



UC Berkeley School of Information

# Google Books: The Metadata Mess

---

**Google Book Settlement Conference**  
**UC Berkeley**  
**August 28, 2009**

**Geoff Nunberg,**  
**School of Information**



## The Last Library

---

"The cost of creating such a library and Google's significant lead time advantage suggest that no other entity will create a competing digital library for the foreseeable future."

Directors of ALA, ACRL, ARL in letter to DOJ Antitrust Division, July 29, 2009

There is no Moore's Law for capture...

Hence the urgency of concerns about pricing, access, exclusivity, privacy...and "quality"



# Whose interests determine "quality"?

Google Book Search is "a tremendous public good for students, for teachers, for scholars, for everyone." Derek Slater, Google

... but students, scholars and "everyone" may have different purposes for using GBS.



# Three ways of using GBS

What "Googling" means: barrelling in sideways



GBS as a borough of Greater Google

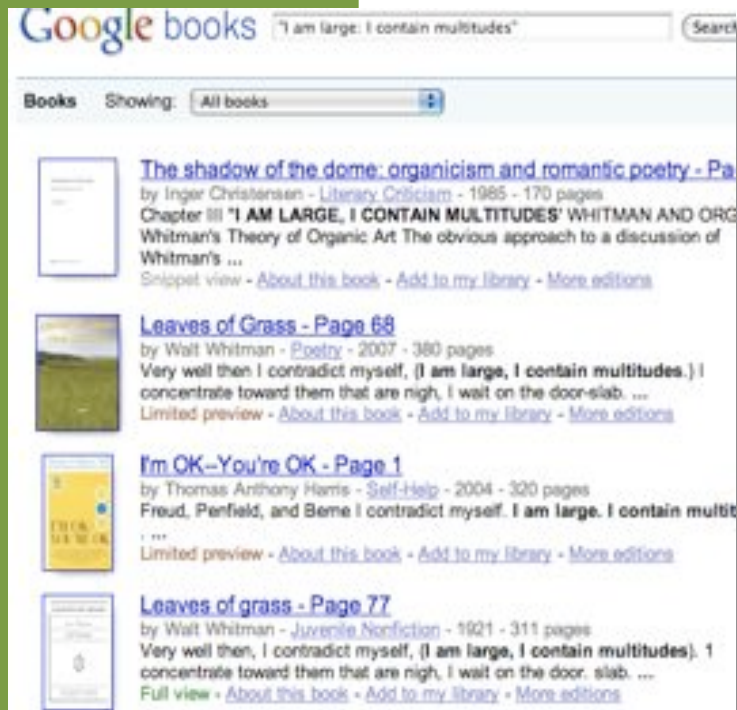
"We just feel this is part of our core mission. There is fantastic information in books. Often when I do a search, what is in a book is miles ahead of what I find on a Web site." Sergey Brin



# Three ways of using GBS

Seeking out works & editions: the "destination experience"

A particular edition of *Leaves of Grass*  
A good edition of *Tristram Shandy*  
18<sup>th</sup>-c. French editions of *Don Quixote*,  
etc.



The importance of metadata: Who, when, where etc.



## Three ways of using GBS

---

"Batch processing": data mining and  
"electronic philology"

"It's only reporters and computational linguists who care if [hit-count estimation] is really precise." Peter Norvig, Google

Text databases and the "new philologies":

The importance of language to social, intellectual, and political history & literary study

Coincides emergence of large-scale historical text databases...

When did *happiness* replace *felicity* in 17<sup>th</sup> c?

Plotting the rise & fall of *propaganda*

How did *liberalism* spread in the early nineteenth-century European context?.



## Good enough for scholarship?

---

Will GBS be an adequate resource for scholarly needs... now and in the future?

Depends on:

- Quality of imaging

- Reliability and robustness of search tools

- Quality and reliability of metadata

  - e.g., date, edition history, author, subject classification, etc.



## Good enough for scholarship?

---

Will GBS be an adequate resource for scholarly needs... now and in the future?

