

Leslie Harka

leslieharka@gmail.com | linkedin.com/in/lharka | Portfolio: lharka.com

PROFESSIONAL SUMMARY

MLIS graduate with 7+ years of cataloging, metadata, and digital collections experience in academic libraries, including team coordination and workflow leadership across multiple units. Skilled in metadata standards, repository systems, and cross-functional collaboration with a strong record of staff mentoring, documentation, and implementing sustainable practices.

EDUCATION

M.S. in Library and Information Science, University of Washington 2025

Concentration: Metadata Design, Digital Archives, Records Management, and Knowledge Organization

M.S. in Community and Regional Planning, University of Oregon 2021

Concentration: Community Archives and Digital Collections

- Completed independent study internships in Digital Scholarship Center and Special Collections at UO Libraries (cataloging, metadata remediation, archival description); led community-engaged planning project developing skills in project coordination and stakeholder engagement.

B.S. in Psychology, University of Oregon 2019

PROFESSIONAL EXPERIENCE

Lab for Computing Cultural Heritage, University of Washington iSchool Seattle, WA

Research Collaborator, Web Archives Analysis Project Mar 2025 – Present

- Analyzing large-scale web archives to identify embedded data (e.g., spreadsheets embedded in PDFs).
- Applying AI/ML and multimodal approaches to improve discovery and metadata assessment.
- Documenting workflows and contributing to publications on digital preservation and access.

iSchool GA Crew, University of Washington iSchool Seattle, WA

Graduate Staff Assistant Sep 2022 – Jun 2025

- Coordinated team workflows and project logistics for faculty research support, including tracking multiple concurrent requests, managing timelines and deliverables, and maintaining active communication with faculty and team members.
- Conducted bibliographic and metadata analysis using Python and SPARQL to support faculty research projects; provided research support including literature reviews, citation formatting, and data analysis.
- Mentored and onboarded incoming graduate assistants, developing training materials and comprehensive workflow documentation to build team capacity and ensure knowledge continuity.
- Coordinated collaborative scholarly workflows using Git/GitHub, Overleaf, and BibTeX for faculty textbooks and publications; managed faculty grant and publication metadata in Activity Insight database.
- Administered Qualtrics surveys and performed qualitative/quantitative analysis; designed graphics for publications and events using industry-standard tools.

Digital Scholarship Center, University of Oregon Libraries Eugene, OR

Metadata Assistant Nov 2017 – Aug 2022

- Assisted with large-scale metadata audits and remediation for thousands of digitized and born-digital materials during the Oregon Digital migration (2021-2022) in the Samvera-based repository.
- Applied linked data principles, controlled vocabularies, authority URIs, and spreadsheet-based workflows to remediate metadata at scale and improve accuracy, interoperability, and discoverability.
- Managed controlled vocabularies and authority URIs in Oregon Digital's vocabulary manager.
- Added new artist records to Getty ULAN, improving representation in UO's Oregon Percent for Art collection.

Special Collections and University Archives, University of Oregon Libraries

Thomas Intern

Eugene, OR
Sep 2021 – Mar 2022

- Processed archival collections and created accession records and finding aids in ArchivesSpace.
- Cataloged and uploaded descriptive records using Alma/Primo and OCLC Connexion.
- Applied cataloging and preservation best practices for both physical and born-digital archival materials.

Special Collections and University Archives, University of Oregon Libraries

Digital Collections Intern

Eugene, OR
Jun 2020 – Sep 2021

- Remediated metadata for 2,500 digitized photographs from the Doris Ulmann collection.
- Designed a public-facing digital exhibit in Spotlight to showcase and contextualize the collection.
- Collaborated with cross-departmental teams on metadata description, preservation, and access.

PORTFOLIO & SELECTED WORK

Selected digital collections, metadata, and archival projects demonstrating technical expertise in metadata remediation, repository management, and emerging technologies. Full portfolio available at lharka.com.

Embedded Spreadsheets in Web Archives (2025-Present)

- Developing methods to detect spreadsheets in End of Term Web Archive PDFs using multimodal models and computer vision. Published tutorial on accessing parquet data on AWS S3. [\[Tutorial\]](#)
- Tools: Python, SQL, AWS S3, Parquet, Web Archives, Data Extraction

ISMIR Metadata Analysis (2024-2025)

- Analyzed ISMIR conference metadata (2000-2024) for author affiliations and topics. Built Python pipeline using dblp Knowledge Graph and SPARQL endpoint to transform RDF data. [\[Project Site\]](#)
- Tools: Python, SPARQL, RDF, Data Extraction

Oregon Digital Migration (2021-2022)

- Metadata auditing and remediation for thousands of cultural heritage materials in the Oregon Digital platform relaunch (Samvera Hyrax, Feb 2023). [\[Repository\]](#) [\[Live Site\]](#)
- Tools: Samvera, Metadata, Digital Archives

Archival Finding Aids (2021-2022)

- Processed collections and created EAD-compliant finding aids in ArchivesSpace for physical and born-digital materials. Selected examples: [\[1\]](#) [\[2\]](#) [\[3\]](#)
- Tools: ArchivesSpace, Metadata, Digital Archives, EAD

Doris Ulmann Digital Exhibit (2020-2021)

- Remediated metadata for 2,400+ Doris Ulmann photographs (early 20th-century American South) and created public Spotlight exhibit. [\[Collection\]](#) [\[Exhibit\]](#) [\[Blog Post\]](#)
- Tools: Spotlight, Metadata, Digital Archives, Exhibits

SKILLS

Metadata Standards & Schema: MARC, RDA, Dublin Core, MODS, RDF, XML, LCSH, LCNAF, AAT

Library & Discovery Systems: OCLC Connexion, Alma/Primo, ArchivesSpace, VIAF, Wikidata

Leadership: Project coordination and workflow management, staff onboarding & mentorship, documentation design, cross-unit collaboration

Technical Skills: Python, SQL, SPARQL, Git/GitHub, command line, Overleaf, BibTeX

Productivity Tools: Microsoft Office, Google Workspace, Adobe Creative Suite (Illustrator, Acrobat)

Accessibility & Preservation: OCR, tagging/alt text, file format standards, digital preservation workflows

Reference Management: Zotero, EndNote