

Indiana Authors and Their Books: The Journey from Print to Digital

Michelle Dalmau, Digital Projects & Usability Librarian, Digital Library Program

Jennifer Liss, Metadata/Cataloging Libraries, Technical Services

Digital Library Brown Bag Series, Fall 2012

September 19, 2012

It all started in 2005 ...

- George W. Bush was inaugurated for a second term
- Pope John Paul II died
- Hurricane Katrina ravaged the Gulf
- Mariah Carey's "We Belong Together" topped *Billboard's* Hot 100 Singles
- *Star Wars: Episode 3 Revenge of the Sith* was the top grossing movie
- Gas topped at \$3.00/gallon

Birth of the *Indiana Authors and Their Books* Project

- An LSTA funded project (June 2005-July 2006)
- Aimed to digitize and encode:
 - 3-volume reference work *Indiana Authors and Their Books*
 - Published by Wabash College: 1949, 1974, 1981
 - Covering: 1816-1916, 1917-1966, 1967-1980
 - Approximately 150 volumes by selected authors from Indiana's Golden Age of Literature (1880-1920) referenced in the encyclopedia

Indiana Authors: Anatomy of the Reference Work

- **Biographical entry** and a complete **bibliography** for all authors who were born, raised, or educated in Indiana, or who lived in the state for a major portion of their lives
 - ~7,000 entries of Hoosier authors from Indiana's first 100 years of statehood
 - ~21,000 citations in the bibliographies

**BLACK****BLACK, WILLIAM HENRY: 1854-1930.**

William Henry Black, son of the Rev. Felix G. and Lydia K. Frederick Black, was born at Centerville, Ind., on Mar. 19, 1854. He graduated from Waynesburg (Pa.) College in 1876, receiving the A.M. degree in 1879, and graduated from Western Theological Seminary in 1878. In 1888 he received the D.D. degree from Cumberland University, in 1903 the LL.D. from Westminster College, and in 1915 the Litt.D. from Waynesburg College. He married Mary Ella Henderson on Apr. 3, 1879.

Ordained to the Presbyterian ministry in 1876, he served as pastor at Pittsburgh from 1877 to 1881 and at St. Louis from 1881 to 1890. After 1890 he was president of Missouri Valley College.

Dr. Black died on June 22, 1930, at Marshall, Mo.

Information from *Who Was Who in America*.

* * *

Sermons for the Sabbath School. 1886.

God Our Father. *Nashville, Tenn.*, 1891.

Womanhood. *Nashville, Tenn.*, 1891.

Outline Life of Paul. 1894.

The Life and Times of Moses. 1901.

The Hebrew Monarchy—a Harmony. 1902.

The Life of Jesus: Tripaschal. 1906.

The Division and Fall of the Hebrew Monarchy—A Harmony. 1910.

**BLACKALL, EMILY LUCAS (MRS. CHRISTOPHER R.): 1832-1892.**

Emily Lucas was born in Salem, Ind., in 1832. She became the second wife of Christopher Rubey Blackall, a newspaper editor, in 1873. She died in New York in 1892.

Information from sketch of Christopher Rubey Blackall in *Who Was Who in America*.

* * *

Superior to Circumstances. *Boston*, 1889.

Stories About Jesus (with C. R. Blackall). *Philadelphia*, 1890.

Won and Not One. *Philadelphia*, 1891.

**BLACKBURN, WILLIAM MAXWELL: 1828-1898.**

"William Maxwell Blackburn, author and educator, was born at Carlisle, Ind., Dec. 30, 1828, son of Alexander and Deliah (Polk) Blackburn . . . He re-

BLACKBURN

ceived his early education in district schools and in an academy at La Porte, Ind., and taught at similar institutions previous to entering . . . Hanover College, where he was graduated, June, 1850. He then studied for the Presbyterian ministry at Princeton Theological Seminary from 1850 to 1854, and after completing his course held pastorates at Three Rivers, Mich., Erie, Pa., Trenton, N. J., Chicago, Ill., and Cincinnati, O. Beginning in his early ministry by contributions to the *PRINCETON REVIEW*, he occupied his leisure continuously with writing; and while preaching and administering the affairs of his pastorates he became known both in America and England, first, as a successful author of Sunday-school books, and later for works on subjects of wider interest. In 1868, the position was offered him of professor of biblical and ecclesiastical history in what is now McCormick Theological Seminary at Chicago . . . He accepted and filled the chair until 1881. From 1884 until 1886, he was president of the Territorial University of North Dakota, and later was appointed president of Pierre University, at East Pierre, S. D., which, in 1898, was removed to Huron, S. D.