Depends on:

- Quality of imaging

- Reliability and robustness of search tools

- Quality and reliability of metadata

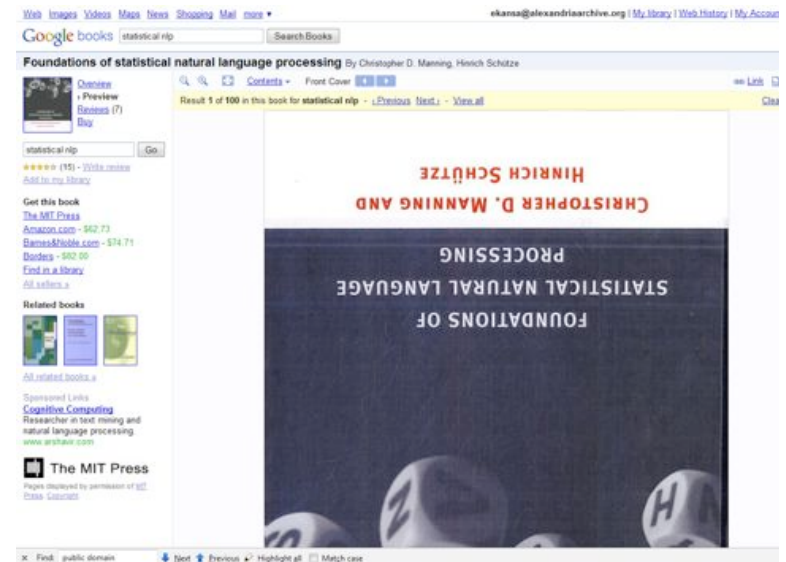
  - e.g., date, edition history, author, subject classification, etc.

**But GBS metadata are awful.**





# Quality Issues : Botched Scans, OCR, &c.



Showing 31 results in this book for Z113NZM - Order by: [relevance](#) | [pages](#)

[Page 1](#)

indústria, capaz de ser facilmente arrasada pela maior força militar  
a do direito à guerra preventiva, é escolher um alvo constantemente  
A mancha mais simples de estabelecer uma nova norma, tal como  
motivos duvidosos, cria-se uma norma. Esse é o significado do poder.  
norma. Mas quando os Estados Unidos bombardeiam a Sérvia por



# Metadata Issues: 1899, annus mirabilis



## KILLER IN THE RAIN

by RAYMOND CHANDLER - Fiction - 1899

It was in the pulp detective magazines of the 1930s that Raymond Chandler's definitive take on the hard-boiled detective story first appeared. ...  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## CULTURE AND SOCIETY 1780-1950 - Page 206

by RAYMOND WILLIAMS - Social Science - 1899

This was the positive result of the life of the family in a small house. ... are not separated from personal **relationships**; and Lawrence knew from this, ...  
[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## CONDITION HUMAINE, LA

by ANDRE MALRAUX - Fiction - 1899 - 352 pages

En mars 1927, l'armée du Kuomintang dirigée par Chang-Kai-Shek s'approche de Shanghai. Les communistes de la ville organisent le soulèvement des ouvriers ...  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## PORTABLE DOROTHY PARKER, THE

by DOROTHY PARKER, SETH - Fiction - 1899 - 600 pages

The second revision in sixty years, this sublime collection ranges over the verse, stories, essays, and journalism of one of the twentieth century's most ...  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## Christine

by Stephen King - Fiction - 1899 - 334 pages

Page 11  
PRÓLOGO Esta é a história de um triângulo amoroso — suponho que este seria o nome — formado por Amie Cunningham, Leigh Cahot e, naturalmente, Christine. ...  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## GRAPES OF WRATH, THE: TEXT AND CRITICISM

by JOHN STEINBECK - Literary Criticism - 1899 - 806 pages

This book includes the restored text corrected by scholar Robert DeMott, a map of the Joad family's journey, Steinbeck's declaration of intent in writing ...  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## COMPLETE SHORTER FICTION OF VIRGINIA WOOLF, THE

by SUSAN DICK - Fiction - 1899

This volume brings together all of Virginia Woolf's short stories and sketches.  
[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## YELLOW SUBMARINE

by THE BEATLES - Juvenile Fiction - 1899 - 40 pages

Uma aventura fascinante com a maior banda de rock de todos os tempos! Era uma vez - ou duas, talvez - um paraíso chamado Pepperland. ...  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



