
Publisher: Bretano
ISBN number?: OCLC 1889177
Content/Disciplinary area: French poetry – 19th century

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Above average.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	It was a little challenging finding a collection of poetry that are otherwise not selections of poems in an anthology. This text seems pretty complete however.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Table of contents very good.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability/ quality of scan**	**anomalies?**	**Other?**	**Other?**
i, xx	Above average	N/A	N/A	N/A
xxx. xl, lv	Above average	N/A	N/A	N/A
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	**Comment**
	(Is OCR search permitted) **Y** / N Online version only, not pdf.
Baudelaire	Excellent
flowers	Excellent
Huneker	Excellent
poetry	Excellent
Paris	Excellent
1919	Excellent

illume	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition
Is the scan on a gray-scale or color scale?	Gray-scale
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming good.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	A little bit challenging due to difficulty in finding poetry collections online.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Yes – online version only not pdf.
Ability to download?	Yes.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	N/A
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://books.google.com/books?id=jeLIjqPGt64C&dq=baudelaire&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 9/7/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/21/08

http://en.wikipedia.org/wiki/Gustavo_Adolfo_B%C3%A9cquer

Gustavo Adolfo Domínguez Bastida, better known as Gustavo Adolfo Bécquer, (Seville February 17, 1836 – Madrid December 22, 1870) was a Spanish post-romanticist writer of poetry and short stories, now considered one of the most important figures in Spanish literature. He adopted the alias of Bécquer as his brother Valeriano Bécquer, a painter, had done earlier. He was associated with the post-romanticism movement and wrote while realism was enjoying success in Spain. He was moderately well known during his life, but it was after his death that most of his works were published.

His best known works are the Rhymes and the Legends, usually published together as Rimas y leyendas. These poems and tales are essential to the study of Spanish literature and common reading for high-school students in Spanish-speaking countries.

His work approached the traditional poetry and themes in a modern way, and is considered the founder of modern Spanish lyricism. Bécquer's influence on 20th century poets of the Spanish language can be felt in the works of Octavio Paz and Giannina Braschi.

Book title: Legends, Tales and Poems
Author: Gustavo Adolfo Bécquer (trans. Everett Ward Olmsted)
Year of Publication: 1907
Place of Publication: Boston, New York
Publisher: Ginn & company
ISBN number?: 1625345
Content/Disciplinary area: Spanish Romanticism; 19th c. Spanish Literature

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Good. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata very good with an accurate table of contents and links to other editions in English and Spanish. In the online version, the table of contents in the text links to its relevant sections.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)	There is an occasional small black mark in the corner of the pages – legibility in no way affected.

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Above average	N/A	N/A	N/A
x, xx	Above average	N/A	N/A	N/A
xxx, xl	Above average	N/A	N/A	N/A

l, lx	Above average	N/A	N/A	N/A
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A
280	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N
	In online version only, not pdf
Becquer	Average. Searchable in text but not in upper margins.
Introduction	Excellent.
año	Excellent.
corza blanca	Excellent. Finds words as a phrase.
xxx	Poor. Page numbers not searchable.
36	Poor. Page numbers not searchable.
Vocabulary	Average. Searchable in text but not in upper margins.
acabar	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes. Just one illustration of author.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Only online version searchable not the pdf. Recognition average.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Probably not.

Are the illustrations searchable?	No.
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	No.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Online version only, not pdf.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	OCR average recognition.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	If OCR is fixed then yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A

Other notes N/A

Book URL:
http://books.google.com/books?id=YIYBAAAAYAAJ&dq=Gustavo+Adolfo+B%C3%A9cquer&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 7/21/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/7/08

http://en.wikipedia.org/wiki/Decameron

The Decameron (subtitle: Prencipe Galeotto) is a collection of 100 novellas by Italian author Giovanni Boccaccio, probably begun in 1350 and finished in 1353. It is a medieval allegorical work best known for its bawdy tales of love, appearing in all its possibilities from the erotic to the tragic. Many notable writers such as Chaucer are said to have drawn inspiration from The Decameron (See Literary sources and influence of the Decameron below).

Book title: Decameron
Author: Giovanni Boccaccio; J. M. Rigg
Year of Publication: 1906
Place of Publication: London
Publisher: Henry F. Bumpus
ISBN number?: OCLC 4570999
Content/Disciplinary area: Italian novella – fourteenth century.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata above average but doesn't readily point to the other volume in set.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata includes table of contents but not labeled by novella, rather by "section". PDF sections link to relevant portions.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	Other?	Other?
i, x, xx, xxx	Above average	N/A	N/A	N/A
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	pg 130 discolored – legibility not affected	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
193-208	**MISSING!**	N/A	N/A	N/A

200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A
280, 290	Above average	N/A	N/A	N/A
300, 310	Above average	N/A	N/A	N/A
320, 330	Above average	N/A	N/A	N/A
340, 350	Above average	N/A	N/A	N/A
360, 370	Above average	N/A	N/A	N/A
380, 390	Above average	N/A	N/A	N/A
400	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word		Comment
	(Is OCR search permitted) **Y** / N Online version only, not pdf.	
Decameron	Excellent	
Pietro	Excellent	
Cyprus	Excellent	
222	Unreadable	
story	Excellent	
Tarisi	Excellent.	
Boccaccio	Excellent	
novel ii	only reads each term separately	

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition
Is the scan on a gray-scale or color scale?	Gray-scale
Quality – does it look like the real thing?	No
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	Yes!
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes

Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming good for text and images.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Below average – missing pages!
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Yes – online version only not pdf.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	Missing pages makes a difference.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	Metadata doesn't readily point to other volume; plus there is a big chunk of pages missing.
Other notes	N/A

Book URL: http://books.google.com/books?id=vKvVpIpr43EC&dq=decameron&source=gbs_summary_s&cad=0
Date accessed: 9/7/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/6/08

http://en.wikipedia.org/wiki/Alvar_Nunez

Álvar Núñez Cabeza de Vaca (Jerez de la Frontera, Cádiz, ca. 1490/1507 – Sevilla, ca. 1557/1559) was an early Spanish explorer of the New World and is remembered as a protoanthropological author.

Book title: Naufragios y Comentarios
Year of Publication: 1922
Place of Publication: Madrid
Publisher: Calpe
ISBN number?: OCLC 738914
Content/Disciplinary area: Narváez, Pánfilo de, d. 1528; Ribera, Hernando de, 16th cent; Indians of North America; America -- Discovery and exploration Spanish; America -- Early accounts to 1600; Argentina -- History 1535-1617; Paraguay -- History To 1811

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata adequate.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	No table of contents included in metadata.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	Other?	Other?
1, 10	Above average	Pencil markings – legibility not affected	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Above average	pg 50 - pencil markings	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Above average	pg 90 – pencil markings	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	pg 120 – pencil markings	N/A	N/A
140, 150	Excellent	pg 140 – pencil markings	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A
200, 210	Excellent	N/A	N/A	N/A
220, 230	Excellent	N/A	N/A	N/A
240, 250	Excellent	N/A	N/A	N/A

260, 270	Excellent	N/A	N/A	N/A
280, 290	Excellent	N/A	N/A	N/A
300, 310	Excellent	N/A	N/A	N/A
320, 330	Excellent	N/A	N/A	N/A
340, 350	Excellent	N/A	N/A	N/A
360	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	(Is OCR search permitted) **Y** / N	Comment
Naufragios	Excellent	
Núñez	Excellent	
tunas	Excellent	
pequeña	Excellent	
333	Excellent	
indios	Excellent.	
Bernal	Excellent	
desembarcó	Excellent	
XXXVIII	Excellent	

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Cover and two maps included.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition
Is the scan on a gray-scale or color scale?	Color scale
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes
Are the illustrations searchable?	No.
Ability to print illustrations? (is there a watermark?)	Yes
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming very nice.

Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	No table of contents included in metadata.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	No table of contents included in metadata.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/naufragioscomen00nbudrich
Date accessed: 9/6/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/28/08

[Today I searched for editions of Pero Vaz de Caminha's letter to the King of Portugal announcing the discovery of Brazil. Latin Americanists always love to compare his letter with that of Columbus – written a mere 8 years earlier.

I am going to do a bit of comparing between the Open-Access Text version of the book with the Google edition posted because they are the same exact edition with some notable differences.]

Book title: *Corografia brazilica, ou, Relação historico-geografica do reino do Brazil* (1817) [originally written in 1500]
Author: Caminha, Pero Vaz de, 15th cent. and Manuel Ayres de Casal.
Year of Publication: 1817 [original date of creation in Portuguese is 1500]
Place of Publication: Rio de Janeiro
Publisher: Impressão regia
ISBN number?: No [OCLC number: 216844038]
Content/Disciplinary area: Brazil -- History -- 1500-1548. | America -- Early accounts to 1600.

AVAILABILITY SEARCH

Source of scanned book:	Open-Access Text Archive.
Comments on condition/provenance of book:	Excellent. [Author: Ayres de Casal, Manuel, b. 1754? Digitizing Sponsor: MSN Usage Rights: See Terms Book Contributor: University of California Libraries Language: Portuguese Keywords: Brazil -- Description and travel]
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Fairly easy. I didn't realize that there were multiple volumes of this book until viewing the item entry in WorldCat. OAT has both volumes and GBS only has the first.
MARC record available?	Yes but very hard to find. I went to WorldCat to see what libraries have holdings of this same edition in order to locate the MARC record. Difficult because not all are accessible to outside users. [http://www.worldcat.org/wcpa/oclc/216844038?page=frame&url=http%3A%2F%2Fcatalog.umd.edu%2FF%3Ffunc%3Dfindb%26local_base%3Dgl%26find_code%3D035%26request%3DOCM00787293&title=University+of+Maryland%2C+College+Park&linktype=opac&detail=UMC%3AUniversity+of+Maryland%2C+College+Park%3AAcademic]
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	The metadata is not very good in the OAT version. It does not link to any information inside or outside the text. It also does not explain that the original was written in 1500. The GBS version does at least reference the original date of creation.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to	The metadata is not very good in the OAT version. It does not link to any information inside or outside the text. It also does not explain that the original was written in 1500. The GBS version does at least reference the original date of creation.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) Page(s):	**readability / quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Above average	I believe an inscription in the upper-left hand corner has been partially cut off by the scanner. Note: the GBS version has a lovely set of fingers included here on several pages.	The same goes for a stamp on one of the blank pages before the title page (?)	N/A
ii, 1	Excellent	N/A	N/A	N/A
10, 20	Excellent	On page 9 there is a fully legible signature in the upper-level hand corner. Very nice. Portions of page 20 are a little bit difficult to read (printing press? age?) and are even more difficult with the gray-scale GBS scan.	N/A	N/A
30, 40	Excellent	N/A	N/A	N/A
50, 60	Excellent	N/A	N/A	N/A
70, 80	Excellent	N/A	N/A	N/A
88, 89	Excellent	N/A	N/A	N/A
90, 100	Excellent	N/A	N/A	N/A
110, 120	Excellent	N/A	N/A	N/A
130, 140	Above average	Page 140 is very faint (ink? Printer? Age?) and hard to read. GBS version is better.	N/A	N/A
150, 160	Excellent	N/A	N/A	N/A
170, 180	Excellent	N/A	N/A	N/A
190, 200	Excellent	N/A	N/A	N/A
210, 220	Excellent	N/A	N/A	N/A
230, 240	Excellent	N/A	N/A	N/A
250, 260	Excellent	N/A	N/A	N/A
270, 280	Excellent	N/A	N/A	N/A
290, 300	Excellent	N/A	N/A	N/A
310, 320	Excellent	N/A	N/A	N/A
330, 340	Excellent	N/A	N/A	N/A
350, 360	Excellent	N/A	N/A	N/A
370, 380	Excellent	N/A	N/A	N/A
390, 400	Excellent	N/A	N/A	N/A
410, 420	Excellent	N/A	N/A	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) **Y** / N** (Important note – please see Appendix II)

Word	Comments
Relação	Cannot be searched with accents, despite the fact that it is part of the title of the book. Non-accented version is searchable (relacao) in both online version and pdf. Word is not a valid search query with GBS.
23	The online version has completely nonsensical responses, the first being "Brazilica", also from the title. A search of the pdf works just fine.
Caminha	Excellent.
Tom. (Volume)	Works great, even including the period.
Introducção	Can't find it. Have to type in "introd" in order to get results. "Introduccao" (without accents) doesn't work either.
( )	Parentheses can be searched – important because all of the footnotes are signaled with parentheses.
Brasil	Excellent.
Brazil	The pdf recognizes the word but not the online version.
Vaz	Excellent.
exemplares	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	OCR. Average recognition due to the fact that it doesn't recognize any of the foreign letters or symbols.
Is the scan on a gray-scale or color scale?	Color.
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for	Zooming excellent.

any advantages that the print version would not allow? (zooming capabilities, filter options)	
Other notes	N/A

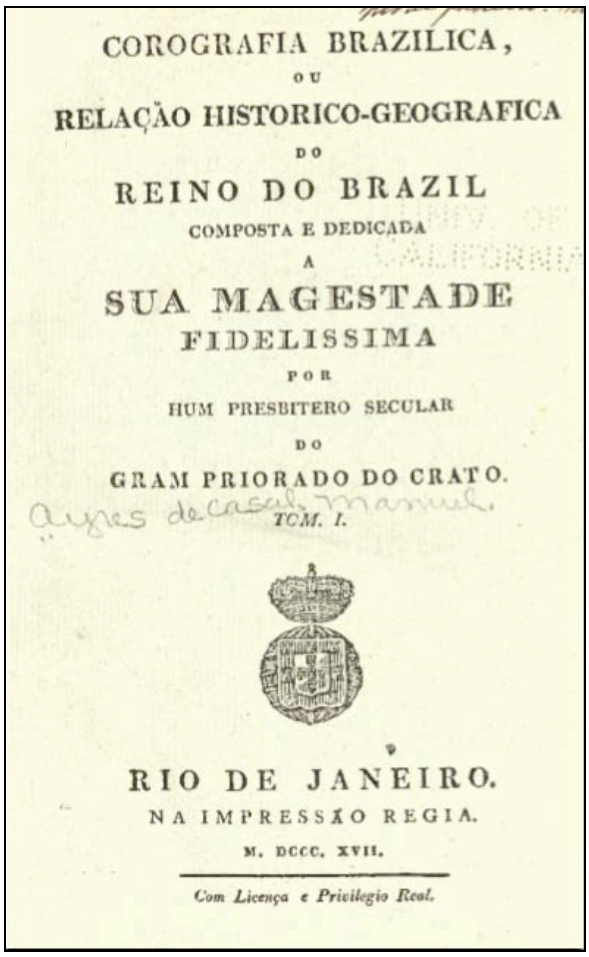
GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Not easy. There are no links to external information on the text so I went to WorldCat to find the particular edition.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes. But cannot search for accented words or symbols. And the pdf and the online version have differing searching abilities.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average. Since the book is a translation to English, nearly all of the book is searchable. However, accented words and other figures/characters are unrecognizable.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Most of it cannot be searched because it is not in English.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes. If its make searchable or if OAT redefined the search terms to allow for foreign language recognition.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	OAT has both volumes. But there is nothing to suggest within the actual entry that OAT has the other volume. There should be a link between them.
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Book URL: http://www.archive.org/details/corografiabrazilica01ayrerich
Date accessed: 6/28/08

Appendix I – OAT version of title page (left) and GBS (right).



COROGRAFIA BRAZILICA,
OU
RELAÇÃO HISTORICO-GEOGRAFICA
DO
REINO DO BRAZIL
COMPOSTA E DEDICADA
A
SUA MAGESTADE
FIDELISSIMA
POR
HUM PRESBITERO SECULAR
DO
GRAM PRIORADO DO CRATO.
TOM. I.

RIO DE JANEIRO.
NA IMPRESSÃO REGIA.
M. DCCC. XVII.

Com Licença e Privilegio Real.

COROGRAFIA BRAZILICA,
OU
RELAÇÃO HISTORICO-GEOGRAFICA
DO
REINO DO BRAZIL
COMPOSTA E DEDICADA
A
SUA MAGESTADE
FIDELISSIMA
POR Ayres de Casal
HUM PRESBITERO SECULAR
DO
GRAM PRIORADO DO CRATO.
TOM. I.

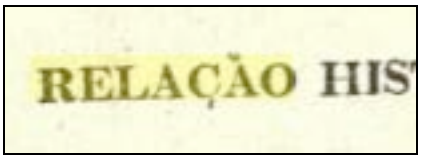
RIO DE JANEIRO.
NA IMPRESSÃO REGIA.
M. DCCC. XVII.

Com Licença e Privilegio Real.

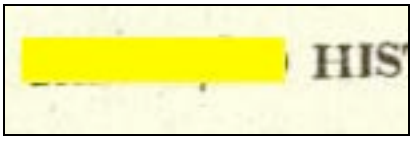
Appendix II – Browser Differences

Missy and I accidentally discovered that, depending on what browser a user is using, the search function capabilities alter. For example, in Firefox, the word being searched (relacao – without the accents) is highlighted but clearly legible. The same cannot be said for Safari:

Firefox



Safari



CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/17/08

http://en.wikipedia.org/wiki/The_Lusiads

Os Lusíadas (usually known by the title The Lusiads in English) is a Portuguese epic poem by Luís Vaz de Camões (sometimes spelled Camoens in English). It is considered one of the most important works of World literature.[citation needed]

Written in Homeric fashion, the poem focuses mainly on a fantastical interpretation of the Portuguese voyages of discovery during the 15th and 16th centuries. Os Lusíadas is often regarded as Portugal's national epic, much in the way as Virgil's Aeneid was for the Ancient Romans. It was first printed in 1572, three years after the author returned from the Indies.

Book title: Os Lusiadas, poema epico de Luis de Camões.
Author: Luís de Camões; Paulino de Souza; José Maria de Souza-Botelho
Year of Publication: 1865
Place of Publication: Paris
Publisher: J. P. Alliaud, Guillard e ca.
ISBN number?: OCLC 12136774
Content/Disciplinary area: Portuguese literature 17th century; Luis de Camões (1524-1580)

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Above average. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Poor. Google gets the author's name wrong, listing him as Domingos Vieira, when it is very clearly Antonio Vieyra. The metadata should also include the fact that there are two common spellings of the author's name (Vieyra or Vieira). This is especially problematic because the other volume is also mislabeled while an earlier edition of the same text (furthermore edited by the same editor) is correctly labeled.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) The page numbers in the upper corners are very hard to make out. I can't tell if it is a printing error or scanning error. Since it's a dictionary, its all the more easier to tell if a page is missing.

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Ex libris	Excellent (in color)	N/A	N/A	N/A
Title page	Excellent (in color)	N/A	N/A	N/A
Illustrations (first two – author and	Above average	Odd half-tone colors.	N/A	N/A

publisher's seal)				
2	Above average	Odd lighter shaded area – legibility not affected	N/A	N/A
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A
280, 290	Above average	N/A	N/A	N/A
300, 310	Above average	N/A	N/A	N/A
320, 330	Above average	N/A	N/A	N/A
340, 350	Above average	N/A	N/A	N/A
360, 370	Above average	N/A	N/A	N/A
380, 390	Above average	N/A	N/A	N/A
400, 410	Above average	N/A	N/A	N/A
420, 430	Above average	N/A	N/A	N/A
440, 450	Above average	N/A	N/A	N/A
460, 470	Above average	N/A	N/A	N/A
480, 490	Above average	N/A	N/A	N/A
500, 510	Above average	N/A	N/A	N/A
520, 530	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	**Comment** (Is OCR search permitted) **Y** / N (Available only in online version, not pdf)
Lusiadas	Excellent
Camões	Excellent
Souza	Excellent
Canto	Excellent
portuguez	Excellent
França	Excellent
271	Unreadable
tão	Excellent
céo	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes – 11 illustrations
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A

Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Above average recognition in online version. No search capabilities available in pdf however.
Is the scan on a gray-scale or color scale?	Gray-scale
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes. However zooming does not aid in the close-detail analysis.
Are the illustrations searchable?	No.
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes. Pixel quality needs to be higher.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming average due to pdf being gray-scale.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy
Overall quality of scan of this book	Above average
Ease of access to full text	Easy
Ability to copy, quote, cite?	No. Search results in right-hand column can be copied and pasted but only that particular fragment of the page
Ability to search?	Yes (only online version)
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Above average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No
Reproduction – is quality of scan good	Yes.

enough that this book could be reproduced into analog form?	
Issues with this book that would affect usefulness for scholars	N/A
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL:
http://books.google.com/books?id=mBNAAAAAIAAJ&dq=camoes&lr=lang_pt&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 8/17/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/04/08

http://en.wikipedia.org/wiki/Celestina

La Celestina (used as title, synecdoche, one of the characters of the book actually called *Tragicomedia de Calisto y Melibea* or *Libro de Calisto y Melibea y de la puta vieja Celestina*) is a novel published anonymously by Fernando de Rojas (about whom we know little) in 1499. This book is considered to be one of the greatest in Spanish literature, and traditionally marks the end of medieval literature and the beginning of the literary renaissance in Spain. The book is written against the servants of the low nobility and advises us to beware of their tricks and lies.

