



MARC Records with ChatGPT: Lessons Learned

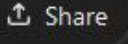

John Kimbrough
Georgetown University Library
DC Ex Libris RUG - October 2024

GEORGETOWN
UNIVERSITY


LIBRARY

Roadmap

- Creating MARC records with ChatGPT-4: two book examples
- Lessons Learned
- AI in Alma:
 - E-Book Metadata
 - Metadata Assistant
- Some Questions to Ponder




ChatGPT 4o  

Thank you. Can you access current Library of Congress Authority headings? If so, what website(s) do you consult?


 Yes, I can access current Library of Congress Authority headings. The primary source for this information is the Library of Congress Authorities website, which provides access to the Library of Congress Name Authority File (NAF), Subject Authority File (SAF), and other authority records. You can access this information through the following websites:

1. [Library of Congress Authorities](#)
2. [Library of Congress Linked Data Service](#)

If you have specific headings you need to look up or any other questions related to Library of Congress Authority headings, please let me know!

Thank you. If I provided a text string, could you tell me if it was a valid LCSH?

 Yes, I can help you determine if a text string is a valid Library of Congress Subject Heading (LCSH). Please provide the text string you would like to check, and I'll look it up for you.

Inputs: Titles to ChatGPT

- E-books (3 titles)
 - 2 non-fiction + poetry
- Print books/items (5 titles)
 - GU Press book with CIP data
 - DVD set
 - Opera program
 - Museum catalog
 - M.A. thesis
- Russian citations (3 citations)
 - Author, title, publisher in Cyrillic

ChatGPT queries:

- Citation only
- Citation + PDF of front matter, TOC, some body pages (not entire book)
- Citation + descriptive summary (from publisher)

Total: 11 titles, 22 queries

“Temporary mode” = ChatGPT doesn’t remember previous queries

Inputs: Prompt

- Same / similar prompt for all books
- No iteration

Adapted from:

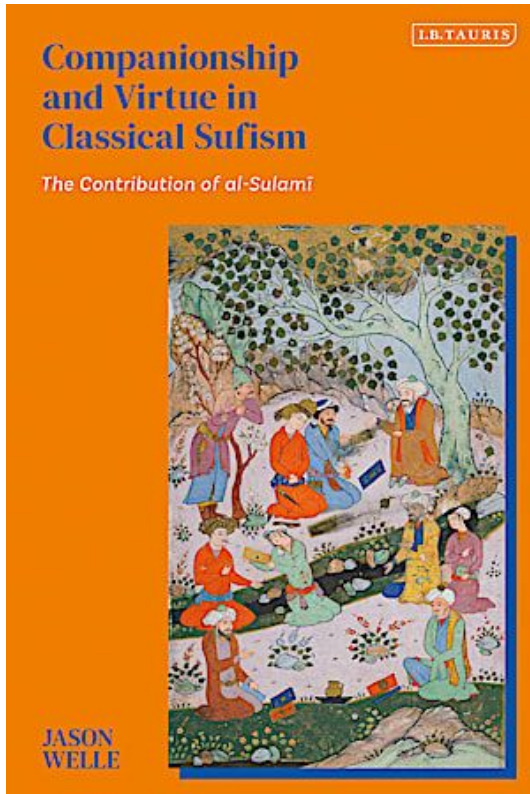
Chow, E. H. C., Kao, T. J., & Li, X. (2024). An Experiment with the Use of ChatGPT for LCSH Subject Assignment on Electronic Theses and Dissertations. arXiv preprint arXiv:2403.16424

As a library cataloger, please create a MARC record for the item listed below. Please provide output according to the following instructions:

- Format the MARC record using Program for Cooperative Cataloging (PCC) standards, including RDA (Resource Description and Access) tags.
- List the cataloging agency in 040 as Georgetown University (DGU).
- Provide names in the form authorized by Library of Congress Name/Title Authority Headings, at authorities.loc.gov.
- Provide subject headings and notes in English, regardless of the book's language.
- Provide subject headings in the form authorized by Library of Congress Subject Headings, at authorities.loc.gov.
- Provide a Library of Congress call number.
- Provide the MARC fields only, without explanations.
- Please add spaces for clarity between field tags, indicators, subfield codes, and associated text.