# Random Dates

## The World According to Drucker



Overview  
[Reviews](#) (0)  
[Buy](#)

☆☆☆☆ (0) - [Write review](#)  
[Add to my library](#)

### Book overview

Peter Drucker Is The Man Who The Economist Calls 'The Biography Airs His Controversial Views On Business, Go Future.

No preview available - 1905 - 220 pages

1905

## What Maisie knew By Henry James



Overview  
[Reviews](#) (3)  
[Buy](#)

### Book overview

No preview available - 1848

1848

## Letters from Virginia Woolf to Jacques and Gwen Raverat

by Virginia Woolf - 1900 - 100 pages

No preview available - [About this book](#) - [Add to my library](#)

1900



## The bonfire of the vanities

by Tom Wolfe - Fiction - 1888 - 690 pages










Tom Wolfe's modern American satire tells the story of Sherman McCoy, a Wall Street "Master of the Universe" who has it all — a Park Avenue apartment, a...

No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)

1888



# The pervasiveness of misdatings

	<b>Internet security</b> by Dan Farmer, WIETSE VENEMA - <a href="#">Computers</a> - 1899 - 800 pages Computer forensics, the art and science of gathering and analyzing digital evidence, reconstructing data and attacks, and tracking perpetrators, ... No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a>	1899
	<b>Multicultural Teaching With The Internet</b> by Sunil Chawan - <a href="#">Internet in education</a> - 1905 - 301 pages No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a>	1905
	<b>Physics</b> by Balfour Stewart - <a href="#">Education</a> - 1878 - 149 pages <a href="#">Full view</a> - <a href="#">About this book</a> - <a href="#">Add to my library</a> - <a href="#">More editions</a>	1878
	<b>Internet For Libraries And Information Centres (book + Floppy)</b> by Jambhekar - <a href="#">Library information networks</a> - 1905 - 174 pages No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a>	1905
	<b>Emma - Page 55</b> <a href="#">Fiction</a> - 1948 - 831 pages Die bekannteste dürfte der <b>Internet</b> Relay Chat (IRC) sein, ... Denn: Das <b>Internet</b> ist das, was seine Nutzerinnen und Nutzer daraus machen ... Snippet view - <a href="#">About this book</a> - <a href="#">Add to my library</a> - <a href="#">More editions</a>	1946
	<b>Internet Technologies And Applications</b> 1905 No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a> - <a href="#">More editions</a>	1905
	<b>Mohit Dictionary Of Internet And Web</b> by Vineet Gupta - <a href="#">Internet</a> - 1905 - 392 pages No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a>	1905
	<b>Project implementation and management - bridging the gap proc. 5th Internet ...</b> 1876 - 305 pages No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a>	1905
	<b>The Mosaic navigator. The essential guide to the internet interface</b> by Sigmund Freud, Katherine Jones - 1939 - 218 pages <b>Includes glossary, index.</b> No preview available - <a href="#">About this book</a> - <a href="#">Add to my library</a>	1939

527 hits returned for  
"Internet" before 1950





# Famous before their lifetime

The screenshot shows a Google Books search interface. The search term is "charles dickens". The "Publication Date" filter is set to "Return content published between" with a red circle around the filter options. The search results show several books, including "Household words: a weekly journal", "A tale of two cities", "Great expectations", "A Christmas carol", and "Our mutual friend". A red arrow points from the filter in the main screenshot to a zoomed-in view of the same filter in the bottom right corner, which shows the year "1812" entered in the "and" field.

