I am a software developer and design researcher with two decades of experience bridging the worlds of libraries and archives with the World Wide Web. I have worked in academia, startups, corporations and government. I work best as a member of agile, highly collaborative teams that focus on furthering social justice work through community building, knowledge sharing and a commitment to sustainability. My research and writing is focused on sociotechnical approaches to understanding digital preservation and memory work.

## Skills

* web technology, APIs, REST, linked data, web archiving
* digital preservation, data management, knowledge graphs, search
* open source, standards development, community building
* automated testing, agile software development, rapid prototyping
* Python, JavaScript, Go, Ruby, Perl, PHP, Java, C, Scheme
* data analysis, Pandas, SciPy, Spark
* data visualization: D3, Plotly, vis.js, Gephi
* Ubuntu, Red Hat, Amazon Web Services

## Work History

Research Faculty  
Maryland Institute for Technology in the Humanities  
University of Maryland  
08/2014 - Present

* Full stack software development (JavaScript, React, Python, PostgreSQL, ElasticSearch)
* Social media data analysis and visualization (Python, JavaScript, D3, VisJS)
* Linked Data publishing (Python, NodeJS)
* Sustainable data preservation workflows (AWS)
* Automated testing and continuous integration (GitHub)
* Configuration management (Git, Docker, Ansible)
* Agile requirements gathering & project management
* Instructor & Mentor: software development (Python, JavaScript), data science (Python, Pandas, Spark)

Information Technology Specialist  
Repository Development Center  
Library of Congress, Washington, DC  
03/2006 - 08/2014

* Digital preservation and repository software design and development (Python, Java, MySQL)
* Lead software developer on the [National Digital Newspaper Project](http://chroniclingamerica.loc.gov/) (Django, MySQL)
* Created a web based quality assurance tool for handling high volume image digitization projects
* social media archiving (Twitter)
* Designed and implemented the [Library of Congress Linked Data Service](http://id.loc.gov/)
* Specification writing and testing [W3C Semantic Web Deployment Group]

Consultant  
George Washington University, Washington, DC  
10/2013 - 12/2014

* Django Web Development
* Voyager ILS integration (Oracle)
* Serial Solutions Summon API integration

Consultant  
Nelson Institute for Environmental Studies  
University of Wisconsin, Madison, WI  
09/2013 - 06/2014

* Web harvesting (Python)
* Data managment (Git, AWS)

Consultant  
Hypothes.is  
San Fransisco, CA  
01/2013 - 10/2013

* JavaScript and Python development
* Development of the open source [Annotator](http://github.com/okfn/anotator) and [Annotator Store](http://github.com/okfn/anotator-store) projects
* Design and implementation of [cross format annotation] annotation support
* Functional/unit testing (Selenium, PyTest, Travis-CI, SauceLabs)

Consultant  
GlueJar  
11/2011 - 04/2012

* Django software development [UnglueIt](http://unglueit.com) [unglueit](http://unglueit.com)
* Amazon AWS Deployment (EC2, RDS, Route 53) for dev, test and production
* Google Books metadata harvesting
* Google, Facebook, Twitter social media integration
* Automated testing and continuous integration (Jenkins)

Consultant  
Canary Project  
Yale School of Medicine, New Haven, CT  
07/2005 - 12/2005

* Python object oriented design
* Automated testing
* Implemented a reporting framework
* Quixote, Pylucene, MySQL, Red Hat administration

Senior Software Engineer  
Follett Corporation, McHenry IL  
05/2005 - 03/2006

* Building and supporting a distributed knowledge management application
* Requirements gathering and project planning
* Web service design and implementation
* ONIX metadata harvesting and ingest
* Java performance tuning
* Automated testing
* Database administration (Oracle, SQL Server)
* Linux/Windows systems administration

Consultant  
Performix, Chicago IL  
02/2004 - 04/2005

* Object oriented design
* Refactored large Perl codebase
* Perl, Apache, mod\_perl, DB2

Consultant  
Queens Borough Public Library, New York NY  
06/2004 - 05/2005

* Character encoding conversion
* Controlled vocabulary mining
* Perl data processing training

Consultant  
Notre Dame University, South Bend IN 06/2004 - 12/2004

* Training in object oriented programming and unit testing
* High volume metadata harvesting
* Guidance and implementation search technologies
* Ported Java Commonn Query Language parser to Perl

Applications Developer  
Follett Library Resources, McHenry IL  
07/2002 - 05/2004

* PHP/Perl development in a Unix/Oracle environment
* Determined project requirements, created task lists and schedules
* Designed and developed library collection analysis application
* Unit and application level automated testing
* Oracle/Apache stress testing and performance tuning
* XML processing
* Distributed application design using web services

Senior Software Engineer  
CheetahMail, New York NY  
02/2000 - 07/2002

* Team lead in a group of 8 programmers
* Developed data processing infrastructure for Fortune 500 clients
* Designed/developed real time reporting system
* Creation and maintenance of web service API
* Expertise in SMTP and HTTP protocols
* Web services development
* Technical representative for partners and sales
* Networking and systems administration as required

Electronic Resources Librarian  
Old Dominion University, Norfolk, VA  
04/1998 - 02/2000

* SGML and XML data processing for the VIVA Consortium
* Created current awareness system to alert students/faculty to new items in their field of interest.
* ILS and OCLC metadata workflow management

Visiting Assistant Professor University of Illinois, Champaign-Urbana, IL  
08/1996 - 03/1998

* Web development (ASP, SQL Server)
* Collection development and cataloging in Sociology and Political Science
* Preservation of 19th century criminology pamphlet collection
* Managed reference services in the Social Science Library

## Selected Projects

These are some projects and experiments where I’ve been directly involved in the technical development.