Book title: *Celestina: or, The Tragi-comedy of Calisto and Melibea*
Author: Rojas, Fernando de.
Year of Publication: 1908 (originally published in 1499)
Place of Publication: London
Publisher: Routledge
ISBN number?: OCLC 634108
Content/Disciplinary area: Picaresque literature

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. No available metadata on other editions, translations.
MARC record available?	Yes. Available through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	It would be nice to be able to link directly from the chapter title in the index to the page itself. That being said, there is no guide to the book provided by metadata, further complicated by the fact the book itself has no table of contents.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
Title page	Excellent	N/A	N/A	N/A
ix, xi	Excellent	slight;y difficult to read parts of margin due to spine tightness	N/A	N/A
x	Excellent	N/A	N/A	N/A
xix, xx	Excellent	pages xix and xx reversed in the pdf version, not the online version	N/A	N/A
xxx	Excellent	N/A	N/A	N/A

xl, l	Excellent	N/A	N/A	N/A
lx, lxx	Excellent	N/A	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
48, 49, 50, 51	Excellent	Pages out of sequence on pdf version (48, 50, 49, 51) not online version	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	Pages 158, 160, 159, 161 our of sequence on pdf version, not online version	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A
200, 210	Excellent	N/A	N/A	N/A
220, 230	Excellent	N/A	N/A	N/A
240, 250	Excellent	N/A	N/A	N/A
260, 270	Excellent	N/A	N/A	N/A
280, 290	Excellent	N/A	N/A	N/A
300, 310	Excellent	N/A	N/A	N/A
320, 330	Excellent	N/A	N/A	N/A
340	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	**Comment**
	(Is OCR search permitted) **Y** / N
Celestina	Excellent.
Cristóbal	Cannot be searched with the accent.
Cristobal	Only reads a couple instances of the word. Others not picked up. Not consistent between pdf and online version.
á, é, í, ú	Unreadable.
xxxix	Excellent.
36	Excellent.
?	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Color stamp.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF. Search recognition average.
Is the scan on a gray-scale or color	Color

scale?	
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming excellent..
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Average search capabilities. Cannot search for accented words.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Cannot search for accented words.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/celestina00rojauoft
Date accessed: 8/04/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/07/08

http://en.wikipedia.org/wiki/Cantar_de_Mio_Cid

El Cantar del Mio Cid is the oldest preserved Spanish epic poem (epopeya). The Spanish medievalist Ramón Menéndez Pidal included the "Cantar de Mío Cid" in the popular tradition he termed the mester de juglaria. Mester de juglaria refers to the medieval tradition according to which popular poems were passed down from generation to generation, being changed in the process. These poems were meant to be performed in public by minstrels (or juglares), who each performed the traditional composition differently according to the performance context--sometimes adding their own twists to the epic poems they told, or abbreviating it according to the situation. On the other hand, some critics (known as individualists) believe "El Cantar del Mio Cid" was composed by Per Abbad who signed the only existing manuscript copy, and as such is an example of the learned poetry that was cultivated in the monasteries and other centers of erudition. Per Abbad puts the date 1207 after his name, but the existing copy forms part of a 14th century codex in the Biblioteca Nacional de España (National Library) in Madrid, Spain. However, it is incomplete. The first page and two other pages in the middle are missing. It is written in medieval Spanish, the ancestor of modern Spanish.

Its current title is a modern invention by Ramón Menéndez Pidal; its original title is unknown. Some call it El Poema del Cid on the grounds that it is not a cantar but a poem made up of three cantares. The title has been translated into English as The Lay of the Cid and The Song of the Cid. Some English translations include the verse translation of W.S. Merwin and prose translation of Rita Hamilton and Janet Perry.

Book title: El cantar de mio cid
Author: Anonymous
Year of Publication: Not stated anywhere! (lacking explanation provided above)
Place of Publication: N/A
Publisher: N/A
ISBN number?: N/A
Content/Disciplinary area: Spanish Medieval Epic Poetry

AVAILABILITY SEARCH

Source of scanned book:	Open-Access Text.
Comments on condition/provenance of book:	Poor. **Up until this point in time, every text that I had used from Open-Access Text was of fantastic quality. However, they all happened to have been digitized by Microsoft originally. This is not the case with the text that I am going to look at today. It is not a pdf but a text document that has no metadata whatsoever. Even though it is awful, I want to look at it precisely because it is not consistent with the rest of what I have seen from Open-Access Text.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. There is no metadata.
MARC record available?	No because I do know the edition of this version, nor date of publication.
Can the book be searched through an OPAC? If so, which one?	N/A

Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	There is no guide to the book provided by metadata.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) This brings up the question – what then is the quality of a "book" that is merely a plain text pdf? When it cease to look like a book?

Page(s):	**readability / quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Poor	Not the actual cover of the book but more like a book report form.	N/A	N/A
1-10 Table of Contents	Above average	Color-coded.	N/A	N/A
10, 20	Excellent.	N/A	N/A	N/A
30, 40	Excellent.	N/A	N/A	N/A
50, 60	Excellent.	N/A	N/A	N/A
70, 80	Excellent.	N/A	N/A	N/A
90, 100	Excellent.	N/A	N/A	N/A
110	Excellent.	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) **Y** / N

Word	**Comments**
Mio Cid	Excellent.
Presentación	Excellent! The accented "ó" was recognized!
á, é, ó, ú, ñ	Only searchable as individual letters in the table of contents. For the remainder of the text, have to be searched as parts of words.
Çid	Recognized.
67	Numbers work for all of the pages involved.
3715	Stanzas work.
¿?	Recognized.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Plain text pdf. Recognition Excellent.
Is the scan on a gray-scale or color scale?	Gray-scale
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	N/A
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	N/A
Overall quality of scan of this book	Above average – everything recognized.
Ease of access to full text	Excellent.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Excellent – it recognizes all of the foreign characters.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Absolutely not.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	N/A bc this text sure must be analog.
Issues with this book that would affect usefulness for scholars	Can't use it because no information on its origins, publication date, etc. Who knows how accurate the text actually may be to editions that contain that information.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No.
Other notes	N/A

Book URL: http://www.archive.org/details/CantarDeMioCid
Date accessed: 8/07/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/10/08

Book title: The life and voyages of Christopher Columbus
Author: Washington Irving
Year of Publication: 1838
Place of Publication: New York
Publisher: Collins, Keese.
ISBN number? OCLC 13531792
Content/Disciplinary area: Columbus, Christopher; America - Discovery and exploration - Spanish.

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	I actually did not know that Washington Irving had an edition on Columbus. I was trying to find Columbus's early writings – either the diary of the first voyage or the letter to the Spanish monarchy – and had a lot of difficulty finding clear metadata on Google Books, Open Access Text, and EEBO.
MARC record available?	Yes. Available through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Very poor. There are countless editions, translations of Columbus's writings – confounded by the spelling variations of Columbus's name depending on the language involved. The metadata provided on Open Access Text is furthermore erroneous. The edition is from 1838 – stated date is 1834. Even within OAT, there are at the most 20 digitized versions of the same text – at times the same edition - from different lending libraries. Each one comes up as a different result – search data results not very clear and each file needs to be opened individually. Metadata very confusing.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
Covers (front, back)	Excellent	N/A	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A
200, 210	Excellent	N/A	N/A	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) Y / N
	Yes
Columbus	Excellent
Irving	Excellent
Introduction	Excellent
Chapter	Excellent
36	Excellent
31st	Excellent
?	Poor. Reads everywhere.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes. 4 illustrations + nice color stamp from one-time owner.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with excellent recognition.
Is the scan on a gray-scale or color scale?	Color.
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes.
Are the illustrations searchable?	Captions searchable.
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming great.
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	None.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	There are at the most 20 digitized versions on Open Access Text of the same text – at times the same edition. Each one comes up as a different result – search data results not very clear and each file needs to be opened individually. Metadata very confusing.

Book URL: http://www.archive.org/details/lifeandvoyages00irvirich
Date accessed: 7/10/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/5/08

Book title: A Translation of the Inferno of Dante Alighieri, in English Verse: With Historical Notes, and the Life of Dante. To which is Added, a Specimen of a New Translation of the Orlando Furioso of Ariosto.
Author: Dante Alighieri, Leonardo Bruni, Henry Boyd, John Scott, Thomas Warton, Lodovico Ariosto.
Year of Publication: 1785
Place of Publication: London
Publisher: C. Dilly
ISBN number?: OCLC 11122267
Content/Disciplinary area: Italy Epic Poetry; Italian Poetry – Fourteenth Century

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Above average. Scan is in gray-scale.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata includes short descriptions of the two volumes of this edition but does not indicate as to which of the two is this particular text. Nor does it link to the other volume.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata is average but entirely lacking in terms of a useful table of contents. There is a "section" division according to the Cantos that is not labeled with the actual word "canto" nor numerically, making it harder to locate the particular canto in question (in order words, because this is volume II, section 1 is canto XVII, not canto 1) The metadata also does not indicate where the other texts included in this publication, such as *Orlando Furioso*, are located by page number.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
165	Poor	Scan illegible	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A

280, 290	Above average	N/A	N/A	N/A
300, 310	Above average	N/A	N/A	N/A
320, 330	Above average	N/A	N/A	N/A
340, 350	Above average	N/A	N/A	N/A
360, 370	Above average	N/A	N/A	N/A
380, 390	Above average	N/A	N/A	N/A
400, 410	Above average	N/A	N/A	N/A
420, 430	Above average	N/A	N/A	N/A
440, 450	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N Online source searchable, not pdf.
breathe	Excellent
Orlando	Poor
person	Poor – many words with an 's' do not read due to the different typography used in the time period. 'person' is instead only found as 'pej-fon'.
pow'r	Excellent
canto	Excellent.
Virgil	Excellent
circle	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Average recognition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A

Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	N/A
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Metadata average but sufficient. Table of contents not very useful.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	OCR recognition bad.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	I can't figure out where the other volume is on Google Books.
Other notes	N/ a

Book URL:
http://books.google.com/books?id=tH8NAAAAQAAJ&dq=A+Translation+of+the+Inferno+of+Dante+Alighieri,+in+English+Verse:+With+Historical+Notes,+and+the+Life+of+Dante.+To+which+is+Added,+a+Specimen+of+a+New+Translation+of+the+Orlando+Furioso+of+Ariosto.&lr=&as_brr=1&source=gbs_summary_s&cad=0

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Date accessed: 9/5/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/24/08

[Today I searched for editions of *Os Sertões* (*Rebellion in the Backlands*) (1902) by Euclides da Cunha – probably the most important book in modern Brazilian literature. Mario Vargas Llosa was so impressed with it that he wrote his own version of it, *La guerra del fin del mundo* (*War at the End of the World*) (1981).
I am going to do a comparison of sorts between the full text version of an English translation available by ACLS Humanities E-Book and Google Book.]

Book title: *Rebellion in the backlands* (*Os sertões*)
Author: Cunha, Euclides da, 1866-1909.
Year of Publication: 1964, c1944 [original date of publication in Portuguese is 1902]
Place of Publication: Chicago
Publisher: University of Chicago Press
ISBN number?: No [OCLC number: 162351816]
Content/Disciplinary area: Brazil -- History -- Canudos Campaign, 1893-1897.

AVAILABILITY SEARCH

Source of scanned book:	ACLS Humanities E-Book. [University of Michigan Library, Scholarly Publishing Office Ann Arbor, Michigan] It can only be accessed through GEORGE however; non-library subscribers do not have permission to view the file.
Comments on condition/provenance of book:	Above average. The entire text version is available through OCR recognition but it does not look like a "page" or a "book" in a traditional sense.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	The best quality digitized versions appear to be of the English translation by Samuel Putnam. There are several available in Portuguese on random fan club websites but they so far appear to be text-only versions that must be downloaded chapter by chapter. I am surprised that there is no digitized version of the Portuguese available online yet, as it was published originally in 1902. The ACLS version is the full text while the Google Book version of the 1944 edition by the same publisher is only a limited preview.
MARC record available?	Yes. Available on GEORGE. [http://catalog.library.georgetown.edu/search~S4?/acunha%2C+euclides/acunha+euclides/1%2C1%2C23%2CB/marc&FF=acunha+euclides+da+1866+1909&13%2C%2C23]
Can the book be searched through an OPAC? If so, which one?	GEORGE and WorldCat both link to the electronic version of this book available on ACLS – subscription needed!
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	The table of contents links to the various sections within the book. It does not however link to any relevant external information regarding the existence of other editions, translations, etc. It also would be nice to have a function where the footnotes link to other pages in the book or external works that are referenced (the same for the appendix at the end). I would like to comment here that the metadata provided by Google Books on the earlier edition of this book by the same publisher is riddled with errors. First of all, the only tag provided is "Africa" – and yes, Africa has *absolutely nothing* to do with this book. Also one of the chapters is spelt wrong in the table of contents, making the word "Moreira" instead "Morkika". Please see Appendix I and II.

SCAN SEARCH

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

search paradigm settings (example: every 10 double-pages) Page(s):	readability/ quality of scan	anomalies?	other?	other?
xxiv	Poor	It is a map and the settings allow you to only zoom so far. Some names of cities remain illegible but it is possible to copy and paste the image and then alter it in a different program (that would be a lot of work however)	N/A	N/A
10, 20	Above average	The remainder of the book is perfectly legible but cannot be searched.	N/A	N/A
30, 40	Above average	N/A	N/A	N/A
50, 60	Above average	N/A	N/A	N/A
70, 80	Above average	N/A	N/A	N/A
88, 89	Above average	N/A	N/A	N/A
90, 100	Above average	N/A	N/A	N/A
110, 120	Above average	N/A	N/A	N/A
130, 140	Above average	N/A	N/A	N/A
150, 160	Above average	N/A	N/A	N/A
170, 180	Above average	N/A	N/A	N/A
190, 200	Above average	N/A	N/A	N/A
210, 220	Above average	N/A	N/A	N/A
230, 240	Above average	N/A	N/A	N/A
250, 260	Above average	N/A	N/A	N/A
270, 280	Above average	N/A	N/A	N/A
290, 300	Above average	N/A	N/A	N/A
310, 320	Above average	N/A	N/A	N/A
330, 340	Above average	N/A	N/A	N/A
350, 360	Above average	N/A	N/A	N/A
370, 380	Above average	N/A	N/A	N/A
390, 400	Above average	N/A	N/A	N/A
410, 420	Above average	N/A	N/A	N/A
430, 440	Above average	N/A	N/A	N/A
450, 460	Above average	N/A	N/A	N/A
470, 480	Above average	N/A	N/A	N/A
490, 500	Above average	N/A	N/A	N/A
510, 520	Above average	N/A	N/A	N/A
530	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) Y / **N**

Word	Comments
N/A	**N/A**

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes. There are several maps included through the book and both ACLS and Google Book include the images.

If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	Since its an OCR, no recognition of other graphics/illustrations.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	OCR. Average reconition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No. Not enough definition to make out the finest lines clearly enough
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Absolutely. Anyone interested in geography, war history, etc. would need to see the maps.
Are the illustrations searchable?	Yes and no. But presumably because Euclides included a subheading for "Maps" you can actually find the maps but not necessarily search the text within them.
Ability to print illustrations? (is there a watermark?)	ACLS yes but GBS no.
Are the images downloadable? (file size, pixel quality)	ACLS allows you to copy and paste the image but not Google Books.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming functions on both versions but could be better because the scan itself is not of the best quality.
Other notes	N/A

GENERAL OBSERVATIONS

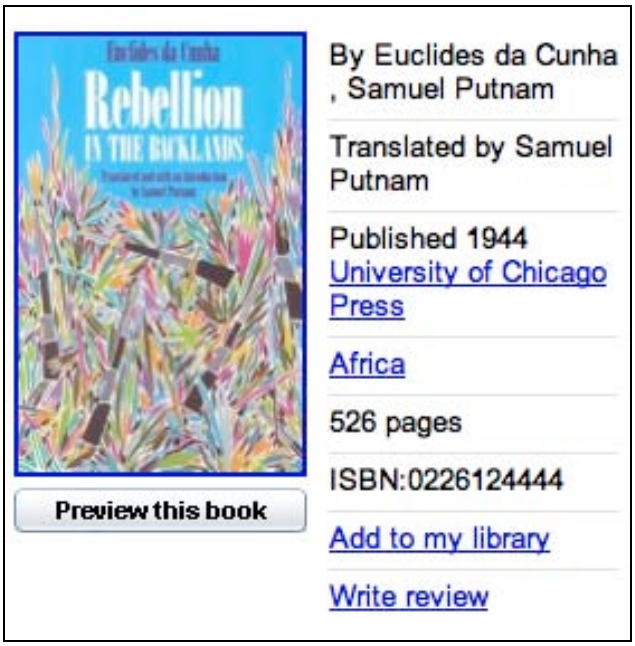
Comment on ease of discovery in integrated interface (OPAC or better)	Not easy. There are no links to external information on the text so I went to WorldCat to find the particular edition.
Overall quality of scan of this book	Above average.
Ease of access to full text	Above average. Would be nice to have the a version in the original language!
Ability to copy, quote, cite?	No.
Ability to search?	No. Can only go by page number or by viewing the table of contents. The GBS search is a little bit better but because it is only a limited preview, no search was done.
Ability to download?	No.
Ability to print?	Yes. Page by page.
Scanned book allow for consulting of contents?	No. Only by table of contents or by individual pages.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

How successful is OCR recognition?	Average. Since the book is a translation to English, nearly all of the book is searchable. However, accented words and other figures/characters are unrecognizable.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Cannot be searched.
Can this book be shared with other scholars/students?	No. I'd rather show them the printed version.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes. If its make searchable or if GBS allows access to the whole text, it would be more useful.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	ACLS does not have any of the other numerous editions of this book. Nor does it mention anywhere that these other editions do indeed exist.
Other notes	N/A

Book URL: http://0-hdl.handle.net.library.lausys.georgetown.edu/2027/heb.02746
Date accessed: 6/24/08

Appendix I – Google's Errors!
http://books.google.com/books?id=ofXM_d8mk5wC&dq=euclides+da+cunha&source=gbs_summary_s&cad=0



Africa has nothing to do with this book; to make matters worse, it is the only tag!