Web Images Videos Maps News Shopping Local more ▾    numbergg@gmail.com | My library | My Account |

Google books    "charles dickens"    Search Books    Advanced Book Search

Publication Date     Return content published anytime     Return content published between    and    1812

Books    Showing: All books    Books 1 - 100 of 182 on "charles dickens". (0.22 seconds)

**Household words: a weekly journal**  
by Charles Dickens - Juvenile Fiction - 1742  
Full view - About this book - Add to my library - More editions

**A tale of two cities**  
by Charles Dickens - 1800 - 378 pages  
No preview available - About this book - Add to my library - More editions

**Great expectations**  
by Anne Marie Mueser, Charles Dickens - 1172 - 30 pages  
No preview available - About this book - Add to my library - More editions

**A Christmas carol**  
by Charles Dickens - Juvenile Fiction - 1135 - 294 pages  
No preview available - About this book - Add to my library - More editions

**Our mutual friend**  
by Charles Dickens - Drama - 1800  
Snippet view - About this book - Add to my library - More editions

Sponsored Links

**CHARLES DICKENS'S**  
Get Charles Dickens's Find & Compare Great Options  
Charles.Ask.com

**Find Charles Dickens**  
Get current address, phone & more  
Easy to use, search for free!  
www.usa-people-search.com

"charles dickens"    Search Books    Advanced Book Search

Publication Date     Return content published anytime     Return content published between    and    1812

Books 1 - 100 of 182 on "charles dickens". (0.22 seconds)

182 hits reported for "Charles Dickens" before birthdate (1812)

Cf Jimi Hendrix, 81; Led Zeppelin, 59 etc.

# Ego-surfing, Edgar Cayce Style



## VIRTUAL COMMUNITY: HOMESTEADING ON THE ELECTRONIC FRONTIER

by HOWARD RHEINGOLD - Computers - 1899  
[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## Do Que É Feito O Pensamento

by STEVEN PINKER - Psychology - 1899 - 44 pages  
Page 411

... no really good? Além disso, como ressaltou o linguista  
Geoffrey Nunberg,

embora de cara imaginar o dilema How brilliant was it? View INão



## VIDA SOCIAL DA INFORMAÇÃO, A

by JOHN SEELY BROWN, PAUL DUGUID - Computers - 1899 -  
300 pages

Com base em suas experiências profissionais como cientista-chefe  
e pesquisador-



## Annual report of the American historical Association - Page 198

History - 1884

... Berkeley Seized Letters as Legal Evidence in the Paris Revolutionary  
Tribunal, 1793-1794. Carla Hesse, University of California, ...

Snippet view - [About this book](#) - [Add to my library](#) - [More editions](#)



## New Jersey history

by New Jersey Historical Society - History - 1971

... Annalee Saxenian, 'In Search of Power: The Organization of Business  
Interests in Silicon Valley and Route 128,' Economy and Society 18 (1989): 25-70

; ...

Snippet view - [About this book](#) - [Add to my library](#) - [More editions](#)

---

"Our reputation  
precedes us"



# The frequency of misdatings

Search on "candy bar" < 1920 yields 66 hits, 46 of them misdated (70%)





# Classification Errors



## Hamlet

by William Shakespeare - [Antiques & Collectibles](#) - 1994 - 71 pages  
très douces manières et de grande mine. En vérité, pour parler de lui avec tact,  
il est le calendrier, la carte de la gentry ; vous trouverez en lui le ...  
[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## MADAME BOVARY:보봐리 부인(불문학작품 40)

by G.FLAUBERT - [Antiques & Collectibles](#) - 1988 - 430 pages  
No preview available - [About this book](#) - [Add to my library](#)