"He received the honorary degree of D.D., from Princeton in 1870 and that of LL.D. from Wooster University . . ."

"Dr. Blackburn was married in 1854 at Valparaiso, Ind., to Elizabeth Powell . . ."

Dr. Blackburn died at Huron, S. D., on Dec. 29, 1898.

Condensed from *The National Encyclopedia of American Biography*, Vol. 9.

* * *

Uncle Alick's Stories. (A series of juveniles listed separately below) *Philadelphia*.

The Early Watermelons; or, Alick Never Afraid of the Truth.

Alick and His Blind Uncle.

Cherry-Bounce; or, Wise Management of Human Nature.

My Dog Rover, and Some Good That He Did in the World.

Teddy, the Bill-Poster, and How He Became Uncle Alick's Right-Hand Man.

Uncle Alick's Sabbath School.

Aonio Paleario and His Friends. *Philadelphia*.

The Holy Child; or, Early Years of Jesus Christ. *Philadelphia*.

Exiles of Madeira. *Philadelphia*, 1862.

Judas the Maccabee and the Asmonian Princess. *Philadelphia*, 1864.

E-Text Explained, Sort Of

The term "e-text" has been co/ad-opted by the Electronic (Text)books market (eText) and by our central IT division at Indiana University.

$(\text{eText} + \text{eBook}) = \text{e-text}$

The origin of "e-text" as in Electronic Texts has been used in the humanities computing arena since the 1970s and is speculatively attributed to Michael Hart who started Project Gutenberg .

e-text \neq eText?

TEI Refresher

Text Encoding Initiative (TEI) / *Guidelines for Electronic Text Encoding and Interchange (TEI)*

- The TEI *Guidelines* "*are addressed to anyone who works with any text in electronic form*. They provide means of representing those features of a text which need to be identified explicitly in order to facilitate processing of the text by computer programs" (Sperberg-McQueen).
- TEI provides elements, attributes, and other mechanisms for encoding prose, poetry, drama, dictionaries, critical apparatus, linguistic corpora, and other scholarly and non-scholarly texts.

Indiana Authors:

Digitization and Text Encoding

- Encyclopedia encoded following TEI P4 standard
 - Structural, Semantic and Stylistic
- Monographs encoded following the TEI Lite P4 standard, a subset of the TEI schema
 - Structural and stylistic markup
 - Genres vary: mostly prose and poetry, some dramas
- TEI and Page Images
 - Initial round of encoding/scanning was vendor outsourced
 - Second round of encoding/scanning happened in-house

Broadening Partnerships

Digital Library
Program

Cataloging
Division

Collection
Development

Technical Services beyond MARC?

A word cloud centered around the acronym 'TEI'. The word 'TEI' is the largest and most prominent, rendered in a dark green color. Surrounding it are several other acronyms in various sizes and colors: 'EAD' in dark teal, 'MODS' in brown, 'MADS' in yellow-green, 'DC' in green, 'METS' in brown, 'PBCORE' in green, 'VRA-CORE' in dark teal, 'PREMIS' in yellow-green, and 'EAC' in brown. The words are arranged in a way that they appear to be floating around the central 'TEI'.

TEI

EAD

MODS

MADS

DC

METS

PBCORE

VRA-CORE

PREMIS

EAC

Catalogers doing e-text?

`<?xml?>`

Introduce XML



Broad roll out



Familiar material format

E-text workflow: **pilot to production**

- **Summer 2009:** 11 catalogers trained
- **Spring 2012:** 4 catalogers trained; hired a student hourly
- More than half of Tech Services catalogers perform text encoding work today
- Staff cataloging job descriptions were rewritten to accommodate non-MARC metadata

Libraries are doing it for themselves: Productionizing an e-text workflow



Selection



Scanning



Encoding



Publication



Selection

- Selection initially done by the collection manager
- In the new workflow, selection became more complicated!



Scanning

- DLP completed scanning in-house
- OCR is run to derive text from scanned images
- Result: a TEI shell (with dirty OCR) and...



Encoding

- ... MARC data! TEI header metadata was derived from MARC data
- Complies with Level 3 encoding, as defined in *Best Practices for TEI in Libraries (version 3.0)*



Publishing

- Once published, these electronic texts are then cataloged as e-books, links can be accessed via IUCAT and WorldCat

Indiana Authors Web Site Design: *Many* False Starts

- 2007
- 2008
- 2009
- Finally 2011!

2007

Indiana Authors and Their Books

Home | Browse | Simple Search | Advanced Search | About | Help Search Encyclopedia:

Advanced Search

Search the Encyclopedia

Search the Biography

<input type="text"/>	Keyword	▼	And	▼
<input type="text"/>	Names	▼	And	▼
<input type="text"/>	Organizations	▼		
<input type="button" value="Search"/>				

Locations
Country
State
Cities
Event Dates
Birth Dates
Death Dates

Search the Bibliography

<input type="text"/>	Keyword	▼	And	▼
<input type="text"/>	Authors	▼	And	▼
<input type="text"/>	Titles	▼		
<input type="button" value="Search"/>				

Editors
Publication Date

Search Tips

Search tips here

Search the Books

<input type="text"/>	Keyword	▼	And	▼
<input type="text"/>	Author	▼	And	▼
<input type="text"/>	Title	▼		
<input type="button" value="Search"/>				

2008

IU Header

[banner image]

Home | Browse | Search | Advanced Search | Project Information | Help Search

Advanced Search

Search for:

keyword and

keyword and

keyword or not

Limit search to full-text books

Publication Year:

All Years to [] [yyyy]

Search Tips

[Search tip here]

[Search tip here]

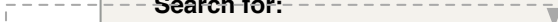
[Search tip here]

[Search tip here]

Version 1.0: Released 23 July 2008
Last Updated: date/time stamp
URL: <http://www.url.edu>
[IU Digital Library Program](#), [IUPUI Digital Library Program?](#)
Comments: diglib@indiana.edu

IU Footer

keyword
author
title



2009

Indiana Authors

Search Results (title)

IU Header

[banner image]

Home | Browse | Search | Advanced Search | Project Information | Help

33 results for Sunshine (title)

Sort by: Results per page:

<< Previous 1 2 3 4 5... Next >>

1. [Stone, Grace \(1871-1945\)](#)
Title: [Eternal **Sunshine** of the Spotless Mind, 1930](#) ▶
Source: Indiana authors and their books 1967-1980, volume 1
2. [Astaire, George \(1884-1957\)](#)
Title: [Sunshine and Rain, 1920](#) Search title in:
IUCAT
Google Books
Worldcat Local
Source: Indiana authors and their books 1967-1980, volume 3
3. [Veer, Sonja \(1840-1921\)](#)
Title: [Sunshine, 1900](#) ▶
Source: Indiana authors and their books 1967-1980, volume 1

<< Previous 1 2 3 4 5... Next >>

Version 1.0: Released 23 July 2008
Last Updated: date/time stamp
URL: <http://www.iu.edu>
IU Digital Library Program, IUPUI Digital Library Program?
Comments: diglib@indiana.edu

IU Footer

Publication Date
Relevance

10
20
50

2011

Search Browse: [Author](#) | [Title](#) | [Year](#)

Search for: My Selections (0)

Results: 247 item(s) Sort by: [Relevance](#)

Page: 1 2 3 4 5 Next Results per page: 20

Format

[Book](#) (230)
[Encyclopedia Entry](#) (17)

ENCYCLOPEDIA ENTRIES

Author

[ADE, GEORGE, 1866-1944](#)
[BIDDLE, HORACE P. \(HORACE PETERS\), 1811-1900](#)
[DREISER, THEODORE, 1871-1945](#)
[EGGLESTON, GEORGE CARY, 1839-1911](#) [more](#)

BOOKS

Publication Year

[1900-1909](#) (90)
[1910-1919](#) (71)
[1880-1889](#) (28)
[1890-1899](#) (26)
[1870-1879](#) (13) [more](#)

Author

[Ade, George, 1866-1944](#) (19)
[McCutcheon, George Barr,](#)

Book

Title: [Last poems](#)
Author: Biddle, Horace P. (Horace Peters), 1811–1900.
Publication Year: 1882
Source: Cincinnati: Robert Clarke, 1882. 435 p. port. 21 cm.
[\[Add to My Selections\]](#)

Search terms in context (787):
...MAKE A YOUNG LOVER OF ME 265 YOU SAY YOU LOVE ME 178 YOUNG LUCY 366 YOUNG ROBBIE 184 PUBLICATIONS OF...
...Mary R.) My Three Angels: Faith, Hope, and **Love**. With full-page illustration. By E. D. Grafton....
...rules her breast. Auburn ringlets, laughing **love**, Soft as winglets of the dove. Eyes, whose flashes...

Encyclopedia Entry

Author: [BIDDLE, HORACE P. \(HORACE PETERS\), 1811-1900.](#)
Source: Volume 3, p. 244.
[\[Add to My Selections\]](#)

Book Titles:
[American Boyhood >>](#)
[Book Not Digitized #1 >>](#)
[Book Not Digitized #2: The Sequel >>](#)
[Elements of Knowledge >>](#)
[Glances at the world >>](#)
[more](#)

Book

Title: [Glances at the world](#)
Author: Biddle, Horace P. (Horace Peters), 1811–1900

Web Design Challenges

- Mixed Content Types
 - Different levels of encoding (whole/parts)
 - Different user expectations for access
- Only small subset of citations available in full text
- Author name normalization across encyclopedia entries (i.e., “transcribed”) and bibliographic access (i.e., “controlled”)

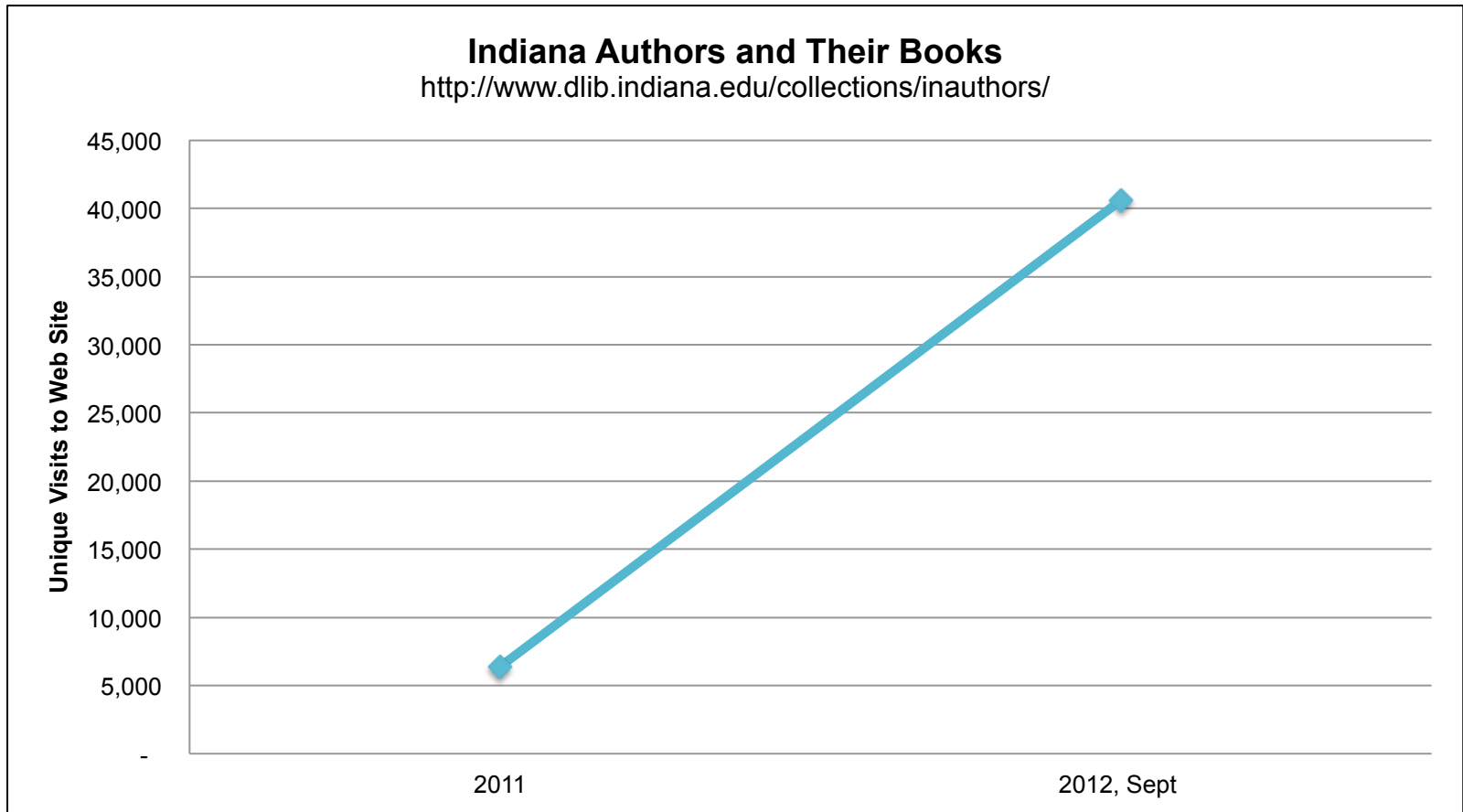
Indiana Authors Demo

- Web site launched in 2 phases:
 - Fall 2011 (monographs-only)
 - April 2012 (encyclopedia & monographs)
- Indiana Authors and Their Books:
 - <http://dlib.indiana.edu/collections/inauthors>

Was it Worth It?