Appendix II – Another Google Book Error

THE MORKIKA CESAR EXPEDITION 225
Floriano Peixoto, Pitombas, Angico

should instead be

CHAPTER V
THE MOREIRA CESAR EXPEDITION

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/18/08

Book title: A Spanish anthology; a collection of lyrics from the thirteenth century down to the present time.
Author: Ford, J. D. M. (or Jeremiah Denis Matthias Ford)
Year of Publication: 1901
Place of Publication: New York
Publisher: Silver, Burdett and company
ISBN number?: OCLC 2549179
Content/Disciplinary area: Spanish poetry.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Good. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	I was looking for works by a specific poet and realized that it would probably be available in an anthology. Not only is this my first look at poetry in a digitized format but this particular edition is interesting because the introduction is in English while the poetry is in the original Spanish. This edition was the first hit on Google Books looking for the poet Nicasio Cienfuegos.
MARC record available?	Yes. Available through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	In the online version, both the table of contents and the index of authors link to the relevant sections. The metadata provided in the "Contents" section (yes, once again the automatically generated table of contents based on keywords) is nonsensical due to spelling errors. See Appendix I.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
Covers	Poor.	Lower half of the cover is unreadable due to blurriness (human error?)	N/A	N/A
Title page	Above average.	N/A	N/A	N/A
x, xx,	Above average.	N/A	N/A	N/A
xxx, xl	Above average.	N/A	N/A	N/A
l	Above average.	N/A	N/A	N/A
1, 10	Above average.	N/A	N/A	N/A
20, 30	Above average.	N/A	N/A	N/A
40, 50	Above average.	N/A	N/A	N/A
60, 70	Above average.	N/A	N/A	N/A
80, 90	Above average.	N/A	N/A	N/A
100, 110	Above average.	N/A	N/A	N/A
120, 130	Above average.	N/A	N/A	N/A
140, 150	Above average.	N/A	N/A	N/A
160, 170	Above average.	N/A	N/A	N/A

180, 190	Above average.	N/A	N/A	N/A
200, 210	Above average.	N/A	N/A	N/A
220, 230	Above average.	N/A	N/A	N/A
240, 250	Above average.	N/A	N/A	N/A
260	Above average.	N/A	N/A	N/A
269, 270, 271, 272	Above average.	Pages out of order.	N/A	N/A
280, 290	Above average.	N/A	N/A	N/A
300, 310	Above average.	N/A	N/A	N/A
320, 330	Above average.	N/A	N/A	N/A
340, 350	Above average.	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N **YES** but only in the online version, not the pdf.
Ford	Excellent.
Spanish	Excellent.
Introduction	Below average. Highlighted in all instances but most do not appear in search results.
Notes	Below average. Also should appear in all of the upper margins of the introduction but does not.
Antología	Below average. Highlighted in all instances but most do not appear in search results.
á	Excellent
año	Excellent
144	Excellent
Anónimo	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Recognition only available in only version not pdf.
Is the scan on a gray-scale or color scale?	Gray scale
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A

Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	No zooming available in online version (which is the searchable one).
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Online version only.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average. Not set for margins so headers unreadable.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	OCR needs to be fine-tuned.
Issues with this book that would affect usefulness for scholars	OCR average.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	OCR needs to be fixed.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Book URL:
Date accessed:

Appendix I – Misspellings in the "Contents" section (nonsensical metadata)

Contents

iNTROIMCTION xv
diphthong, Meléndez Valdés, synalcepha

ÍNDEX OF AUTHORS xlix

Aventura Amorosa Anónimo 3
Eya velar, vilano, tiesta

De las Propiedades que las Duennas 9
Eya velar, yaze, duenna pequenna

Á la Virgen 15
Sennora, Dize la Muerte, tu siruiente

Revelación de un Ermitaño Anónimo 21
feziste, Atanto, agora

more »

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/26/08

http://www.amazon.com/Battle-Tomochic-Memoirs-Lieutenant-Library/dp/0195117433

Tomochic is a controversial and celebrated example of Mexican fiction. *Tomochic* is the fictional narration of the 1892 military campaign that resulted in the massacre of the small village of *Tomochic*, located in the Tarahumara mountains and ordered by the dictatorial regime of Porfirio Diaz. The work is narrated by an eyewitness, the then second lieutenant, Heriberto Frias, and written by him in collaboration with Joaquin Clausell, editor of the newspaper which published it in serial form between March and April of 1893. For a period after the series' publication, the author chose to maintain anonymity. It was expressly this stance which excited more public interest than any other Mexican writer of the 19th century and which eventually led to a drawn out trial to uncover the identity of the author and to implicate him. For, although it is a work of fiction, the general plot of the work, involving a confrontation between a professional army and a handful of citizens, was too similar to the actual massacre as to not be seen by Porfirio Diaz as a reprovement of himself and his regime. As a piece of literature, the novel is also admired for its incorporation of two important trends of the nineteenth century-history as literature and the war novel.

Book title: *Tomóchic: Novela histórica mexicana*
Author: Frías, Heriberto
Year of Publication: 1906 [original date of publication is 1894]
Place of Publication: Buenos Aires
Publisher: Librería La Facultad, de J. Roldán.
ISBN number?: No [OCLC 3787014]
Content/Disciplinary area: Tomóchic (Mexico) -- History -- Fiction.

AVAILABILITY SEARCH

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Source of scanned book:	Google Book Search.
Comments on condition/provenance of book:	Above average.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Metadata is lacking. The only link is to the publishing company that does offer another edition of the book - also full text available on Google. No other metadata provided.
MARC record available?	Yes but had to go through WorldCat to send lending libraries MARC records.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Fantastic. The Google guide follows the book's chapter index.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability / quality of scan	anomalies?	other?	other?
Title page	Above average	Handwritten dedication – hard to transcribe due to gray-scale scan	N/A	N/A
2, 10, 20	Above average	N/A	N/A	N/A
30, 40	Above average	Page 30 has faint lines (possibly drawn?) – does not impact legibility	N/A	N/A
50, 60	Above average	N/A	N/A	N/A
70, 80	Above average	N/A	N/A	N/A
90, 100	Above average	N/A	N/A	N/A
110, 120	Above average	N/A	N/A	N/A
127	Above average	N/A	N/A	N/A
130, 140	Above average	N/A	N/A	N/A
150, 160	Above average	N/A	N/A	N/A
170, 180, 181	Above average	On 181 Part of left leaf on right leaf on top of margin. Does not impact legibility.	On 180, last word in each sentence looks stretched. Does not impact legibility.	N/A
190, 200	Above average	N/A	N/A	N/A

210, 220	Above average	N/A	N/A	N/A
230, 240	Above average	N/A	N/A	N/A
250, 260	Above average	On 260, last word in each sentence looks stretched. Does not impact legibility.	N/A	N/A
270, 280	Above average	Page 270 has some red ink markings.	Page 280 has some red ink markings.	N/A
290, 300	Above average	On 290, underlining. Does not impact legibility.	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) Y / **N**

Word	Comments
N/A	**PDF cannot be searched but the online version can be.** Most foreign characters and symbols are recognized online.
Frías	Recognized.
á, é, ó	Recognized.
ñ	Not recognized in all but one instance.
Extrañas	Recognized. It appears that accented letters by themselves cannot be searched as part of a word. For example, trying to find all of the words that have "ñ" doesn't work but typing the word out "extraña" does work.
¿?	**¿** not recognized.

IMAGE ASSESSMENT

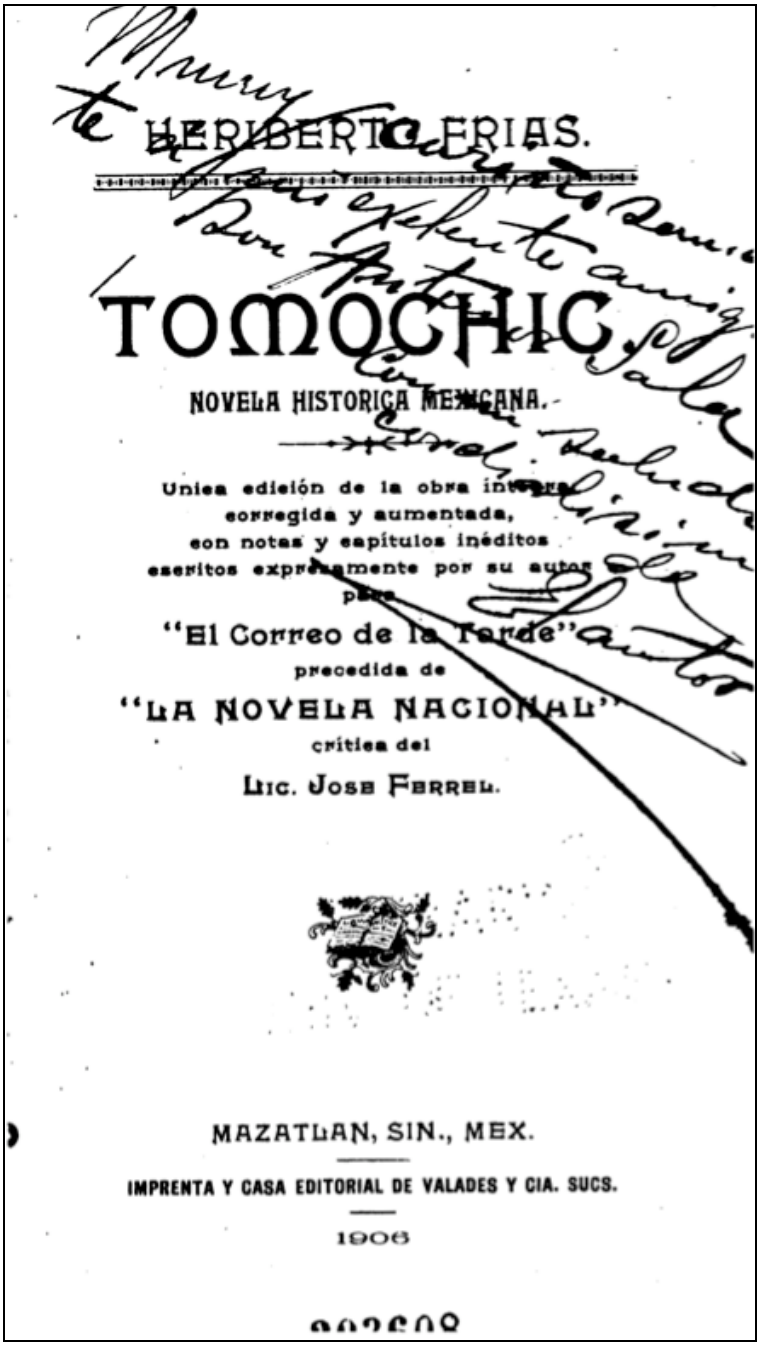
If the original text includes graphics/illustrations, are they included in the electronic version?	Only one illustration of author.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	Can't tell because on gray-scale all markings become indistinguishable from one and other.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with no recognition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Probably not.
Are the illustrations searchable?	No
Ability to print illustrations? (is there a	Yes

watermark?)	
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming not helpful.
Other notes	It would be nice to see the written dedication on the title page in color. Its very hard to make out on a gray-scale if one is trying to do transcription. See Appendix I.

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Above Average.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	Yes.
Ability to search?	No for the pdf and yes for the online version.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Not available in pdf format but excellent online.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Metadata needs to be better.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No.
Other notes	N/A

Book URL:
http://books.google.com/books?id=_NUSAAAAYAAJ&pg=PP2&dq=tomochic&as_brr=1&ei=AvpnSLGjEZDyiwHrht23BA
Date accessed: 6/26/08



HERIBERTO FRIAS.

TOMOCHIC

NOVELA HISTORICA MEXICANA.

Unica edición de la obra integra,
corregida y aumentada,
con notas y capítulos inéditos
escritos expresamente por su autor
para
"El Correo de la Tarde"
precedida de
"LA NOVELA NACIONAL"
crítica del
LIC. JOSE FERREL.

MAZATLAN, SIN., MEX.

IMPRENTA Y CASA EDITORIAL DE VALADES Y CIA. SUCS.

1906

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/19/08

http://en.wikipedia.org/wiki/Inca_Garcilaso_de_la_Vega

Garcilaso de la Vega, (b. Gómez Suárez de Figueroa, April 12, 1539 in Cuzco, Peru, d. 1616) was an illustrious Peruvian historian and acclaimed writer who is recognized primarily for his contributions to Inca history, culture, and society. Although not all scholars agree, many consider Garcilaso's accounts the most complete and accurate available.

Garcilaso de la Vega was born of Spanish aristocratic and royal Inca roots in Cusco, Peru on April 12, 1539.[1] Garcilaso was the illegitimate son of Spanish captain and conquistador Sebastián Garcilaso de la Vega y Vargas (d. 1559).[1] Garcilaso's mother, Inca princess Palla Chimpu Ocllo, baptized as Isabel Suárez Chimpu Ocllo, was descended from Inca nobility, a daughter of Tupac Huallpa and a granddaughter (not a niece) of the powerful Inca Tupac Yupanqui.[1][2] Garcilaso lived with his mother the first ten years of his life and learned to speak both Quechua and Spanish.[3]

Book title: The royal commentaries of Peru…
Author: Vega, Garcilaso de la, 1539-1616.
Year of Publication: 1688
Place of Publication: London
Publisher: Printed by Miles Flesher, for Samuel Heyrick
ISBN number?: OCLC 165942547
Content/Disciplinary area: Incas - Early works to 1800; Indians of South America – Peru; Peru -- History - To 1820.

AVAILABILITY SEARCH

Source of scanned book:	Early English Books Online (EEBO)
Comments on condition/provenance of book:	Good. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy. The record for the book is very complete. Three other editions available on EEBO indicated by search results for author's name.
MARC record available?	No. The only record of this edition in WorldCat is in a German library.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	The record is very complete with information about this specific edition. But does not link to anything inside or outside the text. The book has a fairly long index but EEBO provides no table of contents and each page can only be downloaded separately. There is however a text version available in addition to the pdf.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)	**Page edges visible in the entire text and at times the scan angles are distracting (angles due to two pages being scanned simultaneously; rather than one page at a time). See Appendix I.

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
title page	Above average.	Black spots – legibility slightly affected.	N/A	N/A
10, 20	Above average.	N/A	N/A	N/A
30, 40	Above average.	N/A	N/A	N/A

50, 60	Above average.	N/A	N/A	N/A
70, 80	Above average.	N/A	N/A	N/A
90, 100	Above average.	N/A	N/A	N/A
110, 120	Above average.	N/A	N/A	N/A
130, 140	Above average.	N/A	N/A	N/A
150, 160	Above average.	N/A	N/A	N/A
170, 180	Above average.	N/A	N/A	N/A
190, 200	Above average.	N/A	N/A	N/A
210, 220	Above average.	N/A	N/A	N/A
230, 240	Above average.	N/A	N/A	N/A
250, 260	Above average.	N/A	N/A	N/A
270, 280	Above average.	N/A	N/A	N/A
290, 300	Above average.	N/A	N/A	N/A
310, 320	Above average.	N/A	N/A	N/A
330, 340	Above average.	N/A	N/A	N/A
350, 360	Above average.	360 has black circular mark (wormhole?) – legibility affected.	N/A	N/A
370, 380	Above average.	N/A	N/A	N/A
390, 400	Above average.	N/A	N/A	N/A
410, 420	Above average.	N/A	N/A	N/A
430, 440	Above average.	N/A	N/A	N/A
450, 460	Above average.	N/A	N/A	N/A
470, 480	Above average.	480 has black circular mark – legibility affected.	N/A	N/A
490, 500	Above average	490 slightly blurry – legibility not affected.	N/A	N/A
510, 520	Above average.	N/A	N/A	N/A
530, 540	Above average.	N/A	N/A	N/A
550, 560	Above average.	N/A	N/A	N/A
570, 580	Above average.	N/A	N/A	N/A
590, 600	Above average.	N/A	N/A	N/A
610, 620	Above average.	N/A	N/A	N/A
630, 640	Above average.	N/A	N/A	N/A
650, 660	Above average.	650 slightly blurry – legibility affected.	N/A	N/A
670, 680	Above average.	N/A	N/A	N/A
690, 700	Above average.	N/A	N/A	N/A
710, 720	Above average.	N/A	N/A	N/A
730, 740	Above average.	N/A	N/A	N/A
750, 760	Above average.	N/A	N/A	N/A
770, 780	Above average.	N/A	N/A	N/A
790, 800	Above average.	N/A	N/A	N/A
810, 820	Above average.	N/A	N/A	N/A
830, 840	Above average.	N/A	N/A	N/A
831	Poor	Extremely blurry – legibility affected.	N/A	N/A
850, 860	Above average.	N/A	N/A	N/A
870, 880	Poor	870 extremely blurry – legibility affected.	N/A	N/A
890, 900	Above average.	N/A	N/A	N/A
910, 920	Above average.	N/A	N/A	N/A
930, 940	Above average.	N/A	N/A	N/A
950, 960	Above average.	N/A	N/A	N/A
970, 980	Above average.	N/A	N/A	N/A
990, 1000	Above average.	N/A	N/A	N/A
1010	Above average.	N/A	N/A	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
(Is OCR search permitted) Y / **N** **No**	

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF. No recognition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes.
Are the illustrations searchable?	Yes. They have their own search engine for illustrations versus images! See Appendix II.
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming very nice for pages with just text. However loses detail in illustrations due to gray-scale. See Appendix III.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in	Easy.

integrated interface (OPAC or better)	
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy. However, can't view the whole thing at once – have to open each page one at a time.
Ability to copy, quote, cite?	Yes from the text version not the pdf.
Ability to search?	Only the text version not the pdf.
Ability to download?	Yes. Each page individually.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes. There is a nice feature that allows you to see the thumbnails for every page. So although you can't open the whole thing you can have an overview of the contents.
How successful is OCR recognition?	N/A
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	The scan quality of the illustrations is not very good. Also can't search the text much less easily view it since each page has to be viewed individually.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No because no OCR recognition.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://0-gateway.proquest.com.library.lausys.georgetown.edu/openurl?ctx_ver=Z39.88-2003&res_id=xri:eebo&rft_id=xri:eebo:citation:12374219
Date accessed: 7/19/08

Appendix I – Page Angles

him. *Is it possible,* said he, *that you should fansie and imagine me to be so ungratefull as to forget the kind and charitable entertainment I received from you, since years past in this very City? or to believe, that in return thereof I should not embrace all occasions to make known my gratitude? is it possible for me to be so short of memory as not to remember how kindly you took me and my Family to your own home, (when there was no place to receive us) and entertained us there for dayes and months, untill such time as Marquis* Piçarro, *of glorious memory, had made other provisions for me. And having ever conserved the thoughts hereof in my mind, I carried great respect to every thing wherein you were concerned; for though I had sufficient information, that* Diego Centeno *was concealed within your possession, and though I knew the very Cave it self where he was hid, and nourished by your Indians, yet I winked at it, and took no notice thereof, that I might not give you trouble, nor bring you under a prejudice or ill notion with my Lord the Governour. I might then easily (if I had pleased) have sent Files of Musquetiers and have brought* Diego Centeno *to me; but for your sake I forbore to persecute him, notwithstanding the mortal enmity between us. Nor indeed did I much value him; for he who could submit to take his refuge in a Cave was not much to be feared; not doubting but wheresoever he should adventure again abroad, and dare to take up Arms against my Lord the Governour, I should drive him into another Cave, as I have done since his flight from* Huarina, *where by the blessing of God, and assistance of Friends, I gained a signal Victory. And now, since for your sake I took not vengeance on an enemy whom I had within my power, how much more should I respect your person, and those whom you own for your friends and acquaintance? and to assure you of these my real intentions, I give you free leave to return unto your home, and there repose and rest with as much freedom and security as you can desire. And also you may assure all the Inhabitants of this City, and all such as were taken and brought back in your company, that for your sake they are pardoned and exempted from that punishment which their offences have deserved.* And herewith he acquitted and dismissed *Miguel Cornejo*, and freed the City of the fears and apprehensions they were in for having with zeal and affection shewed themselves active in his Majesty's Service, and for the assistance they had given to *Diego Centeno*. This Story I learned, not onely by common report, but from the particular relation of *Gonçalo Silveyra*, who was a professed enemy to *Francisco de Carvajal*, as he was a stiff and faithfull friend to *Diego Centeno*, having adhered to him and accompanied him in all his adversities and misfortunes till the day of his death, as will appear hereafter. And on this occasion I produce this authentick Witness to confirm the truth of what I have spoke; for I pretend to flatter no man, but really and truly to lay down matters of fact according as they have passed.

Carvajal, having gathered what Arms and Horses he could find in *Arequepa*, returned to *Piçarro*, who was still on his march to *Cusco*: for by reason of the many sick and wounded in the Camp, he was forced to continue longer at *Huarina* than was intended. And here it is very observable, that those of *Piçarro*'s Party who were rich took care and compassion of the wounded men belonging to *Centeno*, and dividing them amongst themselves carried them to their own Tents, where they provided for their maintenance and cure. It was my Father's lot to have twelve assigned to him, of which six dyed on the march, and six happened to live and recover. I knew two of them, one was called *Diego de Papia*, a Gentleman of good Quality, and a person of vertue, who made many gratefull acknowledgments for the care which was taken of him. When I came for *Spain* I left him in the House of *Diego de Silveia*, who was Godfather to me at my Confirmation. The other was called *Francisco de Penna*, whose name agreed well with his nature, which signifies a Rock, and indeed so hardy was he, that though he had received three great slashes on the Crown of his Head, from the first to the last of which there was about three fingers distance; and though his Skull was so broke that the Chirurgeon was forced to take out some Bones, and having no better Instruments to trepan his Head, he made use of a Smith's Pincers in this operation, yet had he the fortune and strength to recover. And this further shewed the soundness and hardiness of his complexion, that, notwithstanding all these wounds, and the unskilfulness of the Instruments that he recovered without Fever, or any other accident; and yet kept no rule or regimen in his Diet, but ate whatsoever was set before him. The which indeed was very strange, and never heard or seen before; and so might well be called *Francisco Penna* rather than *de Penna*, which is *Francis Rock*, and not *Francis of the Rock*. And having said thus much, it will be now time for us to return to the President.

CHAP.

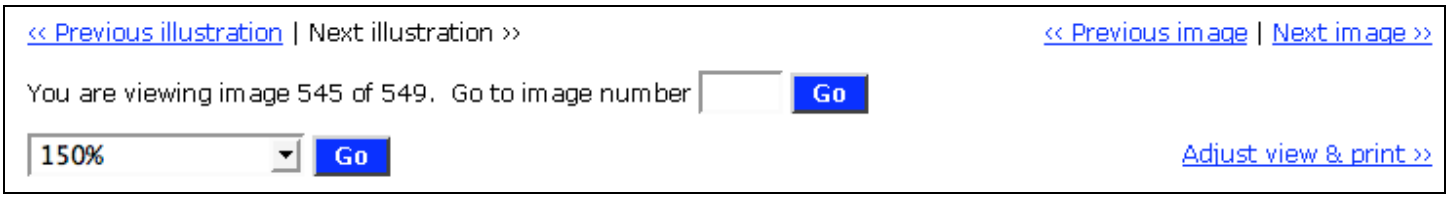
CHAP. XXVI.