## Speculum - Page 283

by Mediaeval Academy of America, JSTOR (Organization) - [Health & Fitness](#) - 1962  
**WERNER JAEGER**, Early Christianity and Greek **Paideia**. Cambridge, Massachusetts:  
Belknap Press of Harvard University Press, 1961. Pp. ii, 154. \$8.25. ...  
[Snippet view](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## Mae West: An Icon in Black and White - Page 267

by Jill Watts - [Religion](#) - 2003 - 400 pages  
Diamond Li's iand **Mae West**'si successful perpetuation of female trans- ... In  
the midst of this revival of the **cult** of ...  
[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## Australian women: contemporary feminist thought - Page 153

by Norma Grieve, Ailsa Burns - [Foreign Language Study](#) - 1994 - 356 pages  
The group moved more or less at random from the writings of Michel **Foucault** to  
lesbian-feminist philosophy, from Freud to Freud's commentators and critics. ...  
[Snippet view](#) - [About this book](#) - [Add to my library](#) - [More editions](#)





# Classification Errors



## [The merchant of Venice](#)

by Christopher Rice, William **Shakespeare** - [Foreign Language Study](#) - 2007 - 72 pages  
No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## [The Century dictionary: an encyclopedic lexicon of the English language - Page 6180](#)

edited by William Dwight **Whitney** - [Family & Relationships](#) - 1891  
**JD Whitney**, The Yosemite Book, p. 24. 2. A device used in dredging, for sweeping  
the sea-bed in order to obtain delicate forms of marine life, ...  
[Full view](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## [Melting the earth: the history of ideas on volcanic eruptions - Page 76](#)

by Haraldur Sigurdsson - [Foreign Language Study](#) - 1999 - 260 pages  
**Dante** also speculated on the powers that had raised the lands above the primeval ocean, opting for an extraterrestrial force. Action was effected "by way of ...  
[Snippet view](#) - [About this book](#) - [Add to my library](#) - [More editions](#)

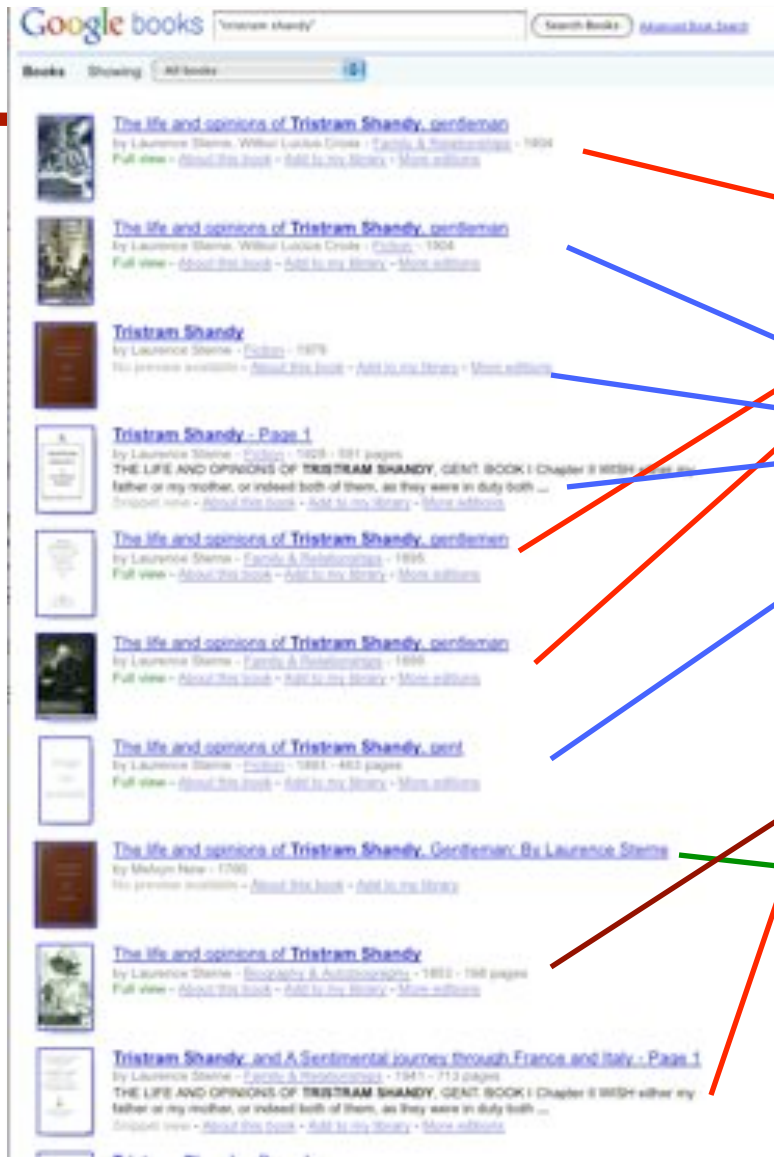


## [Catalogue of copyright entries - Page 744](#)

by Library of Congress. **Copyright Office** - [Drama](#) - 1923  
13, Feb., 1922» 0 Feb. 25, 1922 : 2 c. and aff. Mar. 9. ... **Volume** of production  
— united States ...  
[Full view](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



# The Pervasiveness of Misclassification



family and relationships (4)

fiction (4)

biography and autobiography (1)

Unlabeled (1)  
(others classified as "music,"  
"history," "literary collections")

**Classifications of first 10 hits for  
*Tristram Shandy***



# The Pervasiveness of Misclassification

Web Images Videos Maps News Shopping Gmail more ▾ nunbergg@gmail.com | My library | My Account | Sign out

Google books "leaves of grass" Search Books Advanced Book Search

Books Showing: All books Books 1 - 17 of 17 on "leaves of grass". (0.45 seconds)

List view Cover view

**Leaves of grass**  
by Walt Whitman - *Juvenile Nonfiction* - 1921 - 311 pages  
Unlike many other editions of Leaves of Grass, which reproduce various short, early versions, this Modern Library Paperback Classics "Death-bed" ...  
Full view - About this book - Add to my library - More editions

**Leaves of grass: first and "death-bed" editions: additional poems**  
by Walt Whitman, Karen Karbiener, George Stade - *Poetry* - 2004 - 913 pages  
Here are some of the remarkable features of Barnes & Noble Classics: New introductions commissioned from today's top writers and scholars Biographies...  
Limited preview - About this book - Add to my library - More editions

**Leaves of grass**  
by Walt Whitman - *Fiction* - 1983 - 470 pages  
In 1871 he brought out a fifth *Leaves of Grass*, which included "Passage to India," and in 1881 a sixth, in which the poems ...  
Limited preview - About this book - Add to my library - More editions

**Leaves of grass: including Sands at seventy, first annex, Good-by my fancy...**  
by Walt Whitman - *Literary Criticism* - 1931 - 580 pages  
Snippet view - About this book - Add to my library - More editions

**Leaves of grass**  
by Walt Whitman - *Biography & Autobiography* - 1897 - 455 pages  
Copy is in a slip case, book has no covers.  
Full view - About this book - Add to my library

Sponsored Links

**Ornamental Grass Sale**  
Grow Ornamental Grasses For Less - \$20 Off Now At Michigan Bulb.  
www.MichiganBulb.com

**Trees For Every Season**  
Reinvent Your Living Space Today  
Novel Trees Made with You in mind.  
www.EarthFlorA.com

**Glass Flowers**  
Handcrafted Glass Flowers & Glass Bouquets from Europe. Unique Gifts!  
www.bestartshop.com

**Plant Sod**  
Find Top-Rated Local Sod Pros & Get 4 Free Bids Today!  
www.ServiceMagic.com

**Patch Perfect Grass Seed**  
Official Patch Perfect Website  
As Seen on TV. Buy 1 Get 1 Free  
www.OfficialTVWebsite.com

**Leaves Of Plants**  
Leaves Of Plants Online.  
Shop Target.com.  
www.Target.com

**Plants at fresh&easy™**  
Fresh and healthy produce choices.  
But low prices.

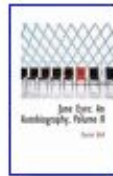
**First 10 hits for *Leaves of Grass* classify it as:**

*Juvenile Nonfiction*  
*Poetry*  
*Fiction*

*Literary Criticism*  
*Biography & Autobiography,*  
*Counterfeits and Counterfeiting*



# More bad metadata



## Jane Eyre: An Autobiography

by Currer Bell - [History](#) - 2008 - 364 pages

[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## Jane Eyre

by Charlotte Brontë - [Governesses](#) - 1986 - 483 pages

No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## Jane Eyre

by Charlotte Brontë - [Love stories](#) - 2000 - 452 pages

No preview available - [About this book](#) - [Add to my library](#)



## Jane Eyre

by Charlotte Brontë, Basil Davenport - [Architecture](#) - 1946 - 474 pages

Snippet view - [About this book](#) - [Add to my library](#) - [More editions](#)



## Jane Eyre

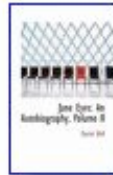
by Charlotte Brontë, Clare West - [Antiques & Collectibles](#) - 1992 - 105 pages

No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)





# More bad metadata



## Jane Eyre: An Autobiography

by Currer Bell - [History](#) - 2008 - 364 pages

[Limited preview](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## Jane Eyre

by Charlotte Brontë - [Governesses](#) - 1986 - 483 pages

No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)



## Jane Eyre

by Charlotte Brontë - [Love stories](#) - 2000 - 452 pages

No preview available - [About this book](#) - [Add to my library](#)



## Jane Eyre

by Charlotte Brontë, Basil Davenport - [Architecture](#) - 1946 - 474 pages

[Snippet view](#) - [About this book](#) - [Add to my library](#) - [More editions](#)



## Jane Eyre

by Charlotte Brontë, Clare West - [Antiques & Collectibles](#) - 1992 - 105 pages

No preview available - [About this book](#) - [Add to my library](#) - [More editions](#)

Reader, I marketed him.



## Other metadata issues

Books ascribed to authors of introductions, or given no author at all.





# Other metadata issues

Titles linked to unrelated works.

**Théorie de l'univers, ou, De la cause primitive du mouvement, et de ses ...** By Jacques Alexander François Allix

Overview  
Reviews (0)  
Buy

☆☆☆☆ (0) - Write review  
Add to my library

Get this book  
AbeBooks  
Alibris  
Amazon  
Barnes&Noble.com  
Find in a library  
All sellers >

**Book overview**  
Snippet view - Edition: 2 - 1818 - 263 pages - [Science](#) **Other editions**  
Edition 2 - 1818 - [Full view](#)

**Search in this book**  
"theorie de l'univers" bradford

Your search - "theorie de l'univers" bradford - did not match any documents.  
Try this search over all volumes: ["theorie de l'univers" bradford](#)

**Reviews**  
We haven't found any reviews in the usual places.  
[Write review](#)

**Common terms and phrases**  
actress Anyway asked **BARBARA TAYLOR BRADFORD**  
Bavaria beautiful breath chair champagne stems Ch  
Claridge's color cottage pie bobby darling Dian  
door Doris dressing table Earl enchilada Estelle exclaimen  
film Francesca Franka gaze gaze girl glanced glas  
Heathcliff Hollywood take term Katharine Tem

**Supervision and Clinical Psychology: Theory, Practice and Perspectives** By Ian Fleming, Linda Steen

Overview  
Preview  
Reviews (1)  
Buy

Search in this book

☆☆☆☆ (1) - Write review  
Add to my library

Get this book  
Psychology Press  
Amazon.com  
Barnes&Noble.com - \$69.95  
Borders  
All sellers >

**Related books**  
  
All related books >

**Psychology Press**  
Taylor & Francis Group  
Pages displayed by permission of  
Dordrecht: Kluwer Academic Publishers

**AMERICAN POLITICS  
in  
HOLLYWOOD FILM**  
  
IAN SCOTT

**Household words: a weekly journal, Volume 10** By Charles Dickens

Overview  
Read  
Reviews (4)  
Buy

Search in this book

☆☆☆☆ (0) - Write review  
Add to my library

Get this book  
AbeBooks  
Alibris  
Amazon  
Barnes&Noble.com

**HISTOIRE  
DE  
L'ACADEMIE  
ROYALE  
DES SCIENCES.**



# Other metadata issues

---

## Strange bedfellows



[The Mosaic navigator: The essential guide to the internet interface](#)

by Sigmund Freud, Katherine Jones - 1939 - 218 pages

Includes glossary, index.

No preview available - [About this book](#) - [Add to my library](#)





# Who is to blame and what is to be done?

"We got the metadata from the libraries":

yes, sometimes... but libraries didn't classify *Hamlet* as "antiques and collectibles" or *Speculum* as "Health & Fitness"

Libraries don't use BISAC headings like "Antiques and Collectibles" and "Health & Fitness" in the first place...

And publishers didn't assign BISAC codes to books published before the 1980's

The screenshot shows two search results from Google Books. The first result is for 'Speculum - Page 283' by Mediaeval Academy of America, JSTOR (Organization) - Health & Fitness - 1962. The second result is for 'Hydriotaphia: urne-buriall, or a discourse of the sepulchral urnes lately ...' by Sir Thomas Browne, published in 1907. The interface includes search bars, navigation links like 'Web Images Videos Maps News Shopping Gmail more', and options to 'Read this book'.



# The world according to BISAC

## Making space for Bambi & Bullwinkle

JNF003260	JUVENILE NONFICTION / Animals / Cows *
JNF003230	JUVENILE NONFICTION / Animals / Deer, Moose & Caribou
JNF003050	JUVENILE NONFICTION / Animals / Dinosaurs & Prehistoric Creatures
JNF003060	JUVENILE NONFICTION / Animals / Dogs
JNF003210	JUVENILE NONFICTION / Animals / Ducks, Geese, etc.
JNF003070	JUVENILE NONFICTION / Animals / Elephants

## ... and Schiller, Petrarch & Verlaine

POE011000	POETRY / Canadian
POE012000	POETRY / Caribbean & Latin American
POE005030	POETRY / Continental European
POE005020	POETRY / English, Irish, Scottish, Welsh
POE014000	POETRY / Epic *



# The world according to BISAC

## Making shelf space for Bambi & Bullwinkle

JNF003260	JUVENILE NONFICTION / Animals / Cows *
JNF003230	JUVENILE NONFICTION / Animals / Deer, Moose & Caribou
JNF003050	JUVENILE NONFICTION / Animals / Dinosaurs & Prehistoric Creatures
JNF003060	JUVENILE NONFICTION / Animals / Dogs
JNF003210	JUVENILE NONFICTION / Animals / Ducks, Geese, etc.
JNF003070	JUVENILE NONFICTION / Animals / Elephants

## ... and scrunching together Schiller, Petrarch & Verlaine

POE011000	POETRY / Canadian
POE012000	POETRY / Caribbean & Latin American
POE005030	POETRY / Continental European
POE005020	POETRY / English, Irish, Scottish, Welsh
POE014000	POETRY / Epic *

Squeezing the universal library into a suburban bookstore



## Correcting the Problem

---

Google: "We're on it (but it isn't a first priority)"

Correcting errors as noticed (like bad scans)?

Crowd Sourcing?

But errors/bad metadata affect 000,000's of records

"Error correction" doesn't address poor & missing metadata, inconsistent/confusing/inappropriate classification schemes

Why should the metadata decisions be left to Google engineers?



# Correcting the Problem

---

HathiTrust to the rescue?

But HathiTrust makes available only out-of-copyright works, has (relatively) limited computational resources

Why should Google have no obligations to do GBS right?

Google Book Search is "a tremendous public good for students, for teachers, for scholars, for everyone."

Derek Slater, Google

But a public good implies a public trust