[BOOK INFORMATION]

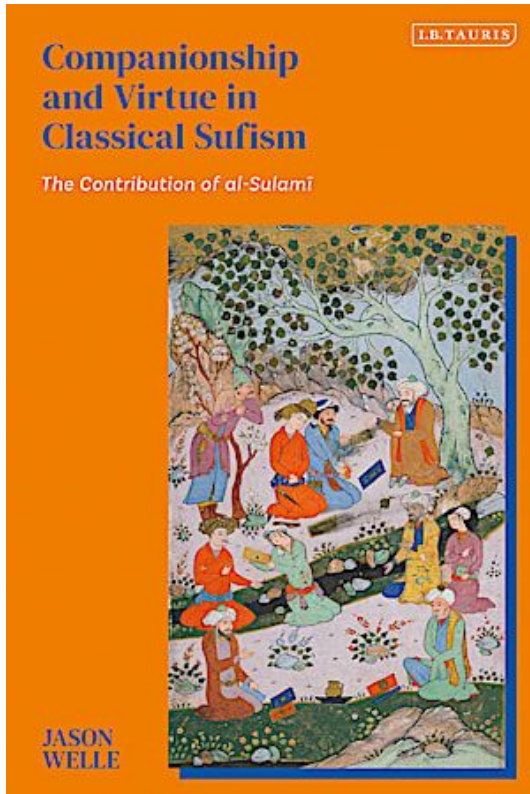
Results from ChatGPT: E-Book



```
=LDR 00000nam a2200000 i 4500
=001 ocn123456789
=003 OCoLC
=005 20240705123456.0
=008 240705s2024 enka b 001 0 eng d
=020 \\$a9780755652273
=035 \\$a(OCoLC)123456789
=040 \\$aDGU$beng$erda$cDGU
=041 0\\$aeng
=042 \\$apcc
=043 \\$aa-ii---
=050 00$aBP189.7.C56$bW45 2024
=082 04$a297.4/8$223
=100 1\\$aWelle, Jason,$eauthor.
=245 10$aCompanionship and virtue in classical Sufism :$bthe contribution of al-Sulamī /$cJason Welle.
=246 3\\$aContribution of al-Sulamī
=264 1$aLondon :$bI.B. Tauris,$c2024.
=300 \\$axi, 214 pages :$billustrations ;$c24 cm
=336 \\$atext$btxt$2rdacontent
=337 \\$aunmediated$bn$2rdamedia
=338 \\$avolume$bnc$2rdacarrier
=490 1\\$aBloomsbury collections
=504 \\$aIncludes bibliographical references (pages 195-208) and index.
```

```
=505 0\\$aIntroduction -- Companionship and community -- Virtue and character -- Conclusion.
=520 \\$a" This study focuses on the Nishapurian shaykh Abū ‘Abd al-Raḥmān al-Sulamī (d. 412/1021), highlighting his notion of companionship (ṣuḥba) and its importance in the spiritual growth within Sufism."
=600 00$aSulamī, Abū ‘Abd al-Raḥmān,$d937-1021.
=650 \0$aSufism$xHistory$yEarly works to 1800.
=650 \0$aSpirituality$xIslam$xHistory$yEarly works to 1800.
=650 \0$aIslamic ethics$xHistory$yEarly works to 1800.
=650 \0$aCompanionship$xReligious aspects$xIslam.
=700 1\\$aSulamī, Abū ‘Abd al-Raḥmān,$d937-1021,$eauthor.
=710 2\\$aGeorgetown University,$eissuing body.
=830 \0$aBloomsbury collections.
=856 40$uhttp://dx.doi.org/10.5040/9780755652303$zConnect to electronic resource
```