How the President and his Army received the ill news of the Victory of Piçarro; *and what new Orders he made.*

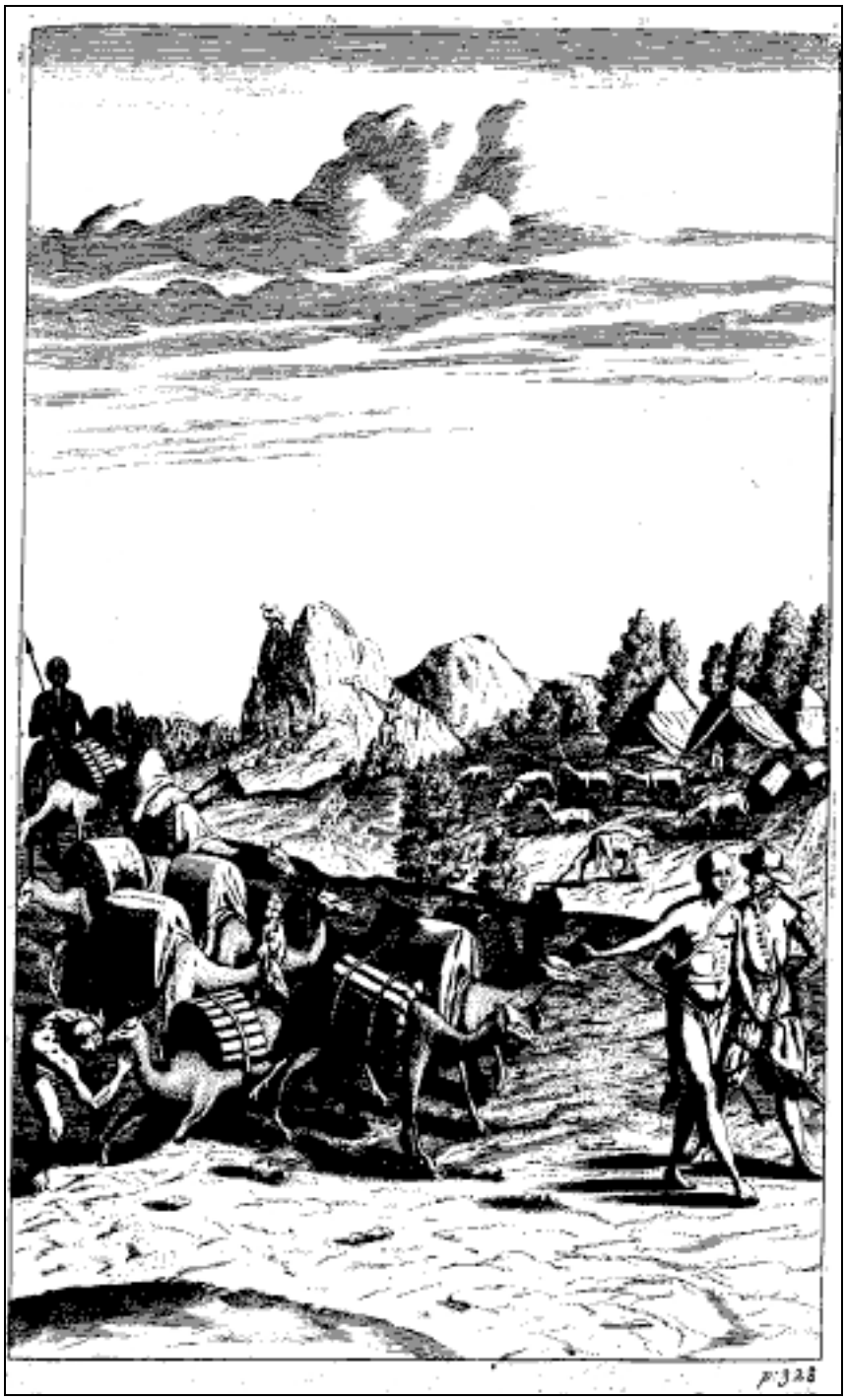
HOW vain and short are the prosperities and successes of this life may appear by what happened in the Valley of *Sauja*, where we left the President and his Army full of joy and contentment upon the good news which *Francisco Piçarro* lately brought concerning the flourishing condition of *Centeno*'s Army, which was much superiour to *Piçarro*'s both in Men, Arms and Horse. In confidence of which the President, and those who were of his Council, resolved to raise no more men, but to disband the Army which they had already formed, and dismiss the Souldiers who were come from such distant Countries and Provinces; for that it seemed an unnecessary and fruitless expence to keep so many men in pay in expectation of an Enemy already presumed to be overcome and defeated. And this course was so far resolved, that (as *Augustine Cavate* saith in the fourth Chapter of his seventh Book) the Army was just upon the point of being disbanded; the which he expresses in these words: At this time came intelligence to the President of the Defeat given to *Diego Centeno*, which he took much to heart, though he expressed not so much publickly, but seemed to carry it outwardly without much concernment: all which happened so contrary to the common expectation of the Army, that many of the Officers were of opinion, that the President had no need to raise an Army, for that *Diego Centeno* had a sufficient force to overthrow *Piçarro*, &c. Thus far *Cavate*.

But as it was the good fortune of the President, so it was the misfortune of his Enemies that this determination was not put in execution; nor was it so much as known that this matter was ever deliberated or designed to be made publick; for if the Army had been disbanded, it had certainly been very difficult to have raised another of the like nature, or so well provided with Men and Ammunition as this was. As matters were in this agitation the Bishop of *Cusco* arrived at *Sauja*, with the ill news of the Rout and Defeat of *Diego Centeno*, of all which he having been an eye-witness was best able to give a relation. The President, and other Inhabitants of that Country, who had Baronies and commands over the *Indians*, were greatly surprised and troubled at the news; finding that the heat of War, which they considered lately as extinguished, was now again inflamed, with such advantage to the Enemy, that their cause seemed almost desperate, and the Rebels in a state not to be reduced. Howsoever the Captains and Souldiers, who live by War, were not in the least disturbed at this news; for War is their livelihood and trade, from whence they derive their Honours and Estates, and especially in the Empire of *Peru*, where the reward of Souldiers and Adventurers is Command, and division of Lands, and a Barony over *Indians*. The President being desirous to hold up the hearts of his people, which he found drooping, put the best countenance he could upon the matter, and made them a short Speech in this manner: That Successes of this nature were usual and common over all the world, and were incident to War; and that these things were intended for the good of God's People, who were to give thanks to his Divine Majesty for having granted to *Piçarro* a Victory, with intent to bring greater mischiefs on his head, and that having filled up the measure of his iniquity, his ruine and downfall would be the greater. And to obtain this great blessing, he advised every Officer to be vigilant and diligent in his station, and to order and dispose all things in such a method and readiness as might enable them to contend with such an Enemy. He added further, that there was no need for him to move and incite Cavaliers so well experienced in War as they were to actions of bravery and to martial exploits; for that he was to follow their example, and take their counsel and advice in matters of greatest importance: In fine, he told them, that he was well-assured that they would direct every thing in the best manner to the Service of his Majesty, who would gratifie their fidelity with rewards agreeable to their merits, and make them Lords of that whole Empire.

Appendix II - Searchable Images vs. Illustrations



Appendix III – Lost Detail



CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/7/08

Huge file – took forever to download.

http://en.wikipedia.org/wiki/Dead_Souls

Dead Souls (Russian language: М ё р т в ы е д у ш и) by Nikolai Gogol was first published in 1842, and is one of the most prominent works of 19th century Russian literature. Gogol himself saw it as an "epic poem in prose", and within the book as a "novel in verse". Despite supposedly completing the trilogy's second part, Gogol destroyed it shortly before his death. Although the novel ends in mid-sentence (like Sterne's Sentimental Journey), it is usually regarded as complete in the extant form.[1]

Book title: Dead souls
Author: Gogol, Nikolai Vasilevich, 1809-1852; Graham, Stephen, 1884-
Year of Publication: 1915 (first edition 1893)
Place of Publication: London
Publisher: Fisher Unwin
ISBN number?: OCLC 37639822
Content/Disciplinary area: Russian literature – 19th century.

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Above average. The copy is marked up with pencil and pen.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata should have included different editions because WorldCat only has first edition listed not later.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	Other?	Other?
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
193-208	Above average	N/A	N/A	N/A

200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A
280, 290	Above average	N/A	N/A	N/A
300, 310	Above average	N/A	N/A	N/A
320, 330	Above average	N/A	N/A	N/A
340, 350	Above average	N/A	N/A	N/A
360, 370	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N Online version only, not pdf.
dead	Excellent
dead souls	Excellent – comes back as a phrase
Gogol	Excellent
Russian	Excellent
Cossack	Excellent
1915	Excellent
222	Excellent.
London	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition
Is the scan on a gray-scale or color scale?	Color-scale
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for	Zooming good.

any advantages that the print version would not allow? (zooming capabilities, filter options)	
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average. Marked up with pen and pencil.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Yes – online version only not pdf.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes but copy should be nicer (i.e. not marked up)
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Marked up. Aesthetically not pleasing but otherwise legible.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/deadsouls00gogouoft
Date accessed: 9/7/08

Researcher name: Patricia A. Soler **Date of investigation**: 8/25/08

40 million copies printed!!

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

http://www.amazon.com/Message-Garcia-Gift-Editions/dp/0880884347

This literary trifle, A Message To Garcia, was written one evening after supper, in a single hour. It was on the 22nd of February, 1899, Washington's Birthday. And on an order of the Mikado, a copy was given to every man in the employ of the Japanese Government, soldier or civilian. Over forty million copies of A Message To Garcia have been printed. This is said to be a larger circulation than any other literary venture has ever attained during the lifetime of an author, in all history.

Book title: A Message to Garcia and Other Essays
Author: Hubbard, Elbert.
Year of Publication: 1924
Place of Publication: New York
Publisher: Thomas Y. Crowell Company
ISBN number?: OCLC 7368272
Content/Disciplinary area: Literature - Spanish American War

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Above average. Scan is in gray-scale.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata incorrectly states that the date of publication is 1916 instead of 1924. 1916 was indeed the date of the first edition from the publisher but this particular edition is from the later date of 1924.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Average but sufficient. Open Access Text has other works by Hubbard but in my opinion there should be at least a blurb on the real-life Andrew Summer Rowan (the main character in Hubbard's piece) and his response work "How I carried the message to Garcia"

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N
Garcia	Excellent
Hubbard	Excellent
Publisher's	Excellent
1915	Excellent
"Little Journeys to the Homes	Excellent. However the whole sentence (which runs two lines) is not searchable. Only words or the words that are in that particular line.

"The Boy From"	Excellent
Rowan	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes. One illustration of author.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes.
Are the illustrations searchable?	No.
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming nice but not as nice as it would be in color.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Metadata average but sufficient. No table of contents included in metadata.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Metadata leaves a lot to be desired.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	

Book URL: http://www.archive.org/details/messagetogarciaa014275mbp
Date accessed: 8/25/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/24/08

[Today I searched for Bartolomé de las Casas's most famous work, The Destruction of the Indies. The only full-text version I could find is on EEBO, with a slightly different title – as history tells us, several Protestant countries used las Casas's works in order to promote their own political agendas. A more accurate translation for the title below would be: *A short account of the destruction of the Indies*. There is another (related) book printed together with this edition but I will only be looking the first work included by las Casas.]

Book title: *An account of the first voyages and discoveries made by the Spaniards in America containing the most exact relation hitherto publish'd, of their unparallel'd cruelties on the Indians, in the destruction of above forty millions of people : with the propositions offer'd to the King of Spain to prevent the further ruin of the West-Indies.*
Author: Casas, Bartolomé de las, 1474-1566
Year of Publication: 1699
Place of Publication: London
Publisher: Printed by J. Darby for D. Brown [and 2 others]
ISBN number?: No
Content/Disciplinary area: Casas, Bartolomé de las, 1474-1566; Early English books online; Spain -- Colonies – America; Indians, Treatment of; Slavery -- Latin America

AVAILABILITY SEARCH

Source of scanned book:	Early English Books Online (EEBO).
Comments on condition/provenance of book:	Below average. It is hard to ascertain the condition of the book with a gray-scale scan set too dark. Either the book is damaged or the scanner set it too dark, resulting in some pages being unreadable.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy. A search for the author's name on GEORGE led me to all of his works; this one had an electronic link directly to EEBO.
MARC record available?	Yes. Available on GEORGE.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata links to other words by author available on EEBO. Quite useful considering that there are several. However, the metadata does offer the option to view the "illustrations" from the text. Upon choosing that option however, it only resulted in two graphs (better, lists) and not in the drawn illustrations.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability / quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Above Average	Not in color but entirely legible.	N/A	N/A
1, 10, 20	Poor	Page 1is too dark. Discoloration makes some words illegible and I	N/A	N/A

1, 10, 20	Poor	Page 1is too dark. Discoloration makes some words illegible and I can make out words from the opposite side of the same page. Page 20 is much better.	N/A	N/A
30, 40	Average	Page 40 has several dark "spots" but is still legible.	N/A	N/A
50, 60	Average	N/A	N/A	N/A
75, 76	Average	Fold page tab in upper corner.		
70, 80	Poor	Page 80 is very discolored. Some words illegible.	N/A	N/A
90, 100	Poor	Page 100 is very discolored. Many words illegible. See Appendix II.	N/A	N/A
110, 120	Average	N/A	N/A	N/A
130, 140	Average	N/A	N/A	N/A
150, 160	Average	N/A		N/A
170, 180	Average	N/A	N/A	N/A
190, 200	Average	N/A	N/A	N/A
210, 220	Average	N/A	N/A	N/A
230, 240	Average	240 blurry. Still legible.	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) Y / **N**

30, 40	Average	Page 40 has several dark "spots" but is still legible.	N/A	N/A
50, 60	Average	N/A	N/A	N/A
75, 76	Average	Fold page tab in upper corner.		
70, 80	Poor	Page 80 is very discolored. Some words illegible.	N/A	N/A
90, 100	Poor	Page 100 is very discolored. Many words illegible. See Appendix II.	N/A	N/A
110, 120	Average	N/A	N/A	N/A
130, 140	Average	N/A	N/A	N/A
150, 160	Average	N/A		N/A
170, 180	Average	N/A	N/A	N/A
190, 200	Average	N/A	N/A	N/A
210, 220	Average	N/A	N/A	N/A
230, 240	Average	240 blurry. Still legible.	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) Y / **N**

Word	**Comments**
N/A	**N/A**

IMAGE ASSESSMENT

If the original text includes Yes

graphics/illustrations, are they included in the electronic version?	
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	Smudges or discoloration is exacerbated with the gray-tone scan.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF. No OCR recognition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes – and they are terrible quality reproductions. See Appendix I.
Are the illustrations searchable?	No
Ability to print illustrations? (is there a watermark?)	Yes but they would look even more terrible printed than on the computer.
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	N/A
Other notes	N/A

GENERAL OBSERVATIONS

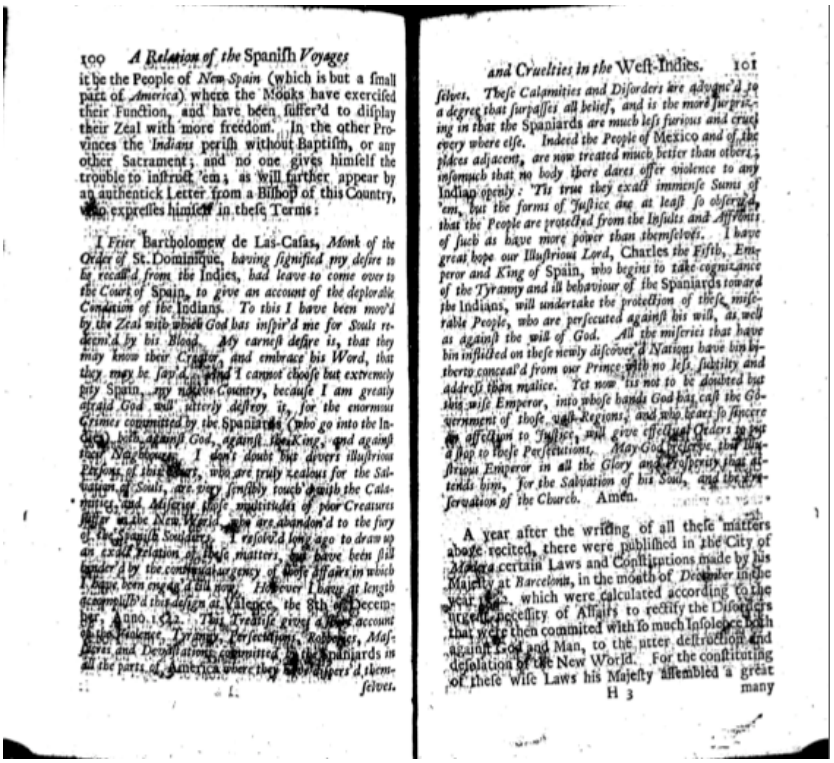
Comment on ease of discovery in integrated interface (OPAC or better)	Not easy. There are no links to external information on the text so I went to WorldCat to find the particular edition. And they got the author information plus, in addition to not including the publisher information.
Overall quality of scan of this book	Poor.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	No.
Ability to search?	No.
Ability to download?	Yes.
Ability to print?	Yes.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	N/A
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	Scanning quality is terrible.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No. The scanning is so blurry/indecipherable in some places it simply needs to be redone not repurposed.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No.
Other notes	N/A

Book URL: http://0-eebo.chadwyck.com.library.lausys.georgetown.edu/search/full_rec?SOURCE=pgimages.cfg&ACTION=ByID&ID=V57865
Date accessed: 6/24/08
Appendix I – Poor quality illustration reproductions

Appendix II: Scanner Difficulties



100 *A Relation of the* Spaniſh *Voyages*

it be the People of *New Spain* (which is but a ſmall part of *America*) where the Monks have exerciſed their Function, and have been ſuffer'd to diſplay their Zeal with more freedom. In the other Provinces the *Indians* periſh without Baptiſm, or any other Sacrament; and no one gives himſelf the trouble to inſtruct 'em; as will further appear by an authentick Letter from a Biſhop of this Country, who expreſſes himſelf in theſe Terms:

I Frier Bartholomew de Las-Caſas, *Monk of the Order of* St. Dominique, *having ſignified my deſire to be recall'd from the* Indies, *had leave to come over to the Court of* Spain, *to give an account of the deplorable Condition of the* Indians. *To this I have been mov'd by the Zeal with which God has inſpir'd me for Souls redeem'd by his Blood. My earneſt deſire is, that they may know their Creator, and embrace his Word, that they may be ſav'd. And I cannot chooſe but extremely pity* Spain, *my native Country, becauſe I am greatly afraid God will utterly deſtroy it, for the enormous Crimes committed by the* Spaniards *(who go into the Indies) both againſt God, againſt their King, and againſt their Neighbours. I don't doubt but divers illuſtrious Perſons of this Court, who are truly zealous for the Salvation of Souls, are very ſenſibly touch'd with the Calamities and Miſeries theſe multitudes of poor Creatures ſuffer in the New World, who are abandon'd to the fury of the Spaniſh Souldiery. I reſolv'd long ago to draw up an exact Relation of theſe matters, but have been ſtill hinder'd by the continual urgency of thoſe Affairs in which I have been engag'd till now. However I have at length accompliſh'd that deſign at* Valence, *the 8th of* December, *Anno* 1542. *This Treatiſe gives a juſt account of the Violence, Tyranny, Perfidiouſneſs, Robberies, Maſſacres and Deſolations committed by the* Spaniards *in all the parts of* America *where they have diſperſ'd themſelves.*

and Cruelties in the Weſt-Indies. 101

ſelves. Theſe Calamities and Diſorders are advanc'd to a degree that ſurpaſſes all belief, and is the more ſurpriſing in that the Spaniards *are much leſs furious and cruel every where elſe. Indeed the People of* Mexico *and of the places adjacent, are now treated much better than others, inſomuch that no body there dares offer violence to any* Indian *openly: 'Tis true they exact immenſe Sums of 'em, but the forms of Juſtice are at leaſt ſo obſerv'd, that the People are protected from the Inſults and Affronts of ſuch as have more power than themſelves. I have great hope our Illuſtrious Lord,* Charles *the Fifth, Emperor and King of* Spain, *who begins to take cognizance of the Tyranny and ill behaviour of the* Spaniards *toward the* Indians, *will undertake the protection of theſe miſerable People, who are perſecuted againſt his will, as well as againſt the will of God. All the miſeries that have bin inflicted on theſe newly diſcover'd Nations have bin hitherto conceal'd from our Prince with no leſs ſubtility and addreſs than malice. Yet now 'tis not to be doubted but this wiſe Emperor, into whoſe hands God has caſt the Government of thoſe vaſt Regions, and who bears ſo ſincere an affection to Juſtice, will give effectual Orders to put a ſtop to theſe Perſecutions. May God preſerve this illuſtrious Emperor in all the Glory and Proſperity that attends him, for the Salvation of his Soul, and the preſervation of the Church.* Amen.

