

JESSICA YI-YUN CHENG

May 2025

yyun.cheng@rutgers.edu

<https://yyunyc2.github.io/>

4 Huntington St., New Brunswick, NJ 08901

EDUCATION

- 2022 **PhD**, Information Sciences, University of Illinois at Urbana-Champaign, USA
-**Dissertation**: *Agreeing to Disagree: Applying a Logic-based Approach to Reconciling and Merging Multiple Taxonomies*
-**Committee**: Bertram Ludäscher (chair), Nico Franz, Allen Renear, Karen Wickett
- 2015 **Master of Arts**, Library and Information Science, National Taiwan University, Taiwan
-**Thesis**: A Study on the Best Practice for Constructing a Cross-lingual Geospatial Information Ontology
-**Committee**: Sherry (Hsueh-Hua) Chen (chair), Kuang-Hua Chen, Chih-Hong Sun,
- 2013 **Bachelor of Arts**, Library and Information Science, National Taiwan University, Taiwan

ACADEMIC APPOINTMENTS

- Sept 2022-present **Assistant Professor**
Department of Library and Information Science, School of Communication and Information, Rutgers, the State University of New Jersey, New Brunswick, NJ, USA

PUBLICATIONS

Books and Book Chapters

- [b3] Bettivia, R.*, **Cheng, Y.-Y.***, Gryk, M*. (under contract & in progress). *Trust, time, and authority in provenance: stories and echoes*. Cambridge University Press. [*Equal contributions]
- [b2] Bettivia, R.*, **Cheng, Y.-Y.***, Gryk, M.* (2022). *Documenting the future: Navigating provenance metadata standards*. Synthesis lectures on information concepts, retrieval, and services. SpringerNature. <https://doi.org/10.1007/978-3-031-18700-1> [*Equal contributions]
- [b1] **Cheng, Y.-Y.** (Accepted). Applying a taxonomy alignment method to reconcile multiple U.S. maps. *Sage research methods: data and research literacy*. Sage.

Refereed Journal Articles

- [j8] **Cheng, Y.-Y.***, & Dinh, L*. (2025). Extracting geographic relations from large social media text data. *International Journal of Geographic Information Science (IJGIS)*, 1-23. <https://doi.org/10.1080/13658816.2025.2510419> [*Equal contributions]
- [j7] **Cheng, Y.-Y.** & Choi, I. (Accepted). From Editorial Records to Structured Provenance Information: Documenting Warrant in Knowledge Organization Systems Using Large Language Models. Knowledge Organization.
- [j6] Bettivia, R.*, **Cheng, Y.-Y.*** & Gryk*, M. (2025). Usage of the term provenance in LIS literature. *Annual Review of Information Science and Technology (ARIST)*. 1-31. <https://doi.org/10.1002/asi.25015> [*Equal Contributions]
- [j5] **Cheng, Y.-Y.**, & Dinh, L. (2025). An experiment on the impact of relation types towards taxonomy alignment problems. *Information Processing & Management Journal (IP&M)*, 62(3), 1-17. <https://doi.org/10.1016/j.ipm.2024.104036>

- [j4] Choi, I.*, & **Cheng, Y.-Y.***. (2024). A conceptual model for tracking the provenance of activities in knowledge organization systems. *Journal of Documentation*. Vol. ahead-of-print No. ahead-of-print. 1-21. <https://doi.org/10.1108/JD-05-2024-0101> [*Equal contributions]
- [j3] **Cheng, Y.-Y.** (2023). Under Whose Wings? A conceptual model for incorporating historical sovereignty information in biodiversity data. *Journal of the Association for Information Science and Technology*. 1-19. <http://doi.org/10.1002/asi.24848>
- [j2] **Cheng, Y.-Y.**, Xia, Y. (2023). A systematic review of methods for aligning, mapping, merging taxonomies in information sciences. *Journal of Documentation*, 76 (6), 1413-1439. <https://doi.org/10.1108/JD-01-2023-0003>
- [j1] Thomer, A., **Cheng, Y.-Y.**, Schneider J., Twidale, M., Ludäscher, B. (2017). Towards alignment of natural history museum database schemas using a logic-based reasoning tool. *Knowledge Organization*, (44) 7, 545-558. <https://doi.org/10.5771/0943-7444-2017-7-545>