* [Unlocking the Airwaves](https://unlockingtheairwaves.org) - NEH funded virtual reunification and minimal computing project
* [Documenting the Now](https://www.docnow.io) - Mellon funded social media archiving practices and tools
* [Chronicling America](https://chroniclingamerica.loc.gov) - Library of Congress access to historical newspapers.
* [Linked Data Service](https://id.loc.gov) - Publishing controlled vocabularies at the Library of Congress using Linked Data
* [Hypothesis](https://web.hypothes.is/) - An annotation tool for the web
* [Wikistream](http://wikistream.wmflabs.org) - visualize realtime edits to all major language Wikipedias
* [Anon](http://github.com/edsu/anon) - track anonymous Wikipedia edits using Twitter
* [Pymarc](http://github.com/edsu/pymarc) - a python library for MARC processing
* [jobs.code4lib.org](https://jobs.code4lib.org) - a jobs board for the code4lib community
* [Linkypedia](http://github.com/edsu/linkypedia) - view how websites are used on Wikipedia
* [Fondz](http://github.com/fondz) - generate an archival description for describing born digital content
* [Microdata](https://github.com/edsu/microdata) - a python library for parsing HTML 5 microdata
* [rdflib](https://github.com/RDFLib) - a Python library for processing RDF

## Selected Publications

* [Appraisal Talk in Web Archives](https://inkdroid.org/papers/appraisal-talk.pdf). Archivaria. May, 2020.
* [Archival Circulation on the Web: The Vine-Tweets Dataset](https://culturalanalytics.org/article/11048). Cultural Analytics. June, 2019.
* [Algorithms and Agenda-Setting in Wikileaks’ #PodestaEmails Release](https://www.tandfonline.com/doi/abs/10.1080/1369118X.2019.1626469). Information, Communication & Society. 2019
* [Appraisal Practices in Web Archives](https://osf.io/preprints/socarxiv/75mjp/). SocArXiv Papers, 2018.
* [Twitter and Tear Gas: Book Review](https://www.tandfonline.com/doi/abs/10.1080/24701475.2017.1383639). Internet Histories, 2017.
* [Ferguson and the death of Michael Brown on Twitter: #BlackLivesMatter, #TCOT, and the evolution of collective identities](http://dx.doi.org/10.1080/01419870.2017.1335422). Ethnic and Racial Studies, 40(10).
* [#sayhername: a case study of intersectional social media activism](http://dx.doi.org/10.1080/01419870.2017.1334934). Ethnic and Racial Studies, 40(10).
* [Bots, Seeds and People: Web Archives as Infrastructure](/papers/bots-seeds-people.pdf). Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 17.
* [Automating Transparency](https://source.opennews.org/articles/automating-transparency/). The Source. April, 2015.
* [MJ no more: Using Concurrent Wikipedia Edit Spikes with Social Network Plausibility Checks for Breaking News Detection](http://arxiv.org/abs/1303.4702). 2013.
* [Linking Things on the Web: A Pragmatic Examination of Linked Data](http://arxiv.org/abs/1302.4591), 2013.
* [Key Choices in the Design of Simple Knowledge Organization System (SKOS)](http://arxiv.org/abs/1302.1224), 2013.
* [Library Linked Data Final Report](http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/), W3C Incubator Group Report, October 2011.
* [SKOS Simple Knowledge Organization System Primer](http://www.w3.org/TR/skos-primer), W3C Working Group Note, August 18th, 2008. http://www.w3.org/TR/skos-primer/
* [LCSH, SKOS and Linked Data](http://dcpapers.dublincore.org/pubs/article/view/916), Proceedings of the International Conference on DublinCore and Metadata Applications, 2008.
* [Following Your Nose to the Web of Data](http://inkdroid.org/journal/2008/01/04/following-your-nose-to-the-web-of-data/), Information Standards Quarterly, 2008
* [Building OAI-PMH Harvesters With Net::OAI::Harvester](http://www.ariadne.ac.uk/issue38/summers) Ariadne, no. 38, Jan 2004.
* Cyril: Expanding the Horizons of MARC21, Library Hi Tech, v. 22, no. 1, 2004.
* A Consortial Approach to Cooperative Cataloging and Authority Control: The Virtual Library of Virginia Experience, Resource Sharing and Information Networks, v. 16, no. 1, 2003, p. 33-53.
* [Paying Homage to Perl (PHP)](http://www.theperlreview.com/articles/php.html), The Perl Review, January, 2003.
* MARC it your way: MARC.pm, Information Technology and Libraries, March, 2002.
* Comprehensive Access to Periodicals: A Database Solution, Library Collections Acquisitions and Technical Services, v. 24, no. 2, 2000, p. 33-44.
* Show Me the Metadata, Against the Grain, v.11, no.1, 1998.
* Gateways to Social Work/Welfare on the Net, C&RL News, v.59, no. 3, 1998, p.163-167.

## Education

* PhD, College of Information Studies, University of Maryland, 2015-2020.
* MLS, School of Communications, Information and Library Science, Rutgers University, 1995-1996.
* BA English and American Literature, Rutgers University, 1987-1991.

## Awards

* 2018 [Community Capacity Award](https://www.diglib.org/announcing-the-2018-dlf-community-capacity-award-winner/). Digital Library Federation.
* 2018 [Exemplary Service Award](https://www2.archivists.org/node/23172). Society of American Archivists.
* 2015 University of Maryland Scholarship and Research Award
* 2015 [Frederick G. Kilgour Award for Research in Library and Information Technology](http://www.ala.org/news/press-releases/2015/03/ed-summers-recipient-2015-kilgour-award). American Library Association.