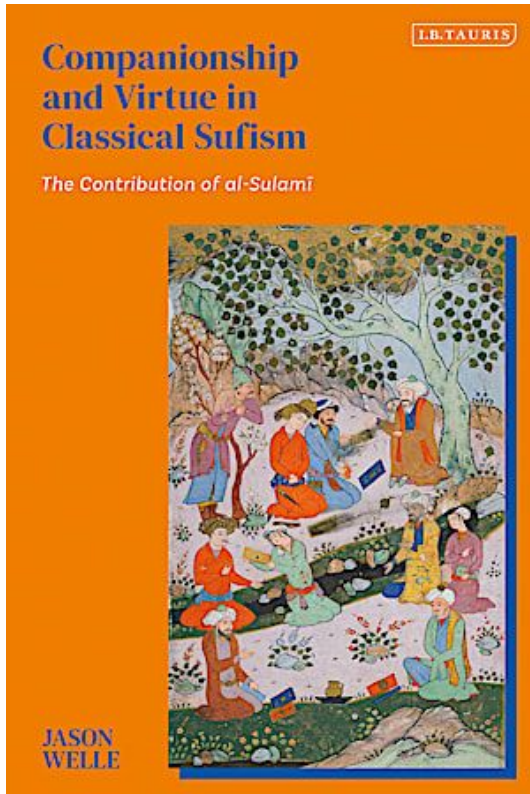
Results from ChatGPT: E-Book



=245 10\$a Companionship and virtue in classical Sufism
: \$b the contribution of al-Sulamī /\$cJason Welle.

=246 3\$a Contribution of al-Sulami

Results from ChatGPT: E-Book

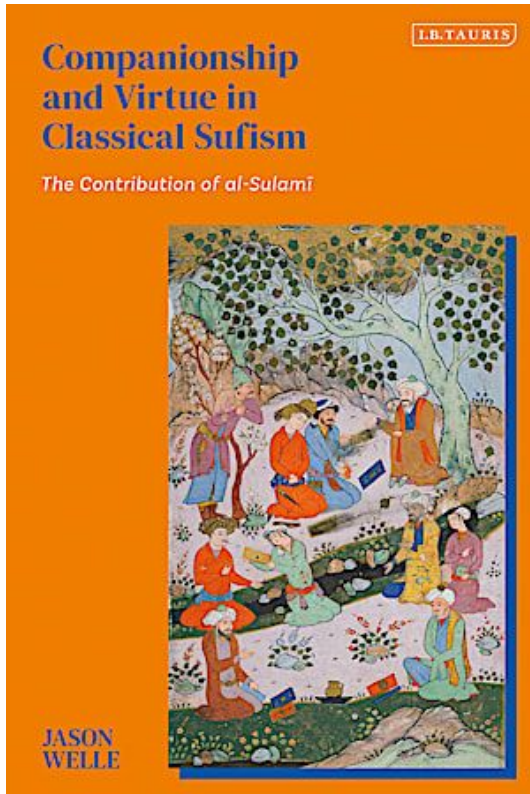


=245 10\$a Companionship and virtue
: \$b the contribution of al-Sulamī
=246 3\$a Contribution of al-Sulamī

=505 0\$a Introduction --
Companionship and community --
Virtue and character --
Conclusion.

Front matter	
Dedication	v–vi
Acknowledgements	ix–x
Transliteration, Dating, Primary Sources and Abbreviations	xi–xiv
Introduction: What does virtue ethics have to do with Sufism?	1–12
1. The emergence of <i>ādāb</i> in al-Sulamī’s spiritual method	13–44
2. Virtue and character in al-Sulamī’s thought ..	45–70
3. Al-Sulamī’s vision of fellowship and spiritual companionship	71–100
4. <i>Ṣuḥḥaba</i> in comparison: al-Sulamī’s near contemporaries	101–126
Conclusion	127–132
Back matter	
Bibliography	189–214

Results from ChatGPT: E-Book



```
=245 10$a Compa  
: $b the contrib  
=246 3$a Contr
```

```
=505 0$a Intro  
Companionship an  
Virtue and chara  
Conclusion.
```

```
=650 \0$a Sufism $x History $y Early works to 1800.  
=650 \0$a Spirituality $x Islam $x History $y Early works to 1800.  
=650 \0$a Islamic ethics $x History $y Early works to 1800.  
=650 \0$a Companionship $x Religious aspects $x Islam.
```

The Library of Congress >> [Go to Library of Congress Online Catalog](#)

LIBRARY OF CONGRESS AUTHORITIES

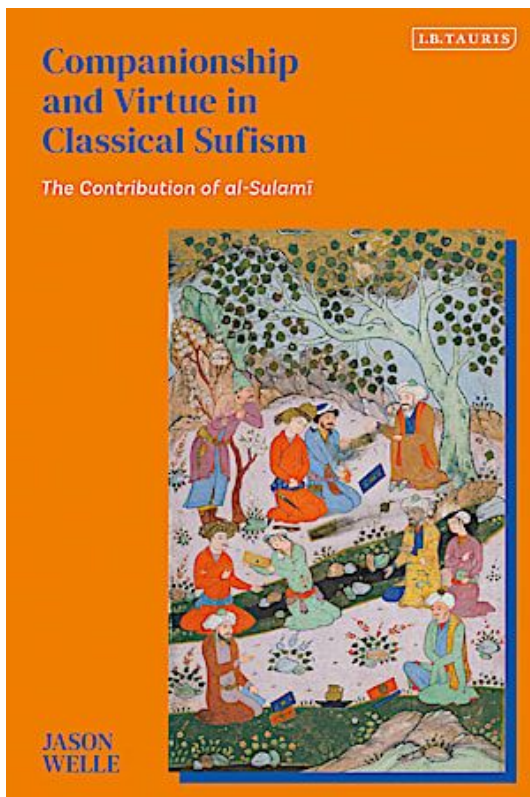
[Help](#) [New Search](#) [Search History](#) [Headings List](#) [Start Over](#)

[← Previous](#) [Next →](#)

[MARC Display](#) [Labelled Display](#)

LC control no.: sh 85047713
LCCN Permalink: <https://lcn.loc.gov/sh85047713>
Topical subject heading: Fellowship
Variant(s): Companionship
Comradeship
See also: [Friendship](#)

Results from ChatGPT: E-Book



=245 10\$a Companionship and virtue in classical Sufism
: \$b the contribution of al-Sulamī /\$cJason Welle.

=246 3\$a Contribution of al-Sulami

=505 0\$a Introduction --
Companionship and community --
Virtue and character --
Conclusion.

Back matter

Bibliography 189-214

=650 \0\$a Sufism \$x History \$y Early works to 1800.

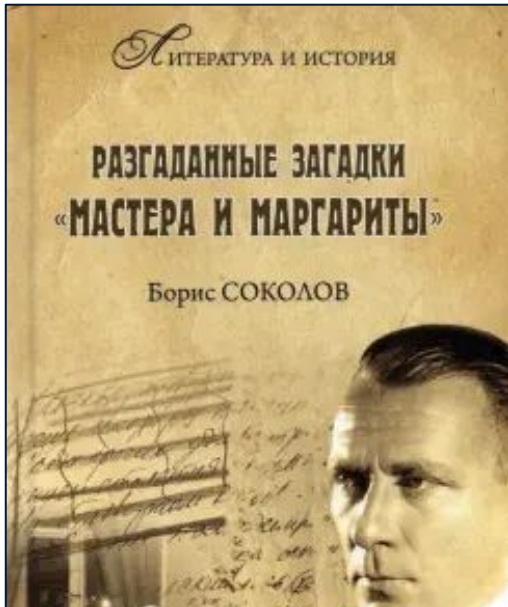
=650 \0\$a Spirituality \$x Islam \$x History \$y Early works to 1800.

=650 \0\$a Islamic ethics \$x History \$y Early works to 1800.

=650 \0\$a Companionship \$x Religious aspects \$x Islam.

=504 \\\$aIncludes bibliographical references (pages 195-208) and
index.

Results from ChatGPT: Russian book (citation only)

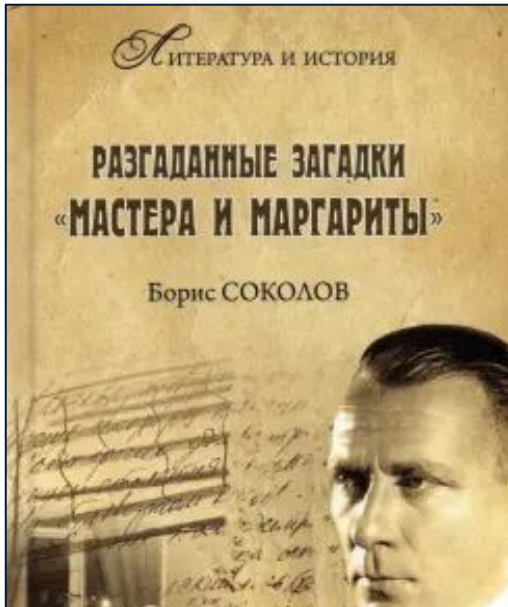


Title: Разгаданные загадки "Мастера и Маргариты"
Author: Борис Соколов.
Publisher: Москва : "Вече"

=001 123456789
=003 DLC
=005 20240726
=008 240726s2023 ru a 000 0 rus d
=040 \\\$a DGU \$b eng \$e rda \$c DGU
=041 1\\\$a rus \$h rus
=100 1\\\$6880-01\$a Соколов, Борис.
=100 1\\\$a Sokolov, Boris.
=245 10\$6880-02\$a Разгаданные загадки "Мастера и Маргариты" / \$c Борис Соколов.
=245 10\$a Razgadannye zagadki "Mastera i Margarity" / \$c Boris Sokolov.
=250 \\\$a Первое издание.
=264 \1\$6880-03\$a Москва : \$b "Вече", \$c 2023.
=264 \1\$a Moskva : \$b "Veche", \$c 2023.
=300 \\\$a 320 pages : \$b illustrations ; \$c 24 cm.
=336 \\\$a text \$b txt \$2 rdacontent
=337 \\\$a unmediated \$b n \$2 rdamedia
=338 \\\$a volume \$b nc \$2 rdacarrier
=504 \\\$a Includes bibliographical references (pages 315-320).

=650 \0\$a Bulgakov, Mikhail, \$d 1891-1940. \$t Master i Margarita \$x Criticism, interpretation, etc.
=650 \0\$a Russian literature \$y 20th century \$x History and criticism.
=700 1\\\$6880-04\$a Соколов, Борис.
=700 1\\\$a Sokolov, Boris.
=880 1\\\$6100-01/\$1\$a Соколов, Борис.
=880 10\$6245-02/\$1\$a Разгаданные загадки "Мастера и Маргариты" / \$c Борис Соколов.
=880 \1\$6264-03/\$1\$a Москва : \$b "Вече", \$c 2023.
=880 1\\\$6700-04/\$1\$a Соколов, Борис.
=942 \\\$2 pcc
=952 \\\$a DGU
=955 \\\$a bsg 20240726

Results from ChatGPT: Russian book (citation only)



Title: Разгаданные загадки "Мастера и Маргариты"
Author: Борис Соколов.
Publisher: Москва : "Вече"

```
=100 1\${6880-01}\$a Соколов,  
Борис.
```

```
=100 1\${a} Sokolov, Boris.
```

```
=300 \${a} 320 pages : \$b illustrations  
; \$c 24 cm.
```

```
=504 \${a} Includes bibliographical  
references (pages 315-320).
```

```
=100 1\${6880-01}\$a Соколов, Борис.
```

```
=245 10\${6880-02}\$a Разгаданные загадки "Мастера и Маргариты" / \$c  
Борис Соколов.
```

```
=264 \1\${6880-03}\$a Москва : \$b "Вече", \$c 2023.
```

```
=880 1\${6100-01}/\$1\$a Соколов, Борис.
```

```
=880 10\${6245-02}/\$1\$a Разгаданные загадки "Мастера и Маргариты" / \$c  
Борис Соколов.
```

```
=880 \1\${6264-03}/\$1\$a Москва : \$b "Вече", \$c 2023.
```

```
=942 \2 pcc
```

```
=952 \a DGU
```

```
=955 \a bsg 20240726
```

Lessons Learned from ChatGPT's MARC records

ChatGPT will produce MARC records, but with problems:

- Hallucinations
- Partial/incomplete information
- MARC rule violations
- Subject heading difficulties

Is ChatGPT a useful metadata tool? It depends...

...but you won't save time using ChatGPT

AI in Alma: E-book metadata enrichment

“AI Bibliographic Records Enrichment”

- Language (008/041)
- Table of Contents (505)
- Summary (520)
- LCSH (650)
- ~8,000 titles/week
- If record already has fields, AI skips
- Enriched fields have |7 “Generated by AI”
- 588|a “Part of the metadata in this record was created by AI, based on the text of the resource”

Subject	Dutch language -- Syntax > Dutch language > Linguistics >
Other title	Syntax of Dutch
Source	Library Catalog
Other note	Part of the metadata in this record was created by AI, based on the text of the resource.
Identifier	ISBN : 90-485-2484-9

https://knowledge.exlibrisgroup.com/Content/Knowledge_Articles/Alma/Knowledge_Articles/AI_Bibliographic_Records_Enrichment

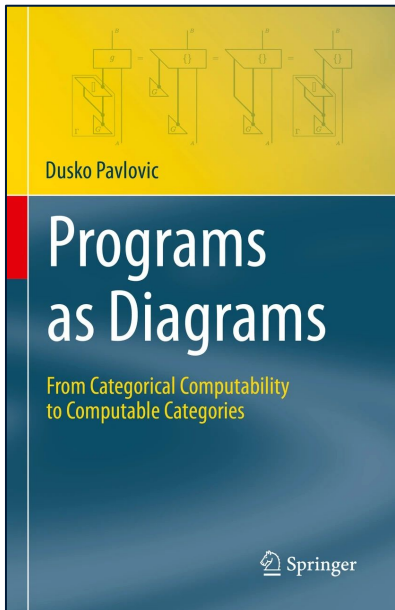
AI in Alma: E-book metadata enrichment

Spot-checked 51 records with AI enrichment tags:

- Language (008/041) = 0 records
- Table of Contents (505) = 33 records
- Summary (520) = 48 records
- LCSH (650) = 46 records (one or more 650)

AI in Alma: E-book metadata enrichment

Table of Contents often truncated (20 records of 33)



Contents
Preface -- What? -- Contents -- 1 Drawing Types and Functions -- 1.1 Types as Strings: Data Passing Without Variables -- 1.1.1 Types -- 1.1.2 Strings Instead of Variables -- 1.1.3 Tupling and Switching -- 1.2 Data Services: Copying and Deleting, Cartesianness -- 1.3 Monoidal Functions as Boxes -- 1.3.1 Composing Functions -- 1.3.2 The Middle-Two-Interchange Law -- 1.3.3 Sliding Boxes -- 1.4 Cartesian Functions as Boxes with a Dot -- 1.5 Categories: Universes of Types and Functions -- 1.6 Workout -- 1.7 Stories -- 1.7.1 Prehistory of Types -- 1.7.2 Categories as Type Universes -- 1.7.3 Logics of Types -- 1.7.3.1 Propositions-as-Types -- 1.7.3.2 Cartesian-Closed Categories -- 1.7.4 Categorical Diagrams: Chasing Arrows and Weaving Strings -- 1.7.4.1 Arrow Diagrams -- 1.7.4.2 String Diagrams -- 2 Monoidal Computer: Computability as a Structure -- 2.1 Computer as a Universal Machine -- 2.2 Universality of Program Evaluation -- 2.2.1 Program Evaluators -- 2.2.1.1 Definition

Subject
C
P

“Table of Contents...will have up to 30 entries...”



Table of contents (8 chapters)

Front Matter
Pages i-xvii

Drawing Types and Functions
Dusko Pavlovic
Pages 1-24

Monoidal Computer: Computability as a Structure
Dusko Pavlovic
Pages 25-54

Fixpoints
Dusko Pavlovic
Pages 55-73

What Can Be Computed
Dusko Pavlovic
Pages 75-101

What Cannot Be Computed
Dusko Pavlovic
Pages 103-113

Computing Programs
Dusko Pavlovic
Pages 115-143

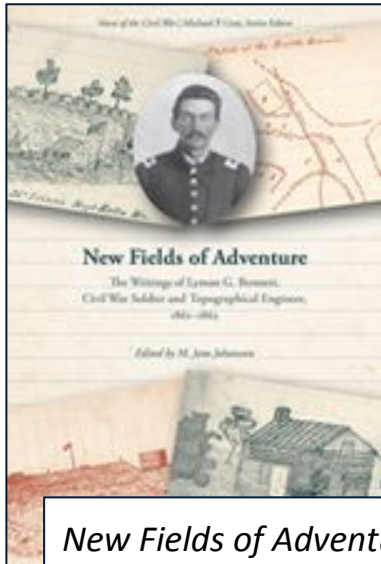
Stateful Computing
Dusko Pavlovic
Pages 145-164

Program-Closed Categories: Computability as a Property
Dusko Pavlovic
Pages 165-190

Back Matter
Pages 191-252

AI in Alma: E-book metadata enrichment

Subject headings accurate, but general - very few subdivisions



New Fields of Adventure: The Writings of Lyman G. Bennett, Civil War Soldier and Topographical Engineer, 1861-1865

Subject	Topographical surveying > Military engineering >
Source	Library Catalog
Other note	Description based on publisher supplied metadata and other sources. Part of the metadata in this record was created by AI, based on the text of the resource.
Identifier	ISBN : 1-62190-862-3

LIBRARY | CATALOG
LIBRARY OF CONGRESS

Bennett, L. G. (Lyman G.)--Diaries.
United States. Army. Illinois Infantry Regiment, 36th (1861-1865)--Biography.
United States. Army--Military life--History--19th century.
Soldiers--United States--Diaries.
Surveyors--United States--Diaries.
United States--History--Civil War, 1861-1865--Personal narratives.
West (U.S.)--History--Civil War, 1861-1865--Campaigns.
West (U.S.)--Description and travel.

The Subject Headings are generated from the book's text, and are formalized to align with the LCSH vocabulary... While in some cases the subject heading will contain subdivisions (subfields \$v, \$x, \$y, \$z), in most cases it will only have the topical term (subfield \$a).

AI in Alma and ChatGPT - Some questions to ponder

- Limited description + more accurate (Alma) vs. richer description + less accurate (ChatGPT)?
- Explicit limits/instructions to AI (e.g. “only do 650\$a”) - helpful or unhelpful?
- Are there ways to productively use AI in metadata creation, without a lot of proofreading?

“Rather than starting with a blank slate which can be filled in from the start by a trained cataloger with the item in hand, the [article] author would have us start with an error-riddled tangle of probability-predicted words and spend our time instead unpicking the mistakes of a generative language model.”

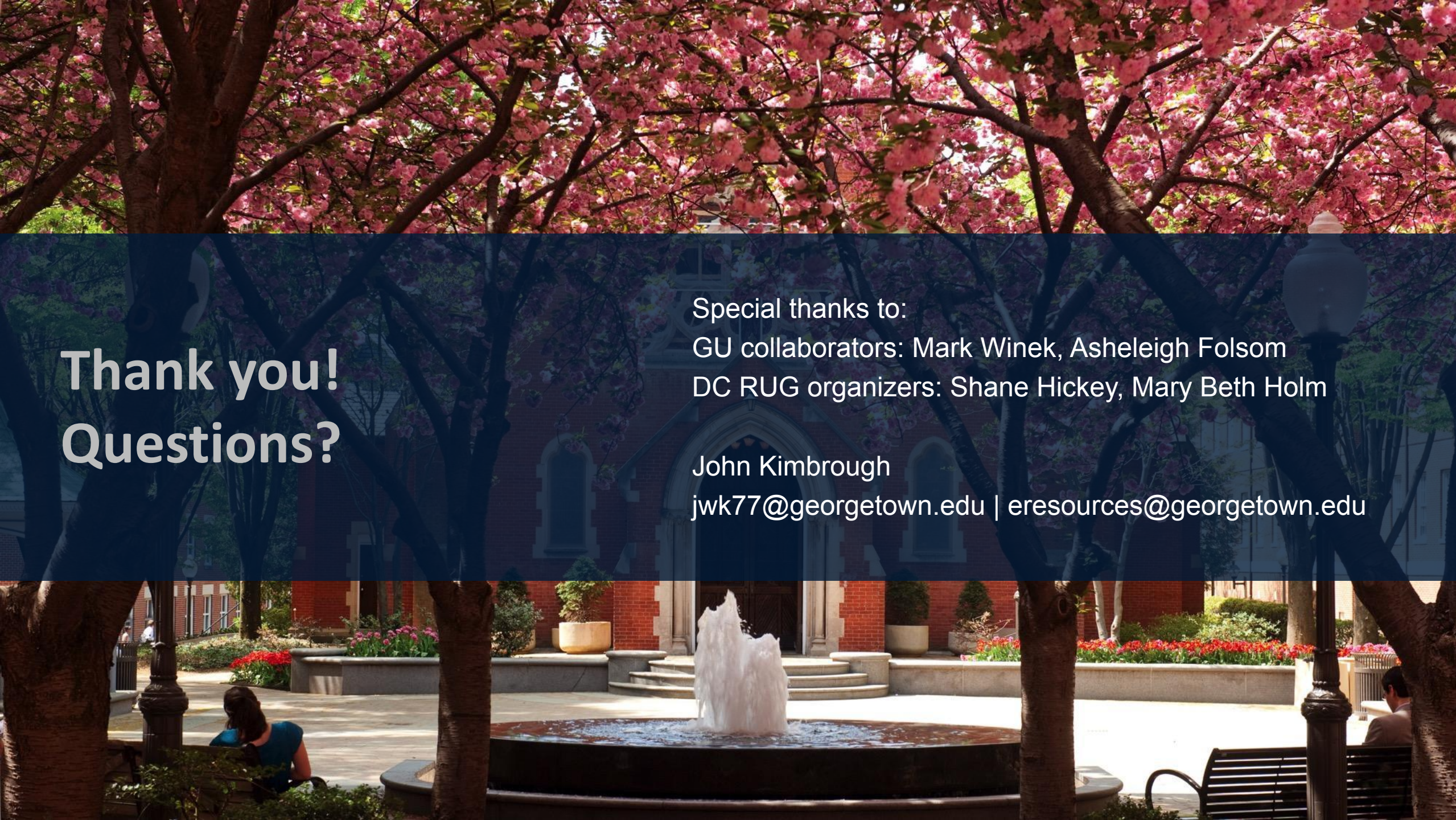
Amram, T., Malamud, R. G., & Hollingsworth, C. (2023). Response to "From ChatGPT to CatGPT". *Information Technology and Libraries*, 42(4).
<https://doi.org/10.5860/ital.v42i4.16983>

A final question - about local fields in ChatGPT records...

```
=942  $$2 pcc  
=952  $$a DGU  
=955  $$a bsg 20240726
```

```
=906  __ |a 0 |b cbc |c orignew |d u |e ncip |f 20 |g y-gencatlg  
=955  __ |a pc10 20240725 clw
```

- Were existing MARC records from one or more libraries part of ChatGPT's training data?



Thank you!
Questions?

Special thanks to:

GU collaborators: Mark Winek, Asheleigh Folsom

DC RUG organizers: Shane Hickey, Mary Beth Holm

John Kimbrough

jwk77@georgetown.edu | eresources@georgetown.edu