Refereed conference papers in proceedings

- [c12] Bettivia, R.*, **Cheng, Y.-Y.***, & Gryk, M.* (2023). I got a letter from my past self: (un)managed change and provenance. *iPRES 2023 Conference Proceedings*, 1-5. Champaign, IL, USA. [Equal contributions]
- [c11] Bettivia, R.*, **Cheng, Y.-Y.***, & Gryk, M.* (2023). What does provenance LACK: how retrospective and prospective met the subjunctive. *Proceedings of iConference 2023 Annual Meeting*, 74-82. Barcelona, Spain. [Equal contributions]
- [c10] **Cheng, Y.-Y.**, & Ludäscher, B. (2020). Reconciling taxonomies of electoral constituencies and recognized tribes of indigenous taiwan. *Proceedings of the ASIS&T 2020 Annual Meeting*. 1-15. Virtual. <https://doi.org/10.1002/pr2.248>
- [c9] **Cheng, Y.-Y.**, Hoang, K.-L., & Ludäscher, B. (2020). Cacao, cacao, or cocoa? reconciliation of taxonomic names in Biodiversity Heritage Library. *Proceedings of the Sixteenth International ISKO Conference*, 1-10. Aalborg, Denmark. <https://doi.org/10.5771/9783956507762-88>
- [c8] **Cheng, Y.-Y.**, Parulian, N., Hsiao, T.-K., Dinh, L., Sarol, J., & Schneider, J. (2019). ReTracker: actively and automatically matching retraction metadata in Zotero. *Proceedings of the ASIS&T 2019 Annual Meeting*, 372-376. Melbourne, Australia. <https://doi.org/10.1002/pr2.32>
- [c7] Dinh, L., Sarol, J., **Cheng, Y.-Y.**, Hsiao, T.-K., Parulian, N., & Schneider, J. (2019). Systematic examination of pre- and post-retraction citations. *Proceedings of the ASIS&T 2019 Annual Meeting*, 390-394. Melbourne, Australia. <https://doi.org/10.1002/pr2.35>
- [c6] **Cheng, Y.-Y.**, & Ludäscher, B. (2019). Exploring geopolitical realities through taxonomies: the case of Taiwan. *Proceedings of North American Symposium on Knowledge Organization (NASKO 2019)*, 7, 1-17. Philadelphia, PA, USA. <http://dx.doi.org/10.7152/nasko.v7i1.15625>
- [c5] Gupta, S., **Cheng, Y.-Y.**, & Ludäscher, B. (2019). Possible worlds explorer: datalog & answer set programming for the rest of us. *Datalog 2.0*, 44-55. Philadelphia, PA, USA.
- [c4] **Cheng, Y.-Y.**, & Ludäscher, B. (2018). Full of beans: a study on the alignment of two flowering plants classification systems. *Proceedings of the 18th European Networked Knowledge Organization Systems Workshop (NKOS 2018)*, 1-11. Porto, Portugal.
- [c3] **Cheng, Y.-Y.**, Franz, N., Schneider, J., Yu, S., Rodenhasen, T., & Ludäscher, B. (2017). Agreeing to disagree: reconciling conflicting taxonomic views using a logic-based approach. *Proceedings of the ASIS&T 2017 Annual Meeting*, 46-56. Washington, D.C., USA. <https://doi.org/10.1002/pr2.2017.14505401006>

- [c2] **Cheng, Y.-Y.**, & Chen, H.-H. (2016). A study on constructing a cross-lingual ontology. *International Conference on Dublin Core and Metadata Applications (DCMI 2016)*, 1 page. Copenhagen, Denmark.
- [c1] Sun, C.-H., **Cheng, Y.-Y.**, Tsai, B, Y., & Chen, H.-H. (2014). Ontology development in geographic information systems: a case study of multi-agent knowledge-oriented cyberinfrastructure project (MAKOCI). (in Traditional Chinese). *12th Conference of Taiwan-China Cross Strait Chinese Association of Library and Information Science Education*, 1-11. Tianjin, China.

Refereed panels in conference proceedings

- [pa4] **Cheng, Y.-Y.**, Choi, I., Bettivia, R., Lee, W. C., & Watson, B. M. (2024). Knowledge organization systems and provenance: experiences and challenges. *Proceedings of the Association for Information Science and Technology*, 61(1), 741-744. (Sponsored by SIG-CMR) <https://doi.org/10.1002/pra2.1091>
- [pa3] Choi, I., **Cheng, Y.-Y.**, Washington, A., & Falconer, J. (in press). Metadata provenance and explainability. *To appear in Proceedings of DCMI 2024 Annual Meeting*, Toronto, Canada.
- [pa2] Bettivia, R.*, **Cheng, Y.-Y.***, & Gryk, M*. (2023). Retrospective, subjunctive, prospective: provenance challenges across time. *iPRES 2023 Conference Proceedings*, 1-2. Champaign, IL, USA. [*Equal contributions]
- [pa1] Bettivia, R.*, **Cheng, Y.-Y.***, Gryk, M.*, Bishop, W., Chassanoff, A., Greenberg, J., & Lischer-Katz, Z. (2022). Storied past, bright future: a provenance jam session. *Proceedings of the ASIS&T 2022 Annual Meeting*, 544-547. Pittsburg, PA, USA. <https://doi.org/10.1002/pra2.622> [*Equal contributions]

Refereed posters or abstract in conference proceedings

- [po10] Hruschka, D., **Cheng, Y.-Y.**, Hsiao, I. H., Bischoff, R., Peebles, M., Kasi, H., & Huang, C. (2024). Tools for integrating data by complex, dynamic categories. *Proceedings of the Association for Information Science and Technology*, 61(1), 934-936. Calgary, Canada. <https://doi.org/10.1002/pra2.1145>
- [po9] **Cheng, Y.-Y.** (2024). Pork roll or Taylor ham? when central new jersey and multiple controlled vocabularies exist in biodiversity data. SPNHC & TDWG Joint Conference. Virtual/Okinawa, Japan.
- [po8] **Cheng, Y.-Y.**, Parulian, N., & Dinh, L. (2023). A text mining approach to uncover the structure of subject metadata in the Biodiversity Heritage Library. *Proceedings of the ASIS&T 2023 Annual Meeting*, 926-928. London, United Kingdom. <https://doi.org/10.1002/pra2.900>
- [po7] **Cheng, Y.-Y.** (2021). Systematic comparison of data models used in mapping knowledge organization systems. *Proceedings of the ASIS&T 2021 Annual Meeting*, 685-687. Salt Lake City, Utah. <https://doi.org/10.1002/pra2.529>
- [po6] **Cheng, Y.-Y.**, & Ludäscher, B. (2020). Through the magnifying glass: exploring aggregations of COVID-19 datasets by county, state, region, and taxonomies. *Proceedings of the ASIS&T 2020 Annual Meeting*, 1-5. Virtual. <https://doi.org/10.1002/pra2.355>
- [po5] **Cheng, Y.-Y.**, Dilliplane, S., & Ludäscher, B. (2020). Snailed it! merging taxonomically organized biodiversity datasets with shifting geopolitical realities. *International Digital Curation Conference (IDCC 2020)*. Dublin, Ireland. [10.5281/zenodo.3666830](https://doi.org/10.5281/zenodo.3666830)
- [po4] Dinh, L., **Cheng, Y.-Y.**, & Parulian, N. (2019). ReTracker: an open-source plugin for automated and standardized tracking of retracted scholarly publications. *Proceedings of ACM/IEEE Joint Conference on Digital Libraries (JCDL 2019)*, 406-407. University of Illinois at Urbana-Champaign. [10.1109/JCDL.2019.00092](https://doi.org/10.1109/JCDL.2019.00092)

- [po3] Dinh, L., & **Cheng, Y.-Y.** (2018). Middle of the (by)line: examining hyperauthorship networks in the human genome project. *Proceedings of the 81st Annual ASIS&T Meeting*, 790-791. Vancouver, Canada. <https://doi.org/10.1002/pra2.2018.14505501118>
- [po2] **Cheng, Y.-Y.**, Weigl, D. M., Downie, J. S., & Page, K. R. (2018). Towards incorporating the notion of feature shapes in music and text retrieval. *Proceedings of iConference 2018 Annual Meeting*, 1-4. Sheffield, U.K. <https://www.ideals.illinois.edu/items/105592>
- [po1] **Cheng, Y.-Y.**, Chen, H.-H. (2016). Towards a seamless multilingual semantic web: a study on constructing a cross-lingual ontology. *Proceedings of the 79th Annual ASIS&T Meeting*, 1-4. Copenhagen, Denmark: <https://doi.org/10.1002/pra2.2016.14505301142>

Public reports

- [pp1] McDowell, K., Bettivia, R., Bonn, M. S., Campbell-Meier, J., **Cheng, Y. Y.**, Corieri, I. L., ... & Sharma, S. (2023). Future applications for data storytelling toolkits. *Information Matters*, 3(6). <https://informationmatters.org/2023/06/future-applications-for-data-storytelling-toolkits/>

GRANTS [Project Total: \$5,155,000; Cheng: \$151,305]

Successful Proposals

- [sp3] (PI) **Y.-Y. Cheng**. (co-PI) R. Bettivia, M. Gryk. **Cheng Storytelling Across Time: Building a Community Around Provenance (B-CAP)**. Institute of Museum and Library Services (IMLS), National Leadership Planning Grants for Libraries. 2024-2026, Project total: \$150,000; Cheng: \$84,306.
- [sp2] (PI) N. Gucunski. **(co-PI) Y.-Y. Cheng**. (co-PI) M. Morgese. (co-PI, TTI) A. Najafi. (co-PI, TTI) S. Hurlebaus. **Advance Bridge Technology Clearinghouse (ATBC) project**. U.S. Department of Transportation. 2024-2029. Project total: \$5,000,000; Cheng: \$61,999.
- [sp1] (PI) **Y.-Y. Cheng**. **Documenting Provenance Through a Reproducible Taxonomy Alignment Workflow**, Rutgers University, School of Communication & Information, Grants for Individual Faculty Research (GIFR), 2023-2024. Total/Cheng: \$5,000

Unsuccessful Proposals

- [up2] (PI) **Y.-Y. Cheng**. Utilizing Taxonomy Alignment to Enhance Provenance and Interoperability in Natural History Collections (TAPIr). IMLS Laura Bush 21st Century Librarian Program (Early Career Research Development) (Successful 1st round proposal; not received in 2nd round 2024.)
- [up1] (PI) **Y.-Y. Cheng**. (co-PI) R. Bettivia, M. Gryk. Storytelling Across Time: Building a Pan-Disciplinary Community Around Provenance Documentation, Metadata, and Modeling. Institute of Museum and Library Services (IMLS), National Leadership Forum Grants for Libraries. (Successful 1st round proposal; Not received in 2nd round in 2023)

FELLOWSHIP AND AWARDS

- 2023 Berner-Nash Doctoral Dissertation Award, \$2000 awarded by School of Information Sciences, University of Illinois at Urbana-Champaign
- 2022 Beta Phi Mu Eugene Garfield Doctoral Dissertation Fellowship, \$3000 Dissertation fellowship awarded by the Beta Phi Mu Honor Society
- 2021 Anita and Marie Hostetter Fellowship, School tuition and fellowship support for doctoral candidates in LIS at the University Illinois Urbana Champaign
- 2019, 2020 LEADS-4-NDP Fellowship, \$10,000 LIS Education and Data Science for the

- National Digital Platform (LEADS-4-NDP) fellowship, an IMLS project led by Dr. Jane Greenberg from Drexel University (RE-70-17-0094-17, RE-246450-OLS-20)
- 2020 Teacher Ranked as Excellent, University of Illinois Urbana Champaign
- 2018 ASIS&T Student Design Competition winner, Winning project: ReTrackers: an automated and standardized solution to tracking retractions
- 2015 Best paper award, Master's Thesis, awarded by the Library Association of the Republic of China (Taiwan)
- 2014 Outstanding Graduate Student Award, awarded by the Chinese Association of Library & Information Science Education (Taiwan)
- 2010,2011, 2012, 2013 National Taiwan University Presidential Award. The award is given each year to students ranking within the top 1% of the class.

INVITED TALKS

- [I10] **Cheng, Y.-Y.** & Lee, W.-C. (2022 October). [Invited panelist] 學思歷程分享 [Research and Studying Abroad in the U.S]. National Taiwan University Department of Library and Information Science Alumni Gathering Panel.
- [I9] **Cheng, Y.-Y.** (Oct, 2021). *Agreeing to disagree: applying a logic-based approach to reconciling and merging multiple taxonomies*. Doctoral colloquium, 2021 Annual Meeting of the Association for Information Science and Technology (ASIS&T). Salt Lake City, UT, October 30 - November 2.
- [I8] **Cheng, Y.-Y.** (Sept, 2021). *Agreeing to disagree: applying a logic-based approach to reconciling and merging multiple taxonomies*. Doctoral colloquium, 2021 Annual Conference of the Association for Library and Information Science Education (ALISE). Virtual, September 20-24.
- [I7] **Cheng, Y.-Y.** (July, 2021). *Agreeing to disagree: applying a logic-based approach to reconciling and merging multiple taxonomies*. Doctoral colloquium, 2021 North American Symposium on Knowledge Organization (NASKO). Virtual, July 9-11.
- [I6] **Cheng, Y. -Y.** (Spring 2021). [Guest Lecture] Metadata, authority control and data quality issues. IS505, Information Organization and Access, Dr. Inkyung Choi. School of Information Sciences, UIUC.
- [I5] **Cheng, Y. -Y.** (2020). [Invited Talk] *Agreeing to disagree: applying a logic based approach to reconciling and merging multiple taxonomies*. Center for Informatics Research in Science and Scholarship (CIRSS) Seminar, UIUC.
- [I4] **Cheng, Y. -Y.** (2020). [Invited Talk] *LEADS fellow examples of impact project – Academy of Natural Sciences and the biodiversity knowledge graphs project*. LEADS-4-NDP Forum, Dr. Jane Greenberg. Metadata Research Center, Drexel University, College of Computing and Informatics.
- [I3] **Cheng, Y. -Y.** (2020). [Invited Talk] ReTracker: actively and automatically matching retraction metadata in Zotero. Center for Informatics Research in Science and Scholarship (CIRSS) Seminar, UIUC.
- [I2] **Cheng, Y. -Y.** (2017). [Invited Workshop Speaker] *Possible world explorer and taxonomy alignment research*. The Future of Systematics in Data-Centric Biology Workshop, Marine Biological Laboratory (MBL), Boston MA.
- [I1] **Cheng, Y. -Y.** (2017). [Invited Talk] *An exploration of cross-lingual ontology development in geographic information systems*. Knowledge Organization and Cultural Heritage: Perspectives of the Semantic Web Conference. Academic Sinica, Taiwan.

TEACHING

School of Communication and Information, Rutgers University

Courses Taught	Semester
Retrieving and Evaluating Electronic Information (ITE) (onsite)	Fall 2024
Independent Study (MI)	Spring 2025, Spring 2024
Metadata for the Information Professional (MI) (online)	Spring 2025, Spring 2024, Fall, 2023, Spring 2023
Organizing Information (MI) (online)	Spring 2025, Fall 2024, Spring 2024, Fall, 2023, Spring 2023, Fall 2022

School of Information Sciences, University of Illinois Urbana Champaign

Courses Taught	Semester
Metadata in Theory and Practice (MLIS) (online synchronous)	Summer 2021
Theory and Practice in Data Cleaning (MSIM) (hybrid), co-taught with Bertram Ludäscher	Fall 2020, Spring 2020
Teaching Assistantships	Semester
Information Organization and Access (MLIS) (online synchronous)	Spring 2021
Theory and Practice Data Cleaning (MCS-DS), (Coursera Online)	Summer 2020
Theory and Practice Data Cleaning (MSIM)	Fall 2019, Spring 2019, Fall 2018, Spring 2018, Fall 2017

Full-Day Workshop and Tutorials

- [w2] Navigating through the Panoply of Provenance Metadata Standards useful for Digital Curation. Workshop at iConference 2021. (online) (co-taught with Rhiannon Bettivia, Michael Gryk).
- [w1] Navigating through the Panoply of Provenance Metadata Standards useful for Digital Curation. Workshop presented at the International Digital Curation Conference (IDCC 2020). Dublin, Ireland, 17 February- 20 February. (co-taught with Rhiannon Bettivia, Michael Gryk).

Guest Lectures

- [L5] **Cheng, Y.-Y.** & Lee, W.-C. (2025 March). [Guest Lecture] Knowledge Organization Research. Information Organization Seminar. National Taiwan University Department of Library and Information Science.
- [L4] **Cheng, Y.-Y.** (December 2nd, 2024) [Guest lecture] *Taxonomy alignment theories, approaches, and issues*. Doctoral Theories and Issues in Library Studies. 16: 194: 656, Dr. Marie Radford. Rutgers University School of Communication & Information, New Brunswick, NJ.
- [L3] **Cheng, Y.-Y.** (September 25th, 2024) [Guest lecture] *Pork roll or Taylor ham: when multiple taxonomies exist in your data*. Proseminar in Communication, Information & Media Studies. 16:194:601, Dr. Melissa Aronczyk. Rutgers University School of Communication & Information, New Brunswick, NJ.
- [L2] **Cheng, Y.-Y.** (November 9th, 2022) [Guest lecture] *Agreeing to disagree: reconciling taxonomies using a logic-based approach*. Proseminar in Communication, Information & Media Studies. 16:194:601, Dr. Jen Theiss. Rutgers University School of Communication & Information, New Brunswick, NJ.

- [L1] **Cheng, Y.-Y.** (October 24th, 2022) [Guest lecture] *Reconciling taxonomies of electoral constituencies and recognized tribes of indigenous Tainan*. Issues in Social Justice & Information Institutions. 17:610:597, Dr. Emil Lawrence. Rutgers University School of Communication & Information, New Brunswick, NJ.

Teaching Certificate

2020 Graduate Teaching Certificate, University of Illinois Urbana-Champaign. Participated in a series of teaching development program and evaluations, and fulfilled additional teaching requirements.

SERVICE

Internal and External Service

Role	Venue/Organization	Time
Program Committee	ACM Hypertext	2025
Member	Rutgers SC&I LIS Doctoral Student Dissertation Proposal Committee	2025
Associate Editor	IEEE Data Descriptions	2024-present
Chair	Rutgers SC&I LIS Research Development Committee	2024-2025
Member	Rutgers SC&I LIS Doctoral Student Qualifying Exam Committee	2024
Member	Rutgers SC&I LIS Faculty search committee	2024-2025
Member	Rutgers SC&I Website committee	2024-present
Member	Advisory board, CatMapper: User-friendly tools for integrating data across complex, dynamic categories (NSF HNDS-I)	2023-present
Member	Rutgers SC&I LIS Planning and Governance Committee	2023-present
Member	Rutgers SC&I MI student awards and fellowship committee	2023-2025
Member	Rutgers SC&I PhD student award committee	2023
Paper session moderator	Annual Conference of the Association for Information Science and Technology (ASIS&T), Research Data Management session	2021
Paper session moderator	International Digital Curation Conference (IDCC), Metadata and Reusability session	2020
Member	Diversity and Inclusivity Task Force, LIS Education And Data Science Integrated Network Group (LEADING), Metadata Research Center, Drexel University (RE-70-17-0094-17, RE-246450-OLS-20)	2020-2023
Organization committee	Student poster session, Women in Data Science Urbana Champaign (WiDS)	2020
Founder and co-chair	UIUC ACM Women Student Chapter	2019-2020
Chair	Conceptual Foundations Group Interest-based research group in School of	2017-2020

	Information Sciences, UIUC	
--	----------------------------	--

Grant proposals, Journal, Conference Papers Reviewing

1. IMLS, Laura Bush 21st Century Librarian program, 2023-2024, 2024-2025
2. Journal of Education for Library and Information Science (JELIS)
3. Neural Processing Letters
4. Computers and Geosciences
5. Data Science Journal
6. Annual Conference of the Association for Information Science and Technology, 2016-present (best panel reviewer 2023)
7. iConference, 2016-present
8. Dublin Core Metadata Initiative (DCMI), 2016-present

Graduate School or Professional Employments

1. 2016-2022, Graduate research assistant and teaching assistant, School of Information Sciences, University of Illinois Urbana-Champaign
2. 2015-2016, Academic librarian, Taiwan Indigenous Peoples Resource Center, National Taiwan University Library, Taiwan

Member

1. Beta Phi Mu Alpha Chapter
2. Association for Information Science and Technology (ASIS&T)
3. International Society of Knowledge Organization (ISKO)
4. Association for Computing Machinery (ACM)
5. Association for Library and Information Science Education (ALISE)