A year after the writing of all theſe matters above recited, there were publiſhed in the City of *Mexico* certain Laws and Conſtitutions made by his Majeſty at *Barcelona*, in the month of *December* in the year 1542. which were calculated according to the urgent neceſſity of Affairs to rectify the Diſorders that were then commited with ſo much Inſolence both againſt God and Man, to the utter deſtruction and deſolation of the New World. For the conſtituting of theſe wiſe Laws his Majeſty aſſembled a great

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/29/08

Book title: *La Araucana*
Author: Alonso de Ercilla y Zúñiga
Year of Publication: 1806. [vol. 2]
Place of Publication: Gotha.
Publisher: Steudel y Keil
ISBN number?: No [OCLC 09443198]
Content/Disciplinary area: Indians of South America Poetry; Spaniards Chile Poetry; Araucanía (Chile) History Poetry.

AVAILABILITY SEARCH

Source of scanned book:	Google Book Search.
Comments on condition/provenance of book:	Below average. It is hard to ascertain the condition of the book with a gray-scale scan set too dark. Either the book is damaged or the scanner set it too dark, resulting in some pages being unreadable.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy. There are three volumes for this particular edition and GBS has them all digitized. It also suggests them!
MARC record available?	Yes but had to go through WorldCat to send lending libraries MARC records.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata links to other volumes all available on Google Books. The only problem is that they are included in the category "Other editions", thus confusing their being volumes, not editions – therefore the other editions are mixed in with the volumes.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability / quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Above Average	Not in color but entirely legible.	Splotchy areas that comes with the territory of being gray-scale.	N/A
1, 10, 20	Average	Part of page 1 doubled up in the opposite side of the margin. See Appendix I.	N/A	N/A
9, 11, 13, 15, 17	Above average	Faint diagonal line running across the page. (scanner? From a previous fold?) – does not impact legibility.	The line is the faintest on 17.	N/A
9, 11, 13, 15, 17	Above average	Faint diagonal line running across the page. (scanner? From a previous fold?) – does not impact legibility.	The line is the faintest on 17.	N/A

30, 40	Above Average	N/A	N/A	N/A
50, 60	Average	Page 60 slightly blurry.	N/A	N/A
51	Average	Off-center and slightly blurry.	N/A	N/A
70, 80	Average	Page 80 contains several dark splotches that make a few words illegible.	N/A	N/A
90, 100	Above average	Page 100 is very discolored. Many words illegible. See Appendix II.	N/A	N/A
109, 110, 120	Poor	Page 109 appears to be torn, rendering several parts illegible. See Appendix II	The ink from the opposite side of the same page is also visible.	N/A
130, 140	Above Average	N/A	N/A	N/A
150, 160	Average	Page 160 is blurry. Legibility slighted impacted.	N/A	N/A
170, 180	Above Average	N/A	N/A	N/A
190, 200	Above Average	N/A	N/A	N/A
210, 220	Average	210 handwritten annotations visible – dashes? Does not impact legibility.	211 large black spot in upper left-hand corner – scanner error?	N/A
230, 240	Above Average	N/A	N/A	N/A
244	Poor	Large black spot in upper right-hand corner obscures several words.	N/A	N/A
250, 260	Above Average	N/A	N/A	N/A
276, 277	Poor	Scanner was apparently flipping from pages 274/275 to 276/277 and was caught half-way. See Appendix III.	Pages 276 and 277 are missing!	N/A
270, 280	Above Average	N/A	N/A	N/A
290, 300	Average	Page 290 has a small black splotch in the lower left, obscuring a couple of words.	N/A	N/A
310, 320	Average	Page 310 slightly blurry – still legible	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) Y / **N**

Word	Comments
N/A	**N/A**

IMAGE ASSESSMENT

If the original text includes N/A

graphics/illustrations, are they included in the electronic version?	
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	Smudges or discoloration are exacerbated with the gray-tone scan. Unclear at times if its scanner error, human error, or textual flaws.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF. No OCR recognition.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming inadequate because the original quality is poor.
Other notes	N/A

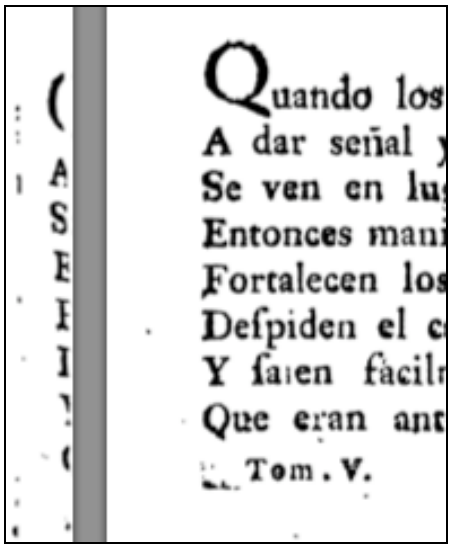
GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Average.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	No.
Ability to search?	No.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of	Yes.

contents?	
How successful is OCR recognition?	N/A
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	Scanning quality is inadequate. Cannot be searched.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No. The scanning is so blurry/indecipherable in some places it simply needs to be redone not repurposed.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	Google Books has all 3 volumes of this edition digitized but the metadata needs to be more specific on that topic.
Other notes	N/A

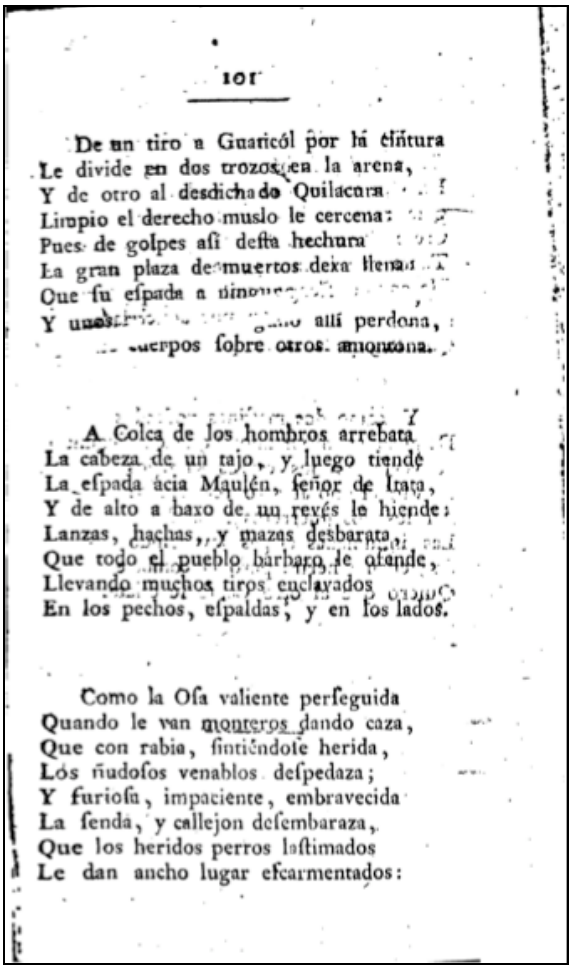
Book URL:
http://books.google.com/books?id=DfkAAAAAMAAJ&pg=PP1&dq=Araucana&as_brr=1&ei=NdVnSM3kC5y8jgGsmcyEBg
Date accessed: 6/29/08

Appendix I: Scanner Difficulties – Image has been partially doubled on the opposite side of the margin.



Quando los
A dar señal
Se ven en lu
Entonces mani
Fortalecen los
Deſpiden el c
Y ſalen facil
Que eran ant
Tom. V.

Appendix II: Damaged page.



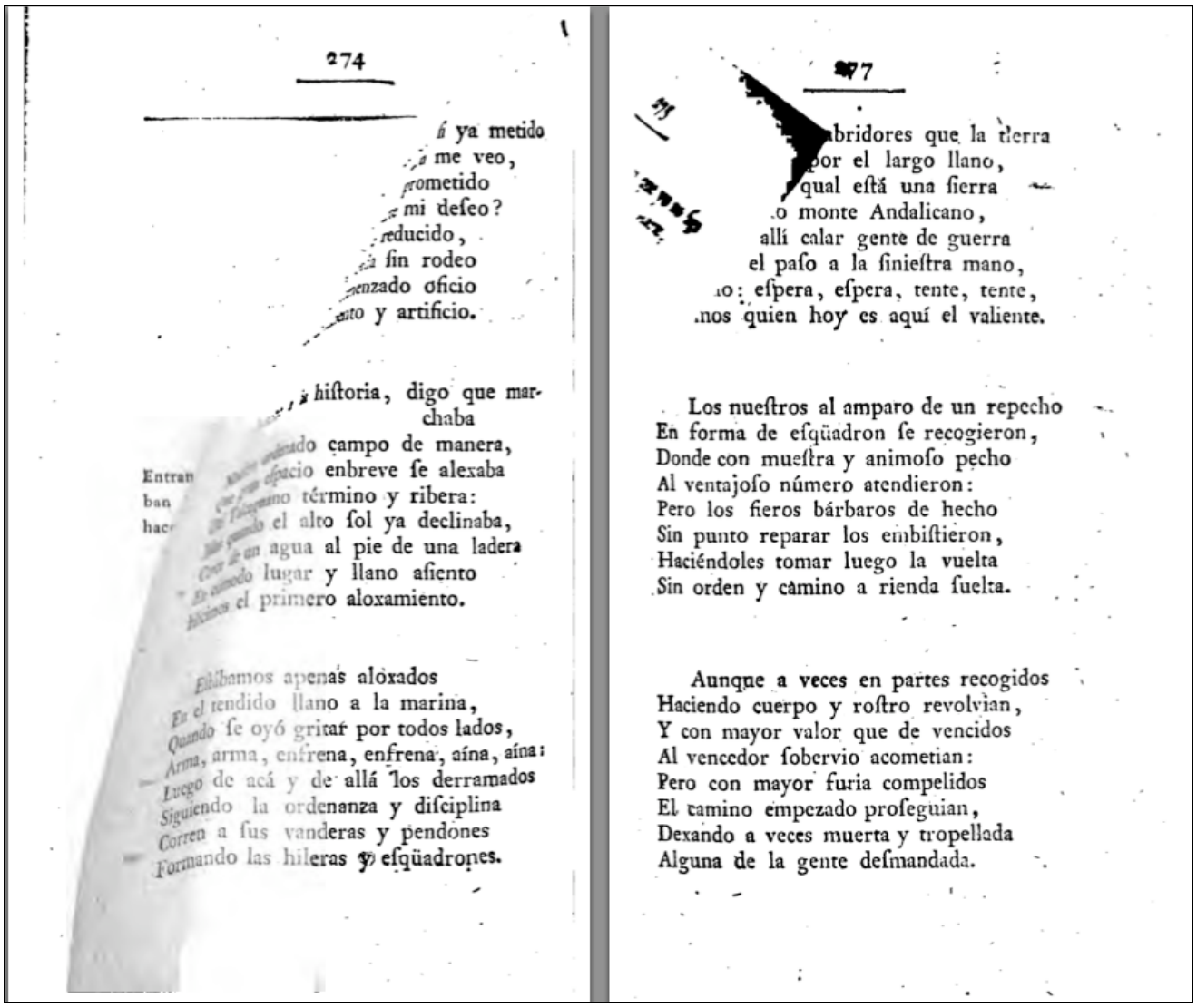
101

De un tiro a Guaricól por la cintura
Le divide en dos trozos en la arena,
Y de otro al desdichado Quilacura
Limpio el derecho muslo le cercena:
Pues de golpes asi desta hechura
La gran plaza de muertos deja llena
Que su espada a ninguno
Y unos allí perdona,
cuerpos sobre otros amontona.

A Colca de los hombros arrebata
La cabeza de un tajo, y luego tiende
La espada acia Maulén, señor de Irata,
Y de alto a baxo de un revés le hiende:
Lanzas, hachas, y mazas desbarata,
Que todo el pueblo bárbaro le ofende,
Llevando muchos tiros enclavados
En los pechos, espaldas, y en los lados.

Como la Osa valiente perseguida
Quando le van monteros dando caza,
Que con rabia, sintiéndose herida,
Los ñudosos venablos despedaza;
Y furiosa, impaciente, embravecida
La senda, y callejon desembaraza,
Que los heridos perros lastimados
Le dan ancho lugar escarmentados:

Appendix III: Caught in Mid-Page



274

…á ya metido
… me veo,
…rometido
… mi deſeo?
…reducido,
… ſin rodeo
…enzado oficio
…nto y artificio.

… hiſtoria, digo que mar-
chaba
… campo de manera,
… eſpacio enbreve ſe alexaba
… término y ribera:
… el alto ſol ya declinaba,
… agua al pie de una ladera
… lugar y llano aſiento
… el primero aloxamiento.

Entran
ban
hac

Eſtábamos apenas aloxados
En el tendido llano a la marina,
Quando ſe oyó gritar por todos lados,
Arma, arma, enfrena, enfrena, aína, aína:
Luego de acá y de allá los derramados
Siguiendo la ordenanza y diſciplina
Corren a ſus vanderas y pendones
Formando las hileras y eſqüadrones.

277

…bridores que la tierra
…or el largo llano,
… qual eſtá una fierra
…o monte Andalicano,
allí calar gente de guerra
el paſo a la ſinieſtra mano,
…o: eſpera, eſpera, tente, tente,
…nos quien hoy es aquí el valiente.

Los nueſtros al amparo de un repecho
En forma de eſqüadron ſe recogieron,
Donde con mueſtra y animoſo pecho
Al ventajoſo número atendieron:
Pero los fieros bárbaros de hecho
Sin punto reparar los embiſtieron,
Haciéndoles tomar luego la vuelta
Sin orden y camino a rienda ſuelta.

Aunque a veces en partes recogidos
Haciendo cuerpo y roſtro revolvian,
Y con mayor valor que de vencidos
Al vencedor ſobervio acometian:
Pero con mayor furia compelidos
El camino empezado proſeguian,
Dexando a veces muerta y tropellada
Alguna de la gente deſmandada.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/13/08

http://en.wikipedia.org/wiki/Lazarillo_de_Tormes

The Life of Lazarillo de Tormes and of His Fortunes and Adversities (La vida de Lazarillo de Tormes y de sus fortunas y adversidades) is a Spanish novella, published anonymously, in 1554 in Alcalá de Henares in Spain, and, in 1555, in Antwerp, Flanders, then under Spanish rule. At this time in Spain, there was a strong social trend towards racial and religious prejudice, and the utmost importance was placed on appearance.

Title: The Life of Lazarillo de Tormes: His Fortunes & Adversities
Author: Anonymous (ed. Clements Markham)
Year of Publication: 1908
Place of Publication: London
Publisher: Adam and Charles Black
ISBN number?: OCLC 1646105
Content/Disciplinary area: Spanish literature – 16th century; Picaresque literature.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Good. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	The metadata erroneously gives the authors names who are in fact the translator and editor. The text is actually anonymous and nowhere is that indicated in the metadata. The section divisions provided by Google metadata do link to the relevant chapter in the text; however, it leads to seemingly random pages within the chapter as opposed to the beginning of the same chapter in question. The online text does offer a table of contents that links to the relevant chapters.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) All of the pages are marked up by someone's notes, underlining, bracketing, etc. Why then choose this edition to scan?

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Cover	Very nicely scanned in color.	N/A	N/A	N/A
x, xx, xxx	Average	All of the pages are marked up by someone's notes.	N/A	N/A
1, 10	Average	""	N/A	N/A
20, 30	Average	""	N/A	N/A

60, 70	Average	Page 60 includes markups	N/A	N/A
80, 90	Average	Page 80 includes markups (page 90 excellent)	N/A	N/A
100, 110	Average	Page 100 markups (page 110 excellent)	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N (Search permitted in online version, but not pdf)
Lazarillo	Excellent
1908	Excellent
Introductory	Poor – some found most not.
Mendoza	Poor – some of the results highlighted within the text but not all included search results column
fermosa	Excellent
Quiñones	Unreadable – accents do not work (incorrect spelling, or by removing the accent is "Quinones" and is fully searchable.
Salamanca	Excellent

80, 90	Average	Page 80 includes markups (page 90 excellent)	N/A	N/A
100, 110	Average	Page 100 markups (page 110 excellent)	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N (Search permitted in online version, but not pdf)
Lazarillo	Excellent
1908	Excellent
Introductory	Poor – some found most not.
Mendoza	Poor – some of the results highlighted within the text but not all included search results column
fermosa	Excellent
Quiñones	Unreadable – accents do not work (incorrect spelling, or by removing the accent is "Quinones" and is fully searchable.
Salamanca	Excellent
47	Average – recognizes number in table of contents and index but not on page itself

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Numerous illustrations included – very good quality scan.
If there are no graphics/illustrations are	N/A

there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Average
Is the scan on a gray-scale or color scale?	Gray-scale
Quality – does it look like the real thing?	No
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Absolutely
Are the illustrations searchable?	Yes they are! Nice bonus.
Ability to print illustrations? (is there a watermark?)	Yes
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	No. Zooming not adequate due to it being in gray-scale.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average
Ease of access to full text	Easy
Ability to copy, quote, cite?	Yes – from plain text link available.
Ability to search?	Yes – but limited due to average recognition
Ability to download?	Yes
Ability to print?	Yes
Scanned book allow for consulting of contents?	Yes – but limited due to average recognition.
How successful is OCR recognition?	Average
Worthy as archive? (i.e., would this copy be acceptable as an archival copy	No.

of the book if physical copies were lost or destroyed?	
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No – OCR recognition needs a lot of work.
Issues with this book that would affect usefulness for scholars	Average recognition. Also can't search the pdf, only the online version. Metadata is also erroneous concerning authorship of the text. Plus, someone wrote notes, etc. all over the book. I would rather find a nicer edition.
Can this book be shared with other scholars/students?	No
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	OCR recognition must be fixed first.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	Google Books includes links to other editions available from their mass digitization project.
Other notes	N/A

Book URL: http://books.google.com/books?id=VzsMAAAAIAAJ&dq=lazarillo&lr=&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 8/13/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/13/08

Book title: Histoire d'un voyage faict en la terre du Brésil [vol 2]
Author: Léry, Jean de
Year of Publication: 1880
Place of Publication: Paris
Publisher: A. Lemerre
ISBN number? OCLC 12874844
Content/Disciplinary area: Brazil -- Description and travel; Indians of South America – Brazil; Tupi language.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Average. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	It took a couple of refined searches to find the text. Since it is a keyword search, I couldn't search by author and thus had too many results. Limited by language (French) I was finally able to find the full text.
MARC record available?	Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Table of contents excellent – it links to the chapter in question; furthermore, in the online version of the text (not the pdf), the table of contents also link directly to the chapter in question.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
1, 10	Above average	On page one, inscriptions – do not affect legibility	N/A	N/A
15	Poor	Left margin didn't get through – affects legibility	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
36, 37	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average	N/A	N/A	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word		Comment
	(Is OCR search permitted) Y / **N**	
N/A	**N/A**	

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Not exactly illustrations per se, but the binding is in color, also the interior of the two book flaps, showing the marbleized paper so typical to nineteenth century bindings.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with no recognition
Is the scan on a gray-scale or color scale?	Text is gray-scale but covers and interior flaps are in color.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	The Ex Libris is always interesting and the marbleized paper – although not an illustration, are visual elements that most definitely would be useful to scholars. See Appendix I.
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	Yes
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	None.
Other notes	

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Above average
Ease of access to full text	Easy
Ability to copy, quote, cite?	Not available. However, I did try a Google feature that allows you to reproduce a selected portion of the screen – kind of like the snap shots that Missy and I have been using all along. See Appendix II.
Ability to search?	No
Ability to download?	Yes
Ability to print?	Yes
Scanned book allow for consulting of contents?	Yes
How successful is OCR recognition?	None
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No – one page is unreadable due to error.
Issues with this book that would affect usefulness for scholars	See above.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No – unless errors are fixed.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	

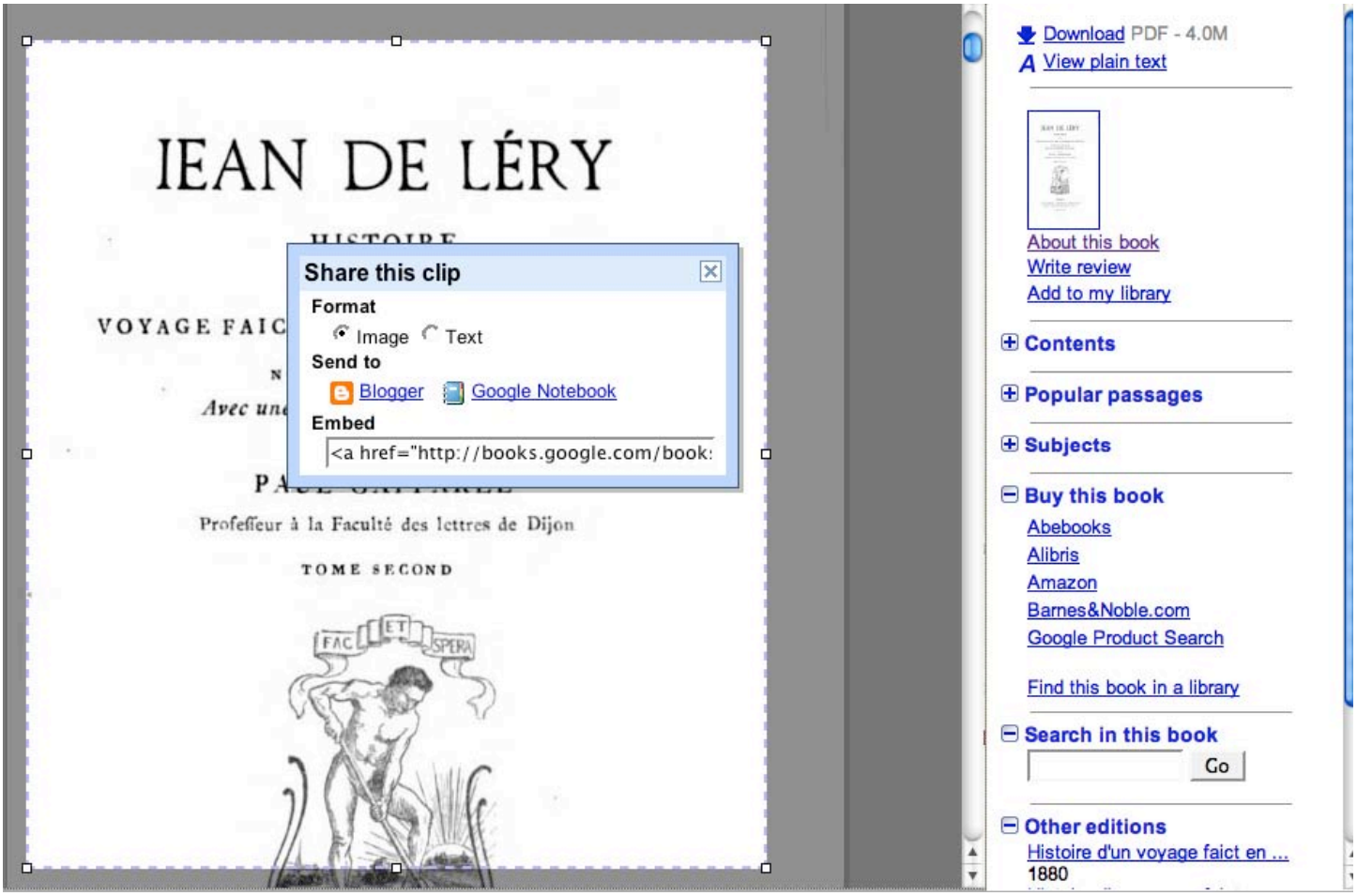
Book URL:
http://books.google.com/books?id=ybgeAAAAMAAJ&dq=jean+de+lery+brazil+bry&lr=&num=30&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 7/13/08

Appendix I – Google uses color!

Appendix II – Say Cheese

1. This is an example of a text that has no highlighting options but it does have a Google feature that allows you to take a "snap shot" of any part of the text and then open it in a program called "Notebook".

2. Below are the steps involved



3. You then have the option of saving the selection as an image or as a text. Both will link you straight back to the original selection chosen.

Conclusion – nice feature if you can't cut and paste from the pdf.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/29/08

http://en.wikipedia.org/wiki/Lugones

Leopoldo Lugones Argüello (13 June 1874 - 18 February 1938) was an Argentine writer and journalist. Lugones was the leading Argentine exponent of the Latin American literary current known as Modernismo. This was a form of Parnassianism influenced by Symbolism. He was also the author of the incredibly dense and rich novel La Guerra Gaucha (1905). He was an impassioned journalist, polemicist and public speaker who at first was a Socialist, later a conservative/traditionalist and finally a supporter of Fascism.

Book title: El imperio jesuítico; ensayo histórico
Author: Lugones, Leopoldo
Year of Publication: 1907 (2nd edition)
Place of Publication: Buenos Aires
Publisher: Arnoldo Moen y hermano
ISBN number: OCLC 3060386
Content/Disciplinary area: Jesuits; Paraguay -- History To 1811

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Above average. The scan is a bit dark – probably the actual book is darker toned. However, it calls into question how much a scan of a book can/should be altered. In other words, should this scan have been lightened?
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Available metadata is adequate but this text is not one of Lugones's better known works and no other ones are available (in full text) on Open Text Access or Google Books.
MARC record available?	No.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata is accurate and somewhat useful. No table of contents available however.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	other?	other?
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A

200, 210	Excellent	N/A	N/A	N/A
220, 230	Excellent	N/A	N/A	N/A
240, 250	Excellent	N/A	N/A	N/A
260, 270	Excellent	N/A	N/A	N/A
280, 290	Excellent	N/A	N/A	N/A
300	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N
imperio	Excellent (however not searchable on pages containing illustrations. Not meaning a captions, but the actual text within the illustration itself.
edición	Excellent (can read accents) (also reads footnote text)
1907	Excellent
año	Excellent
"Garcilaso Inca"	Excellent. The search engine located the one instance of the use of "Garcilaso Inca" (author) but also found every instance of "inca". A search phrase must be put in quotation marks in order to locate the string of words.
"Entre 1724 y 1767 habían"	Excellent
"ción de la Patagonia"	Average. The correct phrase in the book is "Descripción de la Patagonia". However "Descripción" is hyphenated at the end of the sentence and reads "Descrip-ción". Words that are hyphenated at the end of a sentence are not searchable (very interesting!) because the two sections are not treated as one word, but two. See Appendix I.
370.000	Excellent
"río Miriñay"	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes. Some are maps and others illustrations – presumably lithographs.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	None.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	OCR has excellent recognition, especially regarding the ability to search for accented words and characters.
Is the scan on a gray-scale or color scale?	Color
Quality – does it look like the real thing?	The scan is a bit dark – but that could very well be due to the appearance of the txt itself.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Absolutely. This is the first time that I have encountered maps in a digitized text and they seem to be of good quality.
Are the illustrations searchable?	No.
Ability to print illustrations? (is there a	Yes you can print the entire pdf.

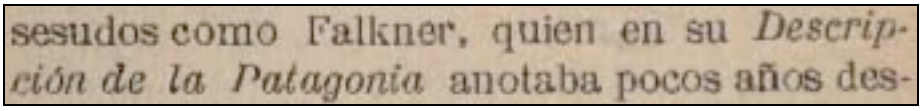
watermark?)	
Are the images downloadable? (file size, pixel quality)	Yes. The pixel quality seems to be above average.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming is excellent.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy
Overall quality of scan of this book	Excellent
Ease of access to full text	Easy
Ability to copy, quote, cite?	Yes
Ability to search?	Yes
Ability to download?	Yes
Ability to print?	Yes
Scanned book allow for consulting of contents?	No table of contents provided.
How successful is OCR recognition?	Excellent
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes
Issues with this book that would affect usefulness for scholars	None that I could think of.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	There are other editions of this book that are not included in the Open Access Text Project.
Other notes	N/A

Book URL: http://www.archive.org/details/elimperiojesut00lugouoft
Date accessed: 8/29/08

Appendix I: Hyphenated words cannot be searched


sesudos como Falkner, quien en su *Descrip-*
ción de la Patagonia anotaba pocos años des-

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/7/08

Sloppiest Google metadata that I have come across so far!

http://en.wikipedia.org/wiki/Joaquim_Maria_Machado_de_Assis

Joaquim Maria Machado de Assis, often known as Machado de Assis, Machado, or Bruxo do Cosme Velho, (June 21, 1839, Rio de Janeiro—September 29, 1908, Rio de Janeiro) was a Brazilian novelist, poet and short-story writer. He is widely regarded as the most important writer of Brazilian literature.[1][2][3] However, he did not gain widespread popularity outside Brazil in his own lifetime.

Machado's works had a great influence on Brazilian literary schools of the late 19th century and 20th century. José Saramago, Carlos Fuentes, Susan Sontag and Harold Bloom are among his admirers[4] and Bloom calls him "the supreme black literary artist to date."[5]

Book title: Memórias póstumas de Brás Cubas
Author: Assis, Machado de
Year of Publication: 1961 (originally published in 1881)
Place of Publication: Rio de Janeiro
Publisher: Tecnoprint Gráfica
ISBN number? OCLC 21313729
Content/Disciplinary area: Novel – Brazil; Machado de Assis.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Average. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Metadata sloppy! I originally wanted to look at this edition because Google metadata said the date of publication was 1880. Upon opening the file, it was very clear that such a modern cover could in no way be from 1880. 1880, in fact, is the address of the publisher. This text is from 1961. It is full text so I am going to look at it anyway. This problem probably arose because whoever was entering data never double checked with an OPAC and instead looked for something that looked like a date.
MARC record available?	Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Table of contents horrible – the metadata again is solely based on keywords that in turn, auto-generate a table of contents.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Cover, inner flap	Excellent (in color)	N/A	N/A	N/A
20, 30	Above average	Page 20 - possible pencil markings – does not affect legibility	Page 30 - possible pencil markings – does not affect legibility	N/A

40, 50	Poor	Page 50 very dark and blurry – legibility affected	N/A	N/A
60, 70	Above average	Page 60 - possible pencil markings – does not affect legibility	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	Page 100 - possible pencil markings – does not affect legibility	N/A	N/A
120, 130	Above average	120 dark but still legible	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	Page 190 - possible pencil markings – does not affect legibility	N/A	N/A
200, 210	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) Y / **N**
	No OCR search in the PDF but yes in the online version
Machado	Excellent
Brás	Excellent (works with accent)
Memórias Póstumas	Excellent
não	Excellent
ânsias	Excellent
67	Unreadable – doesn't recognize page numbers
?	Unreadable
Brasil	Readable

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Not exactly illustrations per se, but the cover is in color, also the interior of the two book flaps, showing that once again, there is no way this book was published in the late 19th c.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with no recognition. Online version is readable however with excellent recognition.
Is the scan on a gray-scale or color scale?	Text is gray-scale but covers and interior flaps are in color.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art	N/A

historians, visual culture researchers?)	
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	Yes
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	None.
Other notes	It is unfortunate that in order to assess the true value of such a digitized book one must have to look at both the pdf and the online version. There are too many inconsistencies between the two as stands.

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Averagef
Ease of access to full text	Easy
Ability to copy, quote, cite?	Not available. However, I did try a Google feature that allows you to reproduce a selected portion of the screen – kind of like the snap shots that Missy and I have been using all along. See Lery form.
Ability to search?	No
Ability to download?	Yes
Ability to print?	Yes
Scanned book allow for consulting of contents?	Yes
How successful is OCR recognition?	Successful online but none available in the pdf.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No – bad scan.
Issues with this book that would affect usefulness for scholars	See above.
Can this book be shared with other scholars/students?	No
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No – unless errors are fixed.

Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	

Book URL:
http://books.google.com/books?id=6J1UAAAAMAAJ&printsec=frontcover&dq=machado+de+assis&lr=lang_pt&num=30&as_brr=1&ei=8wZ7SJSeKYHsiQGgmoGeAQ
Date accessed: 7/7/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/05/08

Book title: *Count Lucanor: Or, The Fifty Pleasant Stories of Patronio*
Author: Manuel, Juan. (Trans. James York)
Year of Publication: 1899 (1868)
Place of Publication: London
Publisher: Gibbons & Company, Limited
ISBN number?: OCLC 2143948
Content/Disciplinary area: Medieval Spanish literature.

AVAILABILITY SEARCH

Source of scanned book:	Google Book Search.
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata for other editions but only English ones available on Google Book Search. Nothing for other translations.
MARC record available?	Yes. Available through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Very nice. The contents section follows book chapters. When searching the online version, the table of contents links to individual chapters.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Title page	Above average	N/A	N/A	N/A
x	Above average	N/A	N/A	N/A
1	Above average	N/A	N/A	N/A
10	Average	Black circular mark that affects legibility	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	120 has some markings (pencil?) – does not affect legibility	N/A	N/A
140, 150	Average	150 has some markings (pencil?) – does not affect legibility	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A

240	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N **
	ONLY WITHIN ONLINE VERSION – NOT PDF
Lucanor	Excellent.
Contents	Excellent.
á, é, í,, ó, ú, ñ	Unreadable.
30	Recognizable within the text but not the margins – page numbers not recognized.
XXXIII; xxxiii	Excellent.
?	Unreadable.
Alhaquima	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF. Search recognition average in online version but not available in pdf.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming average.
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Average search capabilities in online version. No search availability in pdf version.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Very nice. The contents section follows book chapters. When searching the online version, the table of contents links to individual chapters.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Cannot search for accented words. Furthermore can't search within pdf, only online version
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://books.google.com/books?id=CDyQEKxPquAC&dq=conde+lucanor&source=gbs_summary_s&cad=0
Date accessed: 8/05/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/11/08

Book title: Spanish Sketches
Author: Piddington, A. B.
Year of Publication: 1916
Place of Publication: London, New York
Publisher: H. Milford, Oxford University Press
ISBN number? OCLC 2975449
Content/Disciplinary area: Spain -- Description and travel; Velázquez, Diego; Canalejas y Méndez, José.

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	I was actually trying to find works by Hernan Cortés and came across this text – it includes numerous visuals so I want to look more carefully at the digitization quality of a text mostly comprised of illustrations.
MARC record available?	Yes. Available through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Average. Table of contents incomplete and doesn't link to corresponding text.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Covers (front, back)	Excellent	N/A	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	(Is OCR search permitted) Y / N **Yes**	**Comment**
Piddington	Excellent	

Spanish	Excellent
Illustrations	Excellent
José	Not readable with accent. Jose without accent searchable.
36	Excellent
Velazquez	Excellent
vii	Excellent
Señor	Not readable with accent. Senor searchable.
Canalejas	Excellent.
Árabe	Not readable with accent. Arabe searchable.
?	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes. About 30 black and white photographs.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with above average recognition. No recognition of foreign accents/characters.
Is the scan on a gray-scale or color scale?	Color (photographs originally printed in black and white)
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes.
Are the illustrations searchable?	Captions searchable.
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	Yes.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming good. The pdf is very high quality but still not the best for detailed image analysis. Upon first seeing the images, I had to zoom to see if they were lithographs or photographs. Captions ensured that they were in fact photographs. Granted, the quality could reflect the original print text but its hard to tell. I also tried to find out what the original dpi (dots per inch) was when the text was scanned but that information is not available. See Appendix I and II.
Other notes	N/A

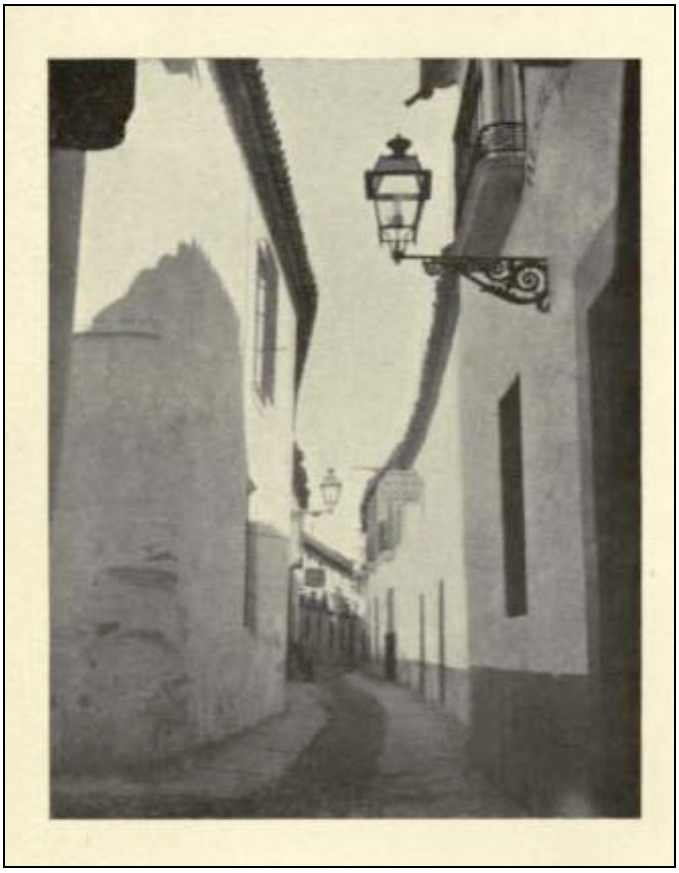
GENERAL OBSERVATIONS

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

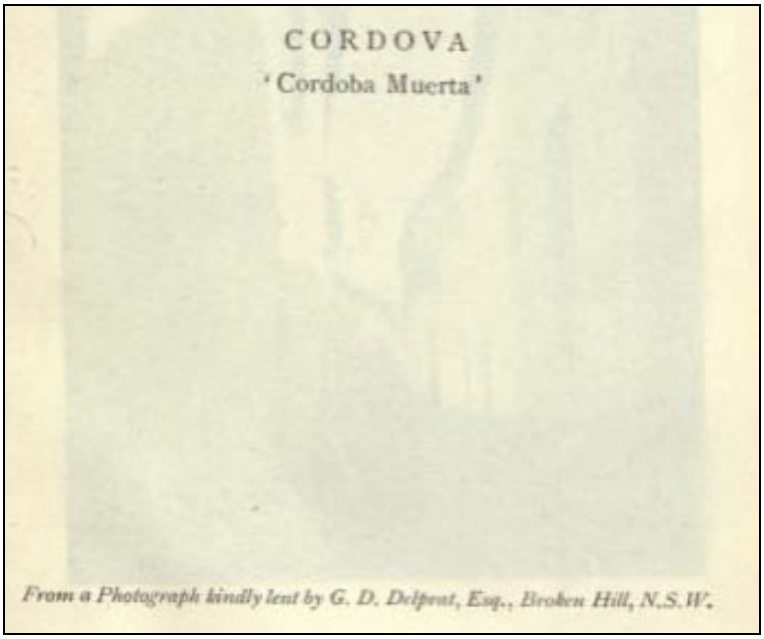
Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average. No foreign words or characters recognized but typing in the same word without the accents is searchable.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	DPI is high but should be higher for the images. However, the same images could be found elsewhere – in other words, not unique to text.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/spanishsketchesb00piddrich
Date accessed: 7/10/08

Appendix I



Appendix II – Caption



CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/17/08

http://en.wikipedia.org/wiki/Joel_Roberts_Poinsett

Joel Roberts Poinsett (March 2, 1779 – December 12, 1851) was a physician, botanist and American statesman. He was a member of the United States House of Representatives, the first United States Minister to Mexico (the United States did not appoint ambassadors until 1896), a U.S. Secretary of War under Martin Van Buren and a cofounder of National Institute for the Promotion of Science and the Useful Arts (a predecessor of the Smithsonian Institution), as well as the namesake of Poinsett County, Arkansas, the historic Poinsett Bridge in Greenville County, South Carolina, Poinsett State Park in Sumter County, SC, and the poinsettia, a popular Christmas flower.

Book title: Manifiesto de los principios políticos del Escmo. Sr. d. J.R. Poinsett
Author: Lorenzo de Zavala; J. R. Poinsett
Year of Publication: 1828
Place of Publication: Mexico
Publisher: Imprenta del Correo
ISBN number?: OCLC 12890024
Content/Disciplinary area: Poinsett, Joel Roberts; Latin America -- Foreign relations -- United States; United States -- Foreign relations -- 1779-1851

AVAILABILITY SEARCH

Source of scanned book:	Open-Text Archive.
Comments on condition/provenance of book:	Good. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Searched for all "manifiestos" [manifesto] in Spanish and came across this text – it intrigued me due to the poinsettia being named after the author!
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Good. It is a short text so there is no table of contents but otherwise the metadata is sufficient.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
1, 10	Above average	N/A	N/A	N/A
20	Above average	N/A	N/A	N/A
24-53	Blank	The text ends on page 24 but it is still odd that there about 30 blank pages following it – printing error?	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Word	Comment
	(Is OCR search permitted) **Y** / N (Available in both online and pdf versions)
	I found the OCR recognition to be excellent in this text. Many times when searching for individual letters, the recognition does not function; that is, I would have to search for the complete word in order to locate it. The recognition here permits searching for parts of words and the letters within them.
manifiesto	Not found.
Poinssett	Excellent.
principios	Excellent.
México	Excellent.
Mexico	Not found. – Good because it means that only correct, accented version functions.
año	Excellent.
7	Excellent.
V. S.	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Yes.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent OCR recognition.
Is the scan on a gray-scale or color scale?	The online version is in color and the pdf is gray-scale.
Quality – does it look like the real thing?	Online yes; pdf no.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes.
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming average due to pdf being gray-scale.
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes (although I am not sure why there are 30 blank pages at the end of the text)
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes but the pdf should also be in color.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	30 blank pages – is it printing error that the scan faithfully reproduces or a scanning error?
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/manifiestodelosp00zavarich
Date accessed: 8/17/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/28/08

[Today I searched for the *Popul Vuh* and included notes on its provenance below. My hypothesis that its contentious and confusing origin will create problems any existing metadata.]

[http://www.newberry.org/collections/PopolVuh.html]

The Popul Vuh, which has been translated as Book of the Council, Book of the Community, Book of the People, and The Sacred Book, is the creation account of the Quiché Mayan people. It contains stories of the cosmologies, origins, traditions and spiritual history of the Mayan people. It is considered by many Mayans as their equivalent to the Christian Bible and is held in deep reverence by them.

The Newberry Library's manuscript of the Popol Vuh is one of the most widely known and possibly the earliest surviving copy. Quiché nobility probably wrote the original manuscript of the Popul Vuh in the mid 16th-century, in the Quiché language, using Latin orthography. The Newberry's Popol Vuh was most likely copied from this original manuscript (now lost) in 1701-03, in the Guatemalan town of Chichicastenango, by Dominican Father Francisco Ximenez. His copy includes the Quiché text and a Spanish translation in side-by-side columns.

In addition to the Popul Vuh, the manuscript also contains a Cakchikel-Quiché-Tzutuhil grammar, Christian devotional instructions, and answers to doctrinal questions and other material by Ximenez.

Book title: *Las historias del origen de los Indios de esta provincia de Guatemala, traducidas de la lengua quiché al castellano …*
Author: Ximenez, Francisco.
Year of Publication: 1857
Place of Publication: Vienna
Publisher: Casa de Carlos Gerold E Hijo [Carlos Gerold and Son]
ISBN number?: No [OCLC number: 21437082]
Content/Disciplinary area: Popol vuh. Quiché mythology. Quiché Indians -- Religion.

AVAILABILITY SEARCH

Source of scanned book:	Google Book Search.
Comments on condition/provenance of book:	Average. It is hard to ascertain the condition of the book with a gray-scale scan. There are also numerous quality concerns with the scanning that I will explore below.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Very difficult. There are different spelling variations of the title that provide different search results. Also the title at times is not what it is commonly called today but instead longer titles that depend on the particular author. The metadata is wretched however, listing the author as *Popol-Vuh* itself [which is not correct], and it doesn't include the publisher information. See Appendix I. It was also difficult finding this particular edition in Google Books. It was only when I searched for "Ximenez, Francisco popol-vuh 1857" that I found the book – thus trying an alternate spelling variation for the title.
MARC record available?	No.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	The metadata is below average. It does not link to any external information or suggest the numerous other editions that exist of the work that exist even within Google Books. Again see Appendix I. It does include a searchable table of contents for each chapter.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

search paradigm settings (example: every 10 double-pages) Page(s):	readability / quality of scan	anomalies?	other?	other?
Title page	Average	Can't make out university stamp/embossment.	N/A	N/A
ii, iv, x	Poor	On pages ii and iv there are color sets of gloved fingers. See Appendix II.	N/A	N/A
1, 10, 20	Above average	I classify it as such because it is not in color.	N/A	N/A
29, 30, 40	Poor	Page 29 barely legible (book moved slightly during scanning?) Pages 30 and 40 fine.	Page 39 barely legible (must be human error if the left leafs are consistently blurred.)	N/A
50, 59, 60	Poor	Pages 60 and 70 fine. Page 59 also blurry. See above.	N/A	N/A
70, 80	Above average	N/A	N/A	N/A
90, 99, 100	Poor	Page 99 illegible. Blurry. Pages 90 and 100 are fine.	N/A	N/A
110, 119, 120	Excellent	Page 119 illegible and blurry. Pages 110 and 120 are fine.	N/A	N/A
125	Average	Page 125 has a large black mark on the bottom side of the page.	N/A	N/A
129, 130, 140	Above average	Page 129 blurry. Pages 130 and 140 fine.	N/A	N/A
137, 149	Average	Pages 137 and 149 have a large black mark on the bottom side of the page.	N/A	N/A
150, 160	Poor.	Page 150 fine. Page 159 has large black mark on the bottom of the page.	Page 160 curves – perhaps the page wasn't lying flat on the scanner? – making right-hand side difficult to read.	N/A
170, 180	Above average	N/A	N/A	N/A
190, 200	Average	190 fine. Page 200, making right-hand side difficult to read.	N/A	N/A
210	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) Y / **N**

Word	Comments
N/A	**N/A**

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A

If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	There are several black dots that appear throughout the book in the same exact spot on certain pages. Could be wormholes but have no way of knowing with the gray-scale scan.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	OCR. Average recognition due to the fact that it doesn't recognize any of the foreign letters or symbols.
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming adequate.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Not easy. There are no links to external information on the text so I went to WorldCat to find the particular edition. And they got the author information plus, in addition to not including the publisher information.
Overall quality of scan of this book	Average.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	No.
Ability to search?	No.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of	Yes.

contents?	
How successful is OCR recognition?	N/A
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	Too many human errors in the scanning process that results in poor quality.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No. The scanning is so blurry in some places it simply needs to be redone not repurposed.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No. But Google Books makes no mention of other editions.
Other notes	N/A

Book URL: http://books.google.com/books?id=yRjj8ca8okcC&pg=PR1&dq=Ximenez,+Francisco+popol-vuh+1857&lr=&as_brr=1&ei=1ZBnSOSRG4uujgGGk_jNBA

Date accessed: 6/27/08

Appendix I – Bad metadata

By Popol-vuh

Translated by Francisco Ximenez

Published 1857

Original from Oxford University

Digitized Apr 24, 2006

Add to my library

Write review

Popul-vuh is the name of the word, not the author

Francisco Ximenez should be listed as the author because he was the one that wrote down the earliest known copy of these oral histories.

Appendix II – Messy Scanning (Also random point: I had no idea that such "finger gloves" existed!)

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/6/08

Book title: History of the Conquest of Mexico: With a Preliminary View of the Ancient Mexican Civilization, and the Life of the Conqueror, Hernando Cortez.
Author: William Hickling Prescott.
Year of Publication: 1886
Place of Publication: New York
Publisher: J.B. Alden
ISBN number?: OCLC 5536269
Content/Disciplinary area: Mexico - History - Conquest, 1519-1540; Cortés, Hernán.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata confusing. At the top of the entry, it states that volumes 2-3 are published by a different publisher: "Publisher varies, v.2-3 published by Harper and Brothers, New York, [1861?]". But on the title page of the book itself, it clearly states that this particular copy is volume 2, published in 1886.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Table of contents in Google Books is also entirely confusing because it labels books and chapters as all being chapters. PDF table of contents actually links to relevant chapters.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	Other?	Other?
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A
200, 210	Excellent	N/A	N/A	N/A
220, 230	Excellent	N/A	N/A	N/A
240, 250	Excellent	N/A	N/A	N/A
260, 270	Excellent	N/A	N/A	N/A
280, 290	Excellent	N/A	N/A	N/A
300, 310	Excellent	N/A	N/A	N/A
320, 330	Excellent	N/A	N/A	N/A
340, 350	Excellent	N/A	N/A	N/A

360, 370	Excellent	N/A	N/A	N/A
380, 390	Excellent	N/A	N/A	N/A
400, 410	Excellent	N/A	N/A	N/A
420, 430	Excellent	N/A	N/A	N/A
440, 450	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N Online source readable, not pdf.
Montezuma	Excellent
Mexico	Excellent
Cortés	Average – not all instances show up (same situation with Cortes without accent)
1886	Unreadable
333	Excellent
Aztecs	Excellent.
Marina	Excellent
Oviedo	Excellent
Bernal Diaz	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Several illustrations of Montezuma and a map of Mexico. The map is on two pages however and one side is a very different tone than the other, making it too light for zooming to be useful.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Excellent recognition
Is the scan on a gray-scale or color scale?	Gray-scale
Quality – does it look like the real thing?	No
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes
Are the illustrations searchable?	No.
Ability to print illustrations? (is there a watermark?)	Yes
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	No.

Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No.
Ability to search?	Yes – online version only.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Metadata poor.
How successful is OCR recognition?	Excellent.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	Metadata poor.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	Metadata appears to erroneously indicate that volumes 2 and 3 are published by a different publisher, when in fact this is volume 2.
Other notes	N/A

Book URL:
http://books.google.com/books?id=WAYOAAAAIAAJ&dq=hernan+cortes&lr=&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 9/6/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/16/08

PDF couldn't open due to file error. Also the file size was huge at a whopping 17.5M

http://en.wikipedia.org/wiki/%C3%81ngel_de_Saavedra%2C_Duke_of_Rivas

Don Ángel de Saavedra y Ramírez de Baquedano, Duke of Rivas (Spanish: Ángel de Saavedra y Ramírez de Baquedano, Duque de Rivas) (March 19, 1791 - June 22, 1865), was a Spanish poet, dramatist and politician born in Córdoba. He is best known for his play Don Álvaro; o, La fuerza del sino (1835), the first romantic success in the Spanish theater.

Book title: Obras completas de D. Ángel de Saavedra
Author: Angel de Saavedra Rivas (Duque de Rivas)
Year of Publication: 1904 (Volume 7)
Place of Publication: Madrid
Publisher: Est. tipográfico "Sucesores de Rivadeneyra"
ISBN number?: OCLC 175165559
Content/Disciplinary area: 19th c. Spanish Literature

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Good. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	No. All copies on WorldCat at German libraries.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Metadata missing place of publication. No table of contents or otherwise useful guide to book. Links to 3 volumes available in full text on Google Books, including this one. There are at least seven volumes but not all are listed on the search results.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)	I only searched the online version because the pdf file wouldn't open. ALL the pages have fingers.

Page(s):	readability/ quality of scan	anomalies?	other?	other?
Cover	Average	Warped.	N/A	N/A
10	Average	Finger. Legibility not affected.	N/A	N/A
16, 17	Below average.	16 completely blurred. No legibility.	17 warped. Still legible.	N/A
20, 30	Average.	30 has a finger. Legibility not affected.	N/A	N/A
40, 50	Above average.	Finger	N/A	N/A

60, 70	Above average.	Finger	N/A	N/A
80, 90	Above average.	Finger	N/A	N/A
100, 110	Above average.	Finger	N/A	N/A
120, 130	Above average.	Finger	N/A	N/A
140, 150	Above average.	Finger	N/A	N/A
160, 170	Above average.	Finger	N/A	N/A
180, 190	Above average.	Finger	N/A	N/A
200, 210	Above average.	Finger	N/A	N/A
220, 230	Above average.	Finger	N/A	N/A
240, 250	Above average.	Finger	N/A	N/A
257	Average.	Weird discoloration. See Appendix I.	N/A	N/A
260, 270	Above average.	Finger	N/A	N/A
280, 290	Above average.	Finger	N/A	N/A
300, 310	Above average.	Finger	N/A	N/A
320, 330	Above average.	Finger	N/A	N/A
340, 350	Above average.	Finger	N/A	N/A
353	Average.	Weird discoloration.	N/A	N/A
260, 370	Above average.	Finger	N/A	N/A
380, 390	Above average.	Finger	N/A	N/A
384	Average	Blurry	N/A	N/A
400, 410	Above average.	Finger	N/A	N/A
420, 430	Above average.	Finger	N/A	N/A
440, 450	Above average.	Finger	N/A	N/A
460, 470	Above average.	470 warped.	N/A	N/A
480, 490	Above average.	Finger	N/A	N/A
500	Above average.	Finger	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N **Online version only.**
Rivas	Average. Search results don't include all of the highlighted hits. Other pages that are blurry are not searchable.
Saavedra	Below average.
Solaces	Average. Search results don't include all of the highlighted hits.
Doña	Excellent.
229	Not readable.
Sí vs. si	Excellent. Accents readable and distinguishable.
Jornada	Excellent.
Alarcón	Excellent.
Escena	Excellent.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the	OCR average. But only available in the online version – pdf not available because it didn't work.

recognition?	
Is the scan on a gray-scale or color scale?	Gray-scale.
Quality – does it look like the real thing?	No.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	No.
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	No zoom on online version.
Other notes	N/A

GENERAL OBSERVATIONS

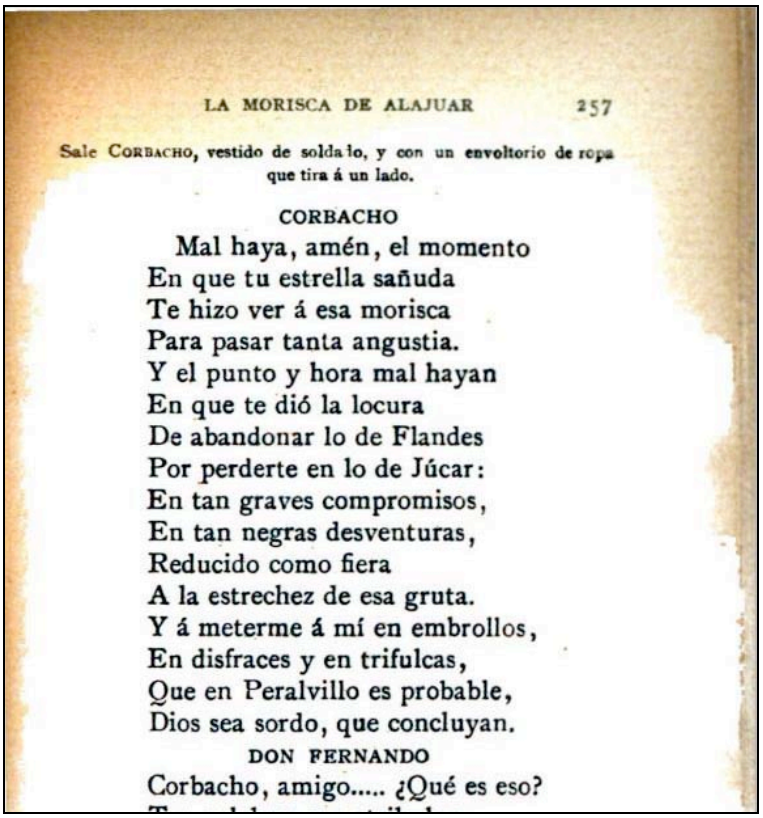
Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Poor. Fingers on every page.
Ease of access to full text	Difficult since the pdf didn't work and the online version allows no zooming capabilities.
Ability to copy, quote, cite?	N/A
Ability to search?	Yes.
Ability to download?	In theory yes but pdf didn't work.
Ability to print?	With difficulty. Page by page in online version.
Scanned book allow for consulting of contents?	No – pdf bad.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Issues with this book that would affect usefulness for scholars	Horrible scan job. Average recognition.
Can this book be shared with other scholars/students?	No.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	At least six other volumes – Google Books doesn't have all them digitized in full.
Other notes	N/A

Book URL: http://books.google.com/books?id=qdgCAAAAYAAJ&dq=editions:0Qwv9bsRm6iP-mtEnBkWaj&source=gbs_summary_s&cad=0
Date accessed: 7/16/08

Appendix I – Weird discoloration (Partially in color – remainder of book is on gray-scale)



LA MORISCA DE ALAJUAR 257

Sale CORBACHO, vestido de soldado, y con un envoltorio de ropa que tira á un lado.

CORBACHO

Mal haya, amén, el momento
En que tu estrella sañuda
Te hizo ver á esa morisca
Para pasar tanta angustia.
Y el punto y hora mal hayan
En que te dió la locura
De abandonar lo de Flandes
Por perderte en lo de Júcar:
En tan graves compromisos,
En tan negras desventuras,
Reducido como fiera
A la estrechez de esa gruta.
Y á meterme á mí en embrollos,
En disfraces y en trifulcas,
Que en Peralvillo es probable,
Dios sea sordo, que concluyan.

DON FERNANDO

Corbacho, amigo..... ¿Qué es eso?

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/29/08

Book title: *Ariel*
Author: José Enrique Rodó
Year of Publication: 1922 [original date of publication is 1900 in Spanish]
Place of Publication: Boston and New York.
Publisher: Houghton, Mifflin Co.
ISBN number?: No [OCLC 491857]
Content/Disciplinary area: Social ethics; Democracy.

AVAILABILITY SEARCH

Source of scanned book:	Open-Access Text and Google Book Search have the same edition! But since the GBS version cannot be searched, I am going to look at the lovely color searchable version from OAT.
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Metadata is poor and does not link to any external information. The only keyword is "William Shakespeare" – which makes sense because the work is named for a character in *The Tempest* – but nothing linking it to social ethics or democracy-related texts.
MARC record available?	Yes but had to go through WorldCat to send lending libraries MARC records.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	This work is a single essay and the edition includes an introduction by the editor but no table of contents. Interestingly, the GBS version created a table of contents for metadata purposes. It makes no sense however and seemingly creates "sections" based on keywords as determined by them. See Appendix I while keeping in mind that there are no chapters in this essay. [http://books.google.com/books?id=7xaX4NjgECwC&dq=Ariel&lr=lang_es&as_brr=1&source=gbs_summary_s&cad=0]

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

I would like to point out here that this book has gone through several rounds of people taking notes in it. Writing in margins, underlining, circling words, etc. But all of that doesn't impact the quality of the scan – strangely, I personally like seeing that all of the defacing is perfectly legible. It points to the quality of the scan but also that nobody tried to clean up the margins (which I don't know would even work in a color scan like this).

Page(s):	readability / quality of scan	anomalies?	other?	other?
Title page	Excellent	Embossment visible.	N/A	N/A
1, 10, 20	Excellent.	A student (professor?) went through a lot of this book and marked it up with a pencil. Everything is still perfectly	N/A	N/A
1, 10, 20	Excellent.	A student (professor?) went through a lot of this book and marked it up with a pencil. Everything is still perfectly legible.	N/A	N/A

30, 40	Excellent	Page 30 - Another individual then used a red pen to underline portions of the text.	Page 40 – blue pen	N/A
50, 60	Excellent	Page 60 has handwritten annotations in blue ink.	N/A	N/A
70, 71, 80	Average	Page 70 – red ink	Page 71 – the centerfold obscures the first letter of every sentence.	Page 80 – pencil annotations
90, 100	Excellent	Page 90 – pencil markings	Page 100 – red ink	N/A
110, 120	Excellent	Page 120 – pencil	The ink from the opposite side of the same page is also visible.	N/A
127	Above average	Folded tab in upper right-hand corner. Barely shadows content.	N/A	N/A
130, 140	Excellent	130 and 140: pencil markings	N/A	N/A
150	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) **Y** / N

Word	Comments
Ariel	Excellent.
José	Not recognized with the accent over the "e".
Jose	Jose (without accent) was erratic. All of the pages of the introduction had the author's name included in the header and "Jose" returned 10 hits while "Jos" returned 18 (highlighting "Jose").
Rodó	Does not work.
Rodo	14 hits with non-accented version of author's last name. "Rod" can't work because it forms the root of other words unrelated to this search.
63	Numbers work fine.
Utilitarianism	Excellent.
Tempest	Excellent.
( )	Parenthesis work fine.
Function	Readable; except for one instance where a reader crossed the word out with ink.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	Looks like the real deal – a palimpsest of different colored inks and pencil markings throughout and all clearly visible.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with great recognition.
Is the scan on a gray-scale or color scale?	Color

Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming excellent.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average. It will not recognize any foreign characters or symbols.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No. The scan is beautiful but it is too marked up to serve as an archive.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes – but question of foreign characters/symbols needs to be addressed.
Issues with this book that would affect usefulness for scholars	Foreign characters/symbols cannot be searched.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

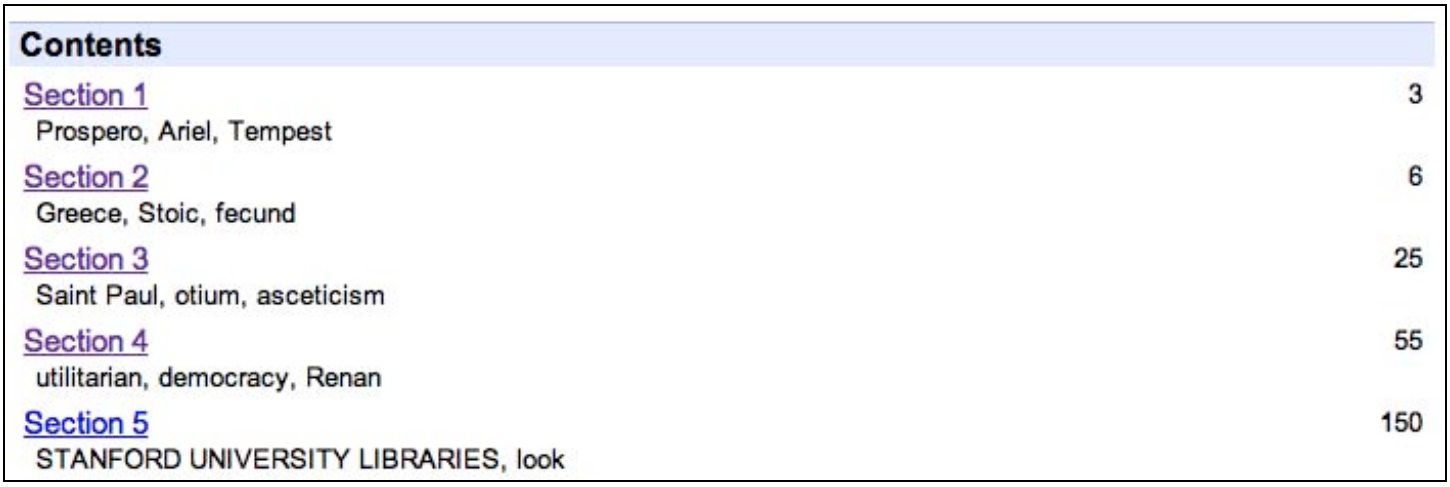
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No.
Other notes	N/A

Book URL: http://www.archive.org/details/arielrod00roduoft
Date accessed: 6/29/08

Appendix I: Google Box Search tagging for *Ariel* (1922).

Close-up of tagging.

Now see "created" sectional metadata – numerous convergences with tagging above.



CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 6/29/08

Best digitized book I have come across thus far!

http://en.wikipedia.org/wiki/Facundo

Facundo: Civilization and Barbarism (original Spanish title: *Facundo: Civilización y Barbarie*) is a book written in 1845 by Domingo Faustino Sarmiento, a writer and journalist who became the seventh president of Argentina. It is a cornerstone of Latin American literature: a work of creative non-fiction that helped to define the parameters for thinking about the region's development, modernization, power, and culture. Subtitled *Civilization and Barbarism*, Facundo contrasts civilization and barbarism as seen in early nineteenth-century Argentina. Literary critic Roberto González Echevarría calls the work "the most important book written by a Latin American in any discipline or genre".[1]

Book title: *Facundo*
Author: Sarmiento, Domingo Faustino
Year of Publication: 1921 [original date of publication is 1845]
Place of Publication: Buenos Aires
Publisher: Librería La Facultad, de J. Roldán.
ISBN number?: No [OCLC 6949295]
Content/Disciplinary area: Argentina -- History -- 1817-1860; Argentina -- Description and travel; Quiroga, Juan Facundo.

AVAILABILITY SEARCH

Source of scanned book:	Open-Access Text.
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Metadata is good – allowing each content/disciplinary area to be searched for other materials within the website. There is for example, an edition in English available on OAT that the metadata provides.
MARC record available?	Yes but had to go through WorldCat to send lending libraries MARC records.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	As with most Spanish-language texts, the table of contents or index is located at the end of the book – not at the beginning. It would be nice to be able to link directly from the chapter title in the index to the page itself. That being said, there is no guide to the book provided by metadata.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	**readability / quality of scan**	**anomalies?**	**other?**	**other?**

Illustration of author	Excellent	N/A	N/A	N/A
Title page	Excellent	Library stamp clearly visible.	N/A	N/A
10, 20	Excellent.	N/A	N/A	N/A
30, 40	Excellent	N/A	N/A	N/A
50, 60	Excellent	Page 50 excellent.	Page 60 has red ink underlining.	N/A
70, 80	Excellent	Page 70 – red ink underlining.	N/A	N/A
90, 100	Excellent	N/A	N/A	N/A
110, 120	Excellent	Page 110 – pencil writing	N/A	N/A
127	Excellent	Folded tab in upper right-hand corner. Barely shadows content.	N/A	N/A
130, 140	Excellent	N/A	N/A	N/A
150, 160	Excellent	N/A	N/A	N/A
170, 180	Excellent	N/A	N/A	N/A
190, 200	Excellent	N/A	N/A	N/A
210, 220	Excellent	N/A	N/A	N/A
230, 240	Excellent	N/A	N/A	N/A
250, 260	Excellent	N/A	N/A	N/A
270, 280	Excellent	Page 270 has some red ink markings.	Page 280 has some red ink markings.	N/A
290, 300	Excellent	N/A	N/A	N/A
310, 320	Excellent	N/A	N/A	N/A
330, 340	Excellent	N/A	N/A	N/A
350, 360	Excellent	N/A	N/A	N/A
370	Excellent	Red ink writing.		

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

(Is OCR search permitted) **Y** / N

Word	Comments
Facundo	Excellent.
Biografía	Excellent! The accented "í" was recognized!
Biografia	Trying the inverse here searching for the word without the accent. No results!
á, é, ó, ú, ñ	All recognized.
67	Numbers work for all of the pages involved.
¿?	Both work!

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	Only one illustration of author.
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	Looks like the real deal.
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF with excellent recognition.
Is the scan on a gray-scale or color scale?	Color
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Probably not.
Are the illustrations searchable?	No
Ability to print illustrations? (is there a watermark?)	Yes
Are the images downloadable? (file size, pixel quality)	Yes
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming excellent.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy. WorldCat.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Excellent.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes for the pdf but strangely the online version search function is not working so I wasn't able to double-confirm my results.
Ability to download?	Yes.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Excellent – it recognizes all of the foreign characters.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No. The scan is beautiful but it is too marked up to serve as an archive.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes.
Issues with this book that would affect usefulness for scholars	None.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	No.
Other notes	N/A

Book URL: http://www.archive.org/details/facundos00sarmuoft
Date accessed: 6/29/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 9/6/08

One of the few full-length Portuguese texts I have found.

Book title: Naufragio, e lastimoso successo da perdiçaõ de Manoel de Sousa de Sepulveda, e Dona Lianor de Sá, sua mulher, e filhos, vindo da India para este reyno na náo chamada o galiaõ grande S. Joaõ, que se perdeo no cabo de Boa-Esperança, na terra do Natal; e a peregrinação, que tiveraõ rodeando terras de Caftes, mais de 300 legoas, té sua morte (1783)
Year of Publication: 1783
Place of Publication: Lisbon
Publisher: Typografia Rollandiana
ISBN number?: OCLC 14394344
Content/Disciplinary area: Sepúlveda, Manuel de Sousa de.

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy to find. Metadata adequate.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	No table of contents included in metadata.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages)

Page(s):	readability/ quality of scan	anomalies?	Other?	Other?
titlepage	poor	Stained/Blotchiness/scrapings – legibility unaffected	N/A	N/A
i, x	Above average	Pencil markings – legibility not affected	N/A	N/A
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average	N/A	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A

280, 290	Above average	N/A	N/A	N/A
300, 310	Above average	N/A	N/A	N/A
320, 330	Above average	N/A	N/A	N/A
340, 350	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	(Is OCR search permitted) **Y** / N	Comment
Naufragios	Unreadable.	
Sepulveda	Excellent	
coração	Excellent	
manoel	Excellent	
Padre	Excellent	
indios	unsearchable.	
Lisboa	unsearchable	
soneto	Excellent	
333	Unreadable	

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Average recognition
Is the scan on a gray-scale or color scale?	Color scale
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming very nice.

Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Above average.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	No table of contents included in metadata.
How successful is OCR recognition?	Average
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No.
Issues with this book that would affect usefulness for scholars	No table of contents included in metadata.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

Book URL: http://www.archive.org/details/http://www.archive.org/details/naufragioelastim00cortiala
d: 9/6/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/17/08

http://en.wikipedia.org/wiki/Unamuno

Miguel de Unamuno y Jugo (September 29, 1864–December 31, 1936) was an essayist, novelist, poet, playwright and philosopher from Bilbao, Biscay, Spain.

Book title: The tragic sense of life in men and in peoples
Author: Unamuno, Miguel de; Flitch, John Ernest Crawford; Madariaga, Salvador de
Year of Publication: 1921
Place of Publication: London
Publisher: Macmillan and Co., Limited
ISBN number?: OCLC 1472494
Content/Disciplinary area: Immortality; Philosophy and religion; Pessimism

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Average. Table of contents not included in metadata.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) The page numbers in the upper corners are very hard to make out. I can't tell if it is a printing error or scanning error. Since it's a dictionary, its all the more easier to tell if a page is missing.

Page(s):	**readability/ quality of scan**	**anomalies?**	**other?**	**other?**
Dedication; signatures	Excellent (in color)	N/A	N/A	N/A
x, xx, xxx	Excellent	N/A	N/A	N/A
Illustrations (first two – author and publisher's seal)	Excellent	N/A	N/A	N/A
2	Excellent	N/A	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A

120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A
200, 210	Excellent	N/A	N/A	N/A
220, 230	Excellent	N/A	N/A	N/A
240, 250	Excellent	N/A	N/A	N/A
260, 270	Excellent	N/A	N/A	N/A
280, 290	Excellent	N/A	N/A	N/A
300, 310	Excellent	N/A	N/A	N/A
320, 330	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment (Is OCR search permitted) **Y** / N (Available only in online version, not pdf)
Tragic	Excellent
Unamuno	Excellent
1921	Excellent
Introductory	Excellent
Aunque	Excellent
Señor	Unreadable (accents cannot be searched)
fia	Unreadable
Malón	Unreadable
Index	Excellent

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	Average recognition – accented words are not searchable.
Is the scan on a gray-scale or color scale?	Color
Quality – does it look like the real thing?	Yes
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	Yes.

Are the images downloadable? (file size, pixel quality)	Yes.
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming excellent.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy
Overall quality of scan of this book	Excellent
Ease of access to full text	Easy
Ability to copy, quote, cite?	Yes
Ability to search?	Yes (only online version)
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No
Issues with this book that would affect usefulness for scholars	Accented words are not searchable.
Can this book be shared with other scholars/students?	Yes – if accented words are fixed.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A
Other notes	N/A

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Book URL: http://www.archive.org/details/thetragicsenseof00unamuoft
Date accessed: 8/17/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 7/08/08

Inspiration for the fabulous movie *Aguirre Wrath of God*.

http://en.wikipedia.org/wiki/Lope_de_Aguirre

Lope de Aguirre (c. 1510 – 27 October 1561) was a Spanish Basque conquistador in South America. Sent, along with other rebellious settlers, on an impossible mission in search of the mythical El Dorado on the Amazon river, he eventually became their leader and rebelled against Philip II, before being defeated and slain.

He joined the 1560 expedition of Pedro de Ursúa down the Marañón and Amazon Rivers with 300 men and hundreds of natives. A year later, he participated in the overthrow and killing of Ursúa and his successor, Fernando de Guzmán, whom he ultimately succeeded. He and his men reached the Atlantic (probably by the Orinoco River), destroying native villages on the way. In March 23, 1561, Aguirre urged 186 captains and soldiers to sign an act which would proclaim him as prince of Peru, Tierra Firma and Chile.

He is reputed to have said in 1561:

I am the Wrath of God,
the Prince of Freedom,
Lord of Tierra Firme and the Provinces of Chile

In 1561, he seized Isla Margarita and brutally suppressed any opposition to his reign. When he crossed to the mainland in an attempt to take Panama, his open rebellion against the Spanish crown came to an end. He was surrounded at Barquisimeto, Venezuela, where he murdered his own daughter Elvira "because someone that I loved so much should not come to be bedded by uncouth people". He also killed several followers who intended to capture him. He was eventually captured and shot. Aguirre's body was cut into quarters and sent to various cities across Venezuela.

Book title: Relación de todo lo que sucedió en la jornada de Omagua y Dorado hecha por el gobernado Pedro de Orsúa
Author: Vázquez, Francisco, fl. 1559
Year of Publication: 1881
Place of Publication: Madrid
Publisher: Impr. de Miguel Ginesta
ISBN number?: OCLC 55272630
Content/Disciplinary area: Aguirre, Lope de; Urzua, Pedro de; Peru – History – Discovery and Conquest, 1519-1536.

AVAILABILITY SEARCH

Source of scanned book:	Open Access Text
Comments on condition/provenance of book:	Excellent.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Easy.
MARC record available?	Yes through WorldCat
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Above average. No table of contents in the actual book but the metadata does link within OAT.

SCAN SEARCH

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

search paradigm settings (example: every 10 double-pages) Page(s):	readability/ quality of scan	anomalies?	other?	other?
title page	Excellent	N/A	N/A	N/A
i - xlviii	Excellent	N/A	N/A	N/A
1, 10	Excellent	N/A	N/A	N/A
20, 30	Excellent	N/A	N/A	N/A
40, 50	Excellent	N/A	N/A	N/A
60, 70	Excellent	N/A	N/A	N/A
80, 90	Excellent	N/A	N/A	N/A
100, 110	Excellent	N/A	N/A	N/A
120, 130	Excellent	N/A	N/A	N/A
140, 150	Excellent	N/A	N/A	N/A
160, 170	Excellent	N/A	N/A	N/A
180, 190	Excellent	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) **Y** / N
Omagua	Excellent.
Dorado	Excellent.
Orsúa	Accents not recognized. Orsua without accents recognized.
MDCCCLXXXI	Not recognized. Erroneously read as MDC**G**CLXXXI.
61	Not recognized. Erroneously recognizes é instead.
á, é, í, ó, ú, ñ	Unreadable.
Aguirre	Excellent.
Lope de Aguirre	Average. Reads each word individually – not as a phrase.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	PDF. Average recognition.
Is the scan on a gray-scale or color scale?	Color.
Quality – does it look like the real thing?	Yes.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	N/A

Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	N/A
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming great.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Excellent.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	Yes.
Ability to search?	Yes.
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes.
How successful is OCR recognition?	Average.
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	Yes.
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	Yes but need to improve character recognition.
Issues with this book that would affect usefulness for scholars	Character recognition lacking.
Can this book be shared with other scholars/students?	Yes.
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	Yes if they fix character recognition.
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	N/A

Other notes	N/A

Book URL: http://www.archive.org/details/relacibondetodo00vazqrich
Date accessed: 7/08/08

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Researcher name: Patricia A. Soler **Date of investigation**: 8/17/08

Note: I previously completed a book search form for the same author – whose last name is spelled either Vieira or Vieyra. My goal is to examine more Portuguese-language texts. This one interested me due to it being a Portuguese-English/English-Portuguese dictionary.

Book title: A new pocket dictionary of the Portuguese and English languages, in two parts, viz. Portuguese and English, and English and Portuguese : abridged from Vieyra's dictionary
Author: Antonio Vieyra; J P Aillaud
Year of Publication: 1837 (Vol. II)
Place of Publication: Paris
Publisher: Printed for J.P. Ailaaud by Casimer
ISBN number?: OCLC 44879365
Content/Disciplinary area: Portuguese language -- Dictionaries – English; English language -- Dictionaries -- Portuguese.

AVAILABILITY SEARCH

Source of scanned book:	Google Books
Comments on condition/provenance of book:	Poor. PDF.
Ease of finding book, appropriateness of available metadata in the project supplying book (i.e. multivolume works, books with different editions/dates/titles)	Average. Since Google Books does not allow for searching for all texts in a given language, I limited the language to Portuguese and searched for the keyword "Portuguese".
MARC record available?	Yes. Through WorldCat.
Can the book be searched through an OPAC? If so, which one?	WorldCat.
Comments on accuracy/ usefulness of book metadata (i.e. table of contents? guide to book?)	Poor. Google gets the author's name wrong, listing him as Domingos Vieira, when it is very clearly Antonio Vieyra. The metadata should also include the fact that there are two common spellings of the author's name (Vieyra or Vieira). This is especially problematic because the other volume is also mislabeled while an earlier edition of the same text (furthermore edited by the same editor) is correctly labeled.

SCAN SEARCH

search paradigm settings (example: every 10 double-pages) The page numbers in the upper corners are very hard to make out. I can't tell if it is a printing error or scanning error. Since it's a dictionary, its all the more easier to tell if a page is missing.

Page(s):	readability/ quality of scan	anomalies?	other?	other?
Ex libris	Excellent (in color)	N/A	N/A	N/A
Marbled paper	Above average	Finger!	N/A	N/A
Dedication	Poor	Too light – can't make it out due to gray-scale	N/A	N/A
2	Above average	Odd lighter shaded area – legibility not affected	N/A	N/A
1, 10	Above average	N/A	N/A	N/A
20, 30	Above average	N/A	N/A	N/A
40, 50	Above average	N/A	N/A	N/A

64, 65, 66	Poor	64 and 65 duplicated	65a - part of margin from opposite page visible.N/A	66 extremely light – legibility affected
60, 70	Above average	N/A	N/A	N/A
80, 90	Above average	N/A	N/A	N/A
102, 105, 106	Beyond poor. All out of order.	102 – odd lighter shaded area – legibility not affected.	105 included 3 times.	106 illegible – totally blurred.
103	Missing!	N/A	N/A	N/A
100, 110	Above average	N/A	N/A	N/A
120, 130	Above average	N/A	N/A	N/A
140, 150	Above average	N/A	N/A	N/A
160, 170	Above average	N/A	N/A	N/A
180, 190	Above average	N/A	N/A	N/A
200, 210	Above average.	210 has lighter shaded patch – legibility not affected.	N/A	N/A
220, 230	Above average	N/A	N/A	N/A
240, 250	Above average	N/A	N/A	N/A
260, 270	Above average	N/A	N/A	N/A
280, 290	Above average	N/A	N/A	N/A
300, 310	Above average	N/A	N/A	N/A
320, 330	Above average	N/A	N/A	N/A
340, 350	Above average	N/A	N/A	N/A
360, 370	Above average	N/A	N/A	N/A
380, 390	Poor	390 – right-side margin cut off.	N/A	N/A
416	Above average	Duplicated	N/A	N/A
400, 410	Above average	N/A	N/A	N/A
420, 430	Above average	N/A	N/A	N/A
440, 450	Above average	N/A	N/A	N/A
460	Above average	N/A	N/A	N/A

WORD SEARCHES DONE WITH THIS BOOK

(at least 10, include footnotes, endnotes, comment on accuracy of finding hits on pages I've already examined above)

Word	Comment
	(Is OCR search permitted) Y / **N** (Available in both online and pdf versions)
	Not searchable in online version or pdf.

IMAGE ASSESSMENT

If the original text includes graphics/illustrations, are they included in the electronic version?	N/A
If there are no graphics/illustrations are there other visual elements missing? (wormholes, ink smears, page aging/mold, paper quality)	N/A
Is the image scan an OCR or a PDF? If OCR, how successful is the recognition?	N/A
Is the scan on a gray-scale or color	Gray-scale

scale?	
Quality – does it look like the real thing?	Online yes; pdf no.
Illustrations? (would the graphics/illustrations be useful to art historians, visual culture researchers?)	Yes.
Are the illustrations searchable?	N/A
Ability to print illustrations? (is there a watermark?)	Yes.
Are the images downloadable? (file size, pixel quality)	N/A
Does the electronic version allow for any advantages that the print version would not allow? (zooming capabilities, filter options)	Zooming average due to pdf being gray-scale.
Other notes	N/A

GENERAL OBSERVATIONS

Comment on ease of discovery in integrated interface (OPAC or better)	Easy.
Overall quality of scan of this book	Poor.
Ease of access to full text	Easy.
Ability to copy, quote, cite?	No
Ability to search?	No
Ability to download?	Yes.
Ability to print?	Yes.
Scanned book allow for consulting of contents?	Yes (only due to the fact that it is a dictionary)
How successful is OCR recognition?	N/A
Worthy as archive? (i.e., would this copy be acceptable as an archival copy of the book if physical copies were lost or destroyed?	No
Reproduction – is quality of scan good enough that this book could be reproduced into analog form?	No
Issues with this book that would affect usefulness for scholars	Very poor quality scan; no OCR recognition; metadata erroneous.

CLIR MASS DIGITIZATION EVALUATION PROJECT
BOOK SEARCH FORM

Can this book be shared with other scholars/students?	No
"web 2.0 functionality" – can digitized text be reprocessed in order to support other kinds of research/ use?	No
Predictability and reliability of collection – any items missing (i.e. volumes in a set, standard words, etc.)	Metadata includes other volume but still erroneous.
Other notes	N/A

Book URL:
http://books.google.com/books?id=Mt0FAAAAQAAJ&dq=portuguese&lr=lang_pt&as_brr=1&source=gbs_summary_s&cad=0
Date accessed: 8/17/08