Indiana Authors Web Stats





Web Publishing



- Super easy! Seamless integration with text encoding workflow
- Streamlined web publishing process



Text Encoding



Submit

View Collection inauthorsbooks

Download | Preview | Validate | Delete | View All Versions | Submit File | Clear Filters | Change Record Status | Add to File Bag | File Bag (0 files) | Change Collection

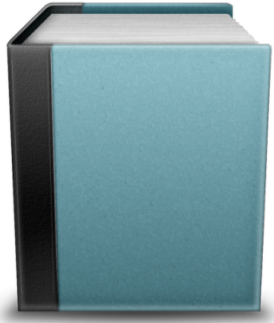
Submitted By	Submission Date	Record Status	Revision	Log	File
bbeamn	09/18/2012	in progress	6	In process	VAC1011
bgucinsk	09/17/2012	in progress	3	p. 120 of 485	VAC0874
bgucinsk	09/14/2012	pending approval	2	p. 20 of 20; ready for review.	VAC0995
sulankec	09/14/2012	in progress	23	done thru p. 136	VAC1311
bgucinsk	09/14/2012	completed	8	Correction made in teiHeader.	VAC0870
tlarson	09/14/2012	in progress	4	In progress - Tim	VAC1006
pcook	09/14/2012	in progress	13	in progress	VAC1130
jlorenz	09/13/2012	in progress	3	p. 25	VAB8218
bgucinsk	09/12/2012	completed	6	p. 376 of 376; ready for review.	VAC0996
tuckerru	09/12/2012	in progress	21	encoding	VAC1014
emiwyat	09/07/2012	in progress	19	in progress	VAB9592
racwheel	09/05/2012	in progress	2	"Rural" is a hard word to pronounce.	VAC0873
jutanaka	08/30/2012	in progress	2	p. 47 of 65	VAC0883
xihu	08/29/2012	in progress	8	p. 218	VAC0992
salibicr	08/28/2012	in progress	3	In Progress	VAC2098
racwheel	08/27/2012	completed	2	The cataloguing department has overflowed from its retired quarters into the reading room spaces, where it is constantly subject to the	VAC2100

Page 1 of 16

Displaying files 1 - 25 of 399

DLP Indiana University

Submit was created by the [Indiana University Digital Library Program](#)
For more info please contact diglib@indiana.edu



Selection



- Should we care if Google has already digitized the book?
- Should we have partnered with other libraries or cultural heritage institution with the holdings?

WHAT ABOUT  ?

THE HUNGER GAMES

Play video




New York Times bestselling author


SUZANNE COLLINS

Introducing Google Play

Now your favorite music, books, magazines, movies, TV shows, apps, and games are all in one place that's accessible from the Web and any Android device. [Explore what's new](#)

GO TO GOOGLE PLAY

 movies & tv

 music

 apps and games

 books

 magazines

[Visit Google Play now](#)



How to be a good e-text vendor.



An e-text vendor doing it wrong.

1	Title	Author	Replacement
2	Husks and nubbins	Greenwood, Alice D. O. (Alice Davis Odekirk), 1851-1936.	\$ 50.00
3	Lyrics of war and peace	Foulke, William Dudley, 1848-1935.	\$ 95.13
4	The path of glory	Haworth, Paul Leland, 1876-1936.	
5	Maya; a drama	Foulke, William Dudley, 1848-1935.	\$ 25.00
6	The redemption of Anthony	Cooke, Marjorie Benton, 1876-1920.	\$ 40.00
7	A bit of finesse; a story of fifty years ago	Lodge, Harriet Newell, 1848-	\$ 25.00
8	Roach & Co.--pirates : and other stories	Fuller, Hector, 1856-	\$ 35.00
9	Amusement of idle hours. The poems of S. Attwood Butterfield	Butterfield, Seymour Attwood.	\$ 47.00
10	The white islander	Catherwood, Mary Hartwell, 1847-1902.	\$ 82.83
11	When the old flag came, and other poems	Dobbins, Douglas, 1860-1927.	\$ 10.00
12	Rocky fork	Catherwood, Mary Hartwell, 1847-1902.	\$ 115.13
13	Protean papers	Foulke, William Dudley, 1848-1935.	\$ 25.00
56	The kid has gone to the colors, and other verse	Herschell, William.	\$ 83.00
57			
58			\$ 1,719.65

What do you think?



THANKS!