

Open Data for Public Good

Grant RE-40-16-0015-16

Interim Performance Report Year 1

July 1, 2016 - June 30, 2017

Submitted September 28, 2017 by:

Dr. Carole Palmer, Co-PI
cpalmer@uw.edu

Dr. Nicholas Weber, Co-PI
nmweber@uw.edu

An Yan, PhD Student
yanan15@uw.edu

Bree Norlander, Project Coordinator
norlab@uw.edu

University of Washington, Information School

Purpose of the Project

The Open Data for Public Good (ODPG) project began July 1, 2016. The purpose of ODPG's work is threefold:

1. To prepare future and current public librarians to curate collections of open data of value to local communities,
2. To gain experience in building the necessary infrastructure and preservation environments to sustain open data collections for long-term sustainability of these valued assets, and
3. To make collaborate with open civic data providers on advocacy and outreach activities that increase awareness about, and use of open data by the public.

Over the course of ODPG's first year the goals of the project are being achieved through the development of new LIS and data science curriculum, as well as practical learning experiences for students enrolled at the University of Washington's iSchool. ODPG's curriculum development activities have focused specifically on new course modules for curating and responsibly managing open civic data. Practical learning experiences have also been facilitated in collaboration with a network of practicing information professionals from the public sector-- each of whom represent a municipal or state government agency engaged in ongoing open data initiatives that can directly benefit from the data expertise of an LIS workforce.

We have completed the first year of our three-year program. During this past year we focused on:

1. Updating graduate-level data curation curriculum
2. Pairing students with external partners for field experiences both as Capstone projects and as summer internships
3. Outreach to external partners, potential partners, and the open data community

Activities

Course Enhancement and Revision

INFX 551 - Fundamentals of Data Curation.

Two sections of the the Fundamentals of Data Curation course at the University of Washington's iSchool were revised to include learning objectives related to open data and public sector information. Enrolled students represented the Masters in Library and Information Science (MLIS), Masters of Science in Information Management (MSIM) graduate degree programs, as well as the Informatics undergraduate major in the iSchool. In total, 35 students were exposed to this revised curriculum.

INFX 598 - Special Topics: Advanced Data Curation.

A special topics course, Advanced Data Curation, was also redesigned to expand the scope to cover open government and civic data, in addition to the previous emphasis on the curation of research data. New content included readings and lectures on access, value, and quality issues and challenges, and technical platforms for open data. Assignments focused on giving students practical experience with curation tools, evaluating open data, and both building and implementing collections designed for a particular user community. The small seminar-style class had 6 students. Additionally, a PhD student at the iSchool - acting as a Teaching Assistant - gained experience with the new course design and content and supported students in experimentation with technical tools.

Data Curation Workbench Development

Students in INFX 551 and 598 also benefited from a new collaboration with the National Data Service to test adapting the NDS LABS workbench to support data curation education. In particular, the platform brings together a core set of tools to support hands-on, practical work with current curation software and tools. It allows both on-campus and online students to gain experience and experiment with the tools without the interference of setup and administration. The Lab Workbench was incorporated into the coursework for students in both the Fundamentals of Data Curation course, and all but one student in INFX 598 opted to use the workbench to extend their final project to include a demonstration of using CKAN to implement a new open data collection.

Targeted Student Development

Phd Student

One PhD student in the UW iSchool served as the project manager and research assistant of ODPG over the last year. The student completed a teaching practicum working with Dr. Palmer to revise the curriculum of INFX 551 to include readings, lab work, and assignments related to open data curation.

Masters Students

Two MLIS students at the UW iSchool completed independent studies working with Dr. Weber. One student helped design and implement a research project studying the open data services offered by USA public libraries. Another student, who holds a JD specializing in Intellectual Property law, developed a method for retrieving and analyzing license information for open data repositories used by cities in the USA.

Undergraduate Student

An undergraduate major in the iSchool's Informatics program completed a research project with Dr. Weber in building an API to retrieve, and transform Seattle City Council data from an existing open data portal managed internally by the City of Seattle. The student's work is contributing to an ongoing personal research project in building civic technologies to support

increased government accountability, as well as a curation workflow for improving city council data that will be used by the City of Seattle's IT department.

Field Experience

Over the last year, ODPG has worked with our project partners to identify and place students in real world open data settings, gaining valuable hands-on experience managing, curating, and visualizing open data in collaboration with a public institution. In total, 11 students from the MLIS, MSIM, and Informatics programs completed a Capstone project, and 4 students completed a full-time summer internship.

Capstone Projects

Four teams of students worked with external partners to address open data challenges. Their work began in January 2017 and culminated in a Capstone poster presentation at the end of May 2017.

Opening Up the Data: Visualizing the effectiveness of Puget Sound restoration efforts

Sponsor: Puget Sound Partnership

Project Description: Interoperability, dispersed data, and inconsistent data formats are common issues across information science. This Open Data Literacy (ODL) Capstone tackled these issues in the realm of environmental restoration. Numerous restoration projects have been undertaken throughout the Puget Sound, but connecting investments in these projects to co-located indicators of habitat viability is challenging. This capstone team leveraged open, found data and open-source tools to build a scalable, sustainable data processing pipeline and an interactive, web-based visualization prototype. This helped their partners better tell the story of Puget Sound restoration efforts and demonstrates that open-source tools can help data curators meet open-data needs.

Total Resource Observation Utility Team

Sponsor: Washington Department of Fish and Wildlife

Project Description: The Total Resource Observation Utility Team (TROUT), created a data visualization using Washington Department of Fish & Wildlife's (WDFW) high-resolution land change data. The interactive visualization makes this important data understandable to the public, policymakers, and environmental scientists. TROUT has unlocked this public data and improved access by disseminating in a clear, concise, user-centered dashboard presenting actionable information. This visualization allows the user to see actual changes in the Puget Sound over time with data-rich GIS information describing the changes. TROUT's use of open-source technologies and code documentation mean the dashboard can be replicated for displaying future public spatial datasets.

WA State Public Meetings System

Sponsor: Washington State Office of the Chief Information Officer

Project Description: The Capstone team working on the Washington State Public Meetings System set a data standard for all the public meetings digital records in Washington State. They created a cloud dataset based on this data standard and populated the dataset with test data scraped from the city and county council websites. They also built a website to visualize the database at its backend. The end result of the project is a platform for users to perform queries about public meetings via the website.

Waste 2 Resources

Sponsor: State of Washington Department of Ecology

Project Description: The Washington State Department of Ecology works to manage and conserve public resources and collects information about the effectiveness of its conservation programs. The Waste2Resources project helped in the Open Data Literacy initiatives of the Department of Ecology by transferring this information to the open data portal of Washington state (data.wa.gov) and designing user-friendly data visualizations. This project created a positive social impact by presenting the existing waste management information through interactive and powerful visualizations to the public. It also helped the department and government officials in tracking of the progress of their conservation efforts more effectively.

Summer Internships

The ODPG team selected four students to work with external partners in eight-week paid summer internships. The students documented their internship progress via blog posts (<https://medium.com/open-data-literacy>) and a short presentation to the ODPG team, partners, and stakeholders in August 2017.

Context for Communities: Seattle and Metadata Standards for Civic Data

Student Program: Master of Science in Information Management Program

Sponsor: City of Seattle

This intern helped the Open Data Program at the City of Seattle apply civic metadata standards to the datasets included on data.seattle.gov. There are numerous standards that are currently available, but the way they are applied varies. Data about some areas of civic life, such as bus transit times, corresponds well to existing metadata standards. But for others, such as the data generated by cutting-edge Internet of Things devices, standards barely exist. The intern researched and reported on the question: How can the City of Seattle reconcile these differing metadata standards and apply them to the city's datasets?

OCIO Geospatial Program

Student Program: Master of Library and Information Science Program

Sponsor: Washington State Office of the Chief Information Officer

Working in the Geospatial Program Office, this intern evaluated and consolidated Washington State's three main geoportals into one. Multiple authoritative portals are expensive, inefficient, and confusing for new users. Transitioning requires indexing geospatial and non-geospatial data from disparate portals, tracking departmental provenance, and constructing a sustainable portal

that can be easily utilized by stakeholders with varying abilities. To that end, the intern reviewed the literature on current geoportal best practices and drafted an eight-week plan for completion. The new singular geoportal can now be found at: <https://ocio.wa.gov/geospatial-program-office/geospatial-data>.

Up and Running with Open Data: Open Data Unconference

Student Program: Master of Science in Information Management Program

Sponsor: Seattle Public Library

This intern worked with Seattle Public Library to develop an Open Data Unconference framework, entitled Up and Running with Open Data. The plan included detailed event logistics, goals, learning objectives, lesson plans, and curriculum. A public *Up and Running with Open Data* event was held September 9, 2017; and a regional library staff training was held September 20, 2017.

Washington State Department of Transportation's Data or Term Search Tool

Student Program: Master of Library and Information Science Program

Sponsor: Washington State Department of Transportation (WSDOT)

In collaboration with technical and business stakeholders, this intern conducted the initial work for a full business analysis of WSDOT Data or Term Search (DOTS), which is managed by the Data Management Services (DMS) group. At its heart, DOTS is a data catalog: an inventory of all of the data objects within DMS-managed environments, along with technical metadata for those objects. The intern's analysis will help data and domain experts deliver greater benefit to knowledge seekers through improvements to DOTS and related information-sharing processes.

Synergistic Activities

Co-PI, Dr. Weber, attended two workshops where he represented ODPG's ongoing work on open data literacy:

1. Data Science in Libraries, May 16-17 in Pittsburgh PA
2. Cyberinfrastructure Workforce Workshop, August 14-15 in Alexandria, VA

Dr. Weber also collaborated with the Technology and Social Change (TASCHA) group at the University of Washington to develop and deliver a data literacy workshop in Yangon, Myanmar on March 16-20, 2017. This ongoing work will provide a set of readings, workshops, and take-home exercises that will be used by public officials throughout the country of Myanmar. Case studies from this activity will be used in ODPG curriculum revisions scheduled for 2018.

Dr. Palmer's role on the Steering Committee for the National Data Service (NDS) resulted in interest within NDS in how they can support data education in iSchools, and potentially beyond. A pilot project proposal was solicited and accepted.¹ A first step in development of the workbench was completed for use by students in our spring 2017 data curation courses who were able to experiment with developing and exposing open data collections using CKAN--one

¹ <http://www.nationaldataservice.org/projects/pilots.html>

of the most widely used platforms for open government data. The aim is to extend the workbench with a larger suite of tools for courses in the next academic year and perhaps to collaborate with at least one another iSchool for testing and further development.

Dissemination and Outreach

Presentation

Palmer, C. L., Weber, N. & Johnson, T. (2017, April 14). *An Educational Workbench for Data Curation*. Workshop presented at 7th National Data Service Consortium Workshop, Chicago, IL.

Abstract: The Data Curation Educational Workbench provides a platform for students to gain hands-on experience with data curation software and tools. The workbench is being piloted with students at the University of Washington (UW) Information School enrolled in the Fundamentals of Data Curation and Advanced Data Curation course. We will provide guidance for students to develop, publish, and manage collections of datasets through the platform's user interfaces and APIs. Students will use the workbench to experiment with and apply tools to a range of curation tasks. The initial phase will focus on building a coherent set of tools around stable instances of CKAN and Dataverse for building collections and complementary tools for performing curation tasks, such as FIDO, SQL Workbench, Oxygen, EIS Archiver, and selected harvesting tools.

Posters

Weber, N. & Yan, A (2017) *Integrating User Feedback with Open Data Quality Models*. Poster accepted to Annual Conference of the Association for Information Science & Technology, Washington, DC.

Abstract: User feedback is critical to improving the quality of open data. However, most open data publishers gather only anecdotal evidence about user experience. This unstructured and informal commentary is, consequently, difficult to translate into actionable steps towards improving data quality. Drawing on user comments collected from Data.Gov - an open data portal providing access to thousands of datasets published by city, state, and federal government agencies in the USA – we inductively develop a classification of reported data quality issues. This poster presents preliminary findings from applying this coding scheme to all issues that users filed on Data.gov in 2015 and 2016. We suggest that our classification scheme can help open data publishers collect structured, actionable information to improve data quality.

Weber, N., Yan, A., & Palmer, C. L. (2017, February 21). *Open Data Literacy Project*. Poster presented at University of Washington eScience Institute Poster and Networking Session. University of Washington, Seattle.

Weber, N., Yan, A., & Palmer, C. L. (2017, March 23). *Open Data Literacy Project*. Poster presented at iConference 2017. Wuhan, China. <https://doi.org/10.9776/17341>

Abstract: The Open Data Literacy project is preparing future and current librarians to advance open data initiatives in public libraries. This poster will provide an overview of our core activities and progress to date, with a focus on strategies iSchools can implement to collaborate with public sector partners to overcome the current lag in data expertise in the public library workforce. Core activities include new curriculum for master's students in Library and Information Science, a slate of fieldwork opportunities at institutions with open data initiatives, and community workshops and open education resources for public librarians and information professionals. The educational framework will improve public accessibility and use of open data while increasing the data capabilities of both new and practicing information professionals in public libraries.

Information Session Fall Quarter 2016

The Open Data Literacy team held an information session for graduate students on November 22, 2016. The agenda for the meeting included:

1. Overview, Motivation, Partners
2. Curriculum (courses you can take)
3. Student Project Opportunities
4. Sample Projects
5. Opportunities at Seattle Public library
6. Q&A

Approximately 25 students attended and we successfully recruited several students for Capstone Projects.

Website Development

We created a comprehensive website to showcase the work of the Open Data Literacy project and to solicit new partners. The website (<https://odl.ischool.uw.edu>) includes pages for each Capstone project and each intern project. It highlights the goals of the Open Data Literacy project to enhance education and research, and to develop partnerships that further student's understanding of open data.

Medium Blog

We set up a publication on Medium.com (<https://medium.com/open-data-literacy>) for publishing blog posts related to Open Data Literacy projects. The summer interns used this publication method to disseminate information about the progress of their projects.

Evaluation

Capstone Evaluations

Upon completion of the Capstone projects, we surveyed both sponsors and students about their experiences.

Sponsors

We had three sponsors respond to the survey from three (out of four) different teams. All respondents were satisfied or very satisfied with their experiences and all three were willing to sponsor another Capstone. One respondent even noted in the comments field, “Great experience from start to finish. I need a higher category than ‘very satisfied.’” Sponsors advised the ODL team that Capstone teams should begin with face-to-face meetings with sponsors and that the Capstone teams would benefit from closer faculty connections.

Students

Six students responded to the Capstone survey. Each of the four teams was represented in the survey responses. All respondents were satisfied or very satisfied with their Capstone experience. The students benefitted from and appreciated the supportive relationship with the sponsor. Respondents included comments such as “Our sponsor was very enthusiastic about our project and genuinely cared about us and our work” and “It was a great, smooth experience. Our sponsor was very helpful in providing the information that we needed”. Students faced challenges with large datasets, messy data, and interoperability between web applications.

Focus Group

Two members of the ODL team interviewed the four ODPG summer interns near the end of their 8-week internships. Focus group questions focused on the internship experience, skills and tools needed to be successful, helpful coursework, and data curation coursework. Three of the four interns found that their internship requirements changed from what was originally discussed in the Spring. They were capable of adjusting to the change and successfully completing the internships. All four interns were satisfied with their experience and deliverable and believe that their sponsors were also satisfied. Most of the skills needed for the internship were acquired in previous work experiences or learned specifically for the work, rather than through coursework. Three of the interns had encountered open data in coursework, while one had not.

Administration and Coordination

ODL and Partner Meeting

The ODL Team gathered with external partners on November 1, 2016 to discuss potential student projects in the form of Directed Fieldwork, Capstone Projects, and Internships. In attendance were members of the ODL team from the **University of Washington** and partners or potential partners from **Washington State Department of Transportation, Washington**

State Historical Society, Seattle Public Library, State of Washington Office of the Chief Information Officer, and The City of Seattle.

Github

Many of the collaborative activities undertaken in year one required access to data, development of software, and central coordination of documentation. ODPG's organizational Github account (<https://github.com/OpenDataLiteracy>) was used for this purpose. These open-access repositories include:

- All materials and course deliverables from the INFX 551 course
- Research related repositories for ongoing work with students in the iSchool
- Internship repositories that host outputs from the ODPG summer internships
- Outreach repositories for past and planned events

Staffing

Nic Weber, who had been originally included as a Post Doc in the proposal was hired as an Assistant Professor at the UW Information School. He then became the co-PI on the project and took on the leadership of a number of the major activities. This greatly increased accomplishments, including implementation of a full phase of Capstone projects a year earlier than expected. It freed up PI Palmer to pursue a productive collaboration with the National Data Service to develop a data curation workbench to support technical enhancements in the curriculum. The shift increased Weber's time devoted to project activities and his cost share contribution.

Bill Howe was originally included in the proposal as the Associate Director of the eScience Institute, serving the project as a technical advisor and manager of the eScience Institute's instructional contributions. He was also hired at the UW Information School, as an Associate Professor. This made for temporary shifts in his contributions, with more direct curriculum advising for year 1 but less direct linking to the eScience Institute activities for this year.

Josh Blumenstock left the University of Washington and his role as Data Science and DataLab advisor in the Information School was assumed by Jevin West, Assistant Professor and current co-director of the DataLab.

Bree Norlander was hired in June 2017 as a temporary 50%-time project coordinator.

Financials

External Partner Cost-share

Washington State Department of Transportation (WSHS)

WSHS committed a total (over three years) of \$20,000 in cost-share to the ODL project. At the conclusion of Year 1, they contributed \$7,427.71 in cost-share. We anticipate WSHS to contribute the full amount committed at their current participation level.

State of Washington Office of the Chief Information Officer (OCIO)

OCIO committed a total (over three years) of \$22,000 in cost-share to the ODL project. At the conclusion of Year 1, they contributed \$12,000 in cost-share. This is above and beyond what we expected and we anticipate that OCIO will easily contribute the full amount committed at their current participation level.

Washington State Historical Society (WSHS)

WSHS committed a total (over three years) of \$19,200 in cost-share to the ODL project. At the conclusion of Year 1, they contributed \$573.49 in cost-share. During Year 1 we were unable to craft a feasible Capstone or Internship project that met our specifications for scope and support. We have been working with WSHS in an attempt to create a feasible Capstone project for Year 2. We are also seeking alternative sponsors in an effort to ensure this cost-share is covered.

Seattle Public Library (SPL)

OCIO committed a total (over three years) of \$42,380 in cost-share to the ODL project. At the conclusion of Year 1, they contributed \$6,172.23 in cost-share. While this is less than expected, the summer internship had just begun by the end of Year 1. We anticipate a much greater contribution in Year 2 from the culmination of the summer internship and potential future projects.

Internal Partner Cost-share

Due to personnel changes noted in the interim performance report, there have been shifts in cost-share allocations internally. We expect the cost-share to even out by the end of the grant.

Appendix: Posters, Flyers, Handouts



Open Data Literacy Project

Information School, University of Washington
 Nicholas Weber, An Yan, Carole L. Palmer
 nweber@uw.edu, yanan15@uw.edu, cpalmer@uw.edu



1. Introduction

Open Data Literacy (ODL) is a collaborative project led by the Information School (iSchool) at the University of Washington (UW) in partnership with public institutions that create, manage, and publish open data.

The three-year project (2016-2019) is positioned at the **intersection of two new strategic areas** within the UW Information School:

Data for Social Good and Future of Libraries

The goal is to build a **robust model for innovative data education** to prepare information professionals to improve public accessibility and use of open data, and the data capabilities of both data producers and consumers.

ODL will **advance data literacy** through

- **new curriculum** for iSchool students
- **internships and fieldwork** at public institutions with open data initiatives
- **open education resources** for public libraries and information professionals

2. Partners & collaborators

Chief Information Officer

iSchool Contributors

- Chris Coward
- Mike Crandall
- Bill Howe
- Jevin West

Acknowledgement:
 This work was funded by a grant from the Institute of Museum and Library Services, Laura Bush 21st Century Librarians Program, Grant number: 67-5285

3. Open Data Literacy Model

1. Open Data Curriculum

The Curriculum Team is developing a curricular framework that covers collecting, managing, curating, preserving, and providing meaningful access to open data.

Core Courses

- INFX 591 A: Fundamentals of Data Curation
- INFX 598 D: Advanced Data Curation
- INFX 598: Open Data Infrastructures and Services (under development)

Recommended Courses

- INFX 592: Introduction to Data Science
- INFX 593: Data Science II: Theoretical Foundations
- INFX 594: Data Science III: Machine Learning and Econometrics
- INFX 595: Data Science III: Scaling, Applications, & Ethics
- INFX 598: Digital Preservation

2. Fieldwork Options

Students have three fieldwork options with public sector and iSchool partners.

- **paid summer internships** for intensive 8-week projects
- **directed fieldwork** for 2-6 month projects
- **capstone projects**

Areas of focus for fieldwork projects include:

Data Collections on Civic Concerns	Professional Best Practices
Environment & wildlife monitoring	Open access and data licensing
Natural resources and hazard mapping	Resource interoperability
Regional cultural heritage curation	Digital preservation

To find out more about ODL:
 Please email odl@ischool.uw.edu
 Or visit our website: <http://odl.ischool.uw.edu>

4. Current Progress

1. Curriculum Development

The Fundamentals of Data Curation and Advanced Data Curation courses have been expanded to cover open data topics on transparency, privacy, open civic data, portal architecture, etc., with curation tools and hands-on activities introduced into online and on-campus courses.

2. Capstone Projects

ODL is co-sponsoring four iSchool capstone projects with Washington state agencies in 2017.

16 students from across all three iSchool programs are building tools, publishing data, and creating interfaces for agencies, including the Department of Ecology, Department of Fish and Wildlife, and the Governor's Salmon Recovery Office.

Project Name	Students' Academic Programs	Key words
BEYOND WASTE-OPEN DATA LITERACY PROJECT	MSIM	Open data publishing, visualization, data analysis
OPENING UP THE DATA: VISUALIZING THE EFFECTIVENESS OF PUGET SOUND RESTORATION EFFORTS	MSIM and MLIS	Visualization prototype, open source platform
OPEN DATA AT WASHINGTON DEPARTMENT OF FISH AND WILDLIFE	MSIM and Informatics	interactive information visualization, UX design, data-driven story-telling
WA STATE PUBLIC MEETINGS CALENDAR	MLIS	Open access, cloud database, visualization, open government

ODL Impacts

- **Innovation in curriculum and instruction** - draws on and feeds back to network of professional partners; reaches students and current practitioners; graduates diverse cohorts of professionals prepared to lead and mentor colleagues.
- **Enhanced student learning** - practice-based class work; strongly mentored field experiences; carefully designed real-world projects; engagement with experts in LIS, data science, and public sector information services.
- **Models for open data products and processes** - exemplar data resources; documentation of best practices; case studies produced through field experiences; adaptable and sustainable partnership approach.
- **Strengthened workforce** - information and library professionals with strong expertise and experience; network of professionals active in data initiatives for public good.

5. Next Step

- Implement evaluation plan
- Plan and execute summer internships with partners
- Curriculum development for Digital Preservation and Advanced Data Curation courses
- Launch open data literacy and workforce study

|| **RECOMMENDED COURSES**

INFX 551 A: Fundamentals of Data Curation

Description: Data curation is an emerging area of expertise for information professionals across sectors. This course examines the curation of digital research data for professionals working in institutions involved in data intensive research, scholarship, and innovation. The course covers principles and practices in data management and data services, and emphasizes curation expertise for development of reusable and accessible data resources, within the context of current trends in data sharing and open access.

Example Topics: Institutional and Disciplinary Perspectives on Data Curation; Metadata & Provenance; Data Services; Architecture and Preservation; Attribution and Reproducibility; Sustainability and Costs

INFX 598 D: Advanced Data Curation

Description: This course examines a broad range of practical and conceptual issues in the emerging field of data curation. It focuses on recent advances and challenging problems in the curation of research data across disciplines and new trends in open data resources and services for research communities and the public. The course will draw on current and foundational literature and research, as well as current initiatives, to examine key challenges in the field and practical solutions applied by data professionals. In-depth study of topics will be driven by student interests and their contributions to the course through discussion, current awareness briefings, and their final project.

Example Topics: Data Curation Evaluation; Cultures of Data/Evidence; Inventories, Catalogs, & Interoperability; Data Levels, Provenance, Process; Value & Quality; Enhanced Publications and Complex Research Objects; Physical Samples and Data Rescue; Open Data & Public Libraries

INFX 598 OD – Open Data Infrastructures and Services (under development)

Description: This course will prepare students to design, implement, and sustain open data initiatives. This will be achieved through a combination of lectures and laboratory work. Lectures will be based on case-studies from institutions that have engaged in open data publishing, and address topics such as privacy, transparency, interoperability, trustworthiness, stewardship, and costing & sustainability.

Hands-on laboratory instruction will include two components:

1. Techniques for retrieving, cleaning, and visualizing open data.
2. Designing and deploying open data portals through the use of CKAN software.

The course will be designed around a single group project which will require students to launch an open data portal using the open-source software platform CKAN. This exercise will require students to gather requirements from a potential user community, locate necessary data and metadata necessary to serve that community, and then design policies for publishing and providing access to these collections.

Example topics: Best Practices for Publishing Open Data; Retrieving and Cleaning Open Data; Privacy, Politics, and Economics of Open Data; Describing and Creating typed relationships between Open Datasets; Preservation, and Sustainability of Open Data Initiatives.

Related recommended courses (not required)

INFX 572 Introduction to Data Science (4)

Introduces a broad, non-technical overview of key concepts, skills, and technologies used in "data science". Provides a high-level introduction to common data science pipelines, such as experimental design, data collection and storage, basic analytics, machine learning, and data visualization, focusing on analyzing in real-world datasets using industry standard statistical packages.

INFX 573 Data Science I: Theoretical Foundations (4)

Provides an overview of key concepts, skills, and technologies used by data scientists, including inference; machine learning and pattern recognition; storage and scaling; experimental design; and data visualization.

INFX 574 Data Science II: Machine Learning and Econometrics (4)

Provides skills required to analyze and derive insight from large-scale, heterogeneous data. Covers key concepts of functional and imperative programming for storing, extracting, analyzing, and presenting large data projects; and data analysis using inferential statistics, supervised and unsupervised machine learning. Students gain experience modeling social and behavioral data.

INFX 575 Data Science III: Scaling, Applications, and Ethics (4)

Advanced skills for analyzing and deriving insight from large-scale, heterogeneous data. Provides methods, tools, and frameworks for analyzing data, with topics including scaling and distributed computing, network analysis, and visualization. Student gain experience applying methods to real-world data.

INFX 598 – P Digital Preservation

Description: Libraries, archives, and museums are increasingly charged with the curation and long-term preservation of digital objects – including collections of cultural heritage objects, born-digital scientific data, new digital media (such as video games), and software. This course focuses on tools, techniques, and policies for delivering trustworthy preservation in these settings. Practical activities will include the design and implementation of a digital preservation plan, using established frameworks to audit preservation services, and learning digital preservation standards such as reference models, file formats, metadata schemas, and certifications for trustworthy digital repositories.

Example Topics: a basic introduction to preservation strategies, such as migration and emulation of digital resources; preservation standards and specifications for providing long-term access to digital material; forensic approaches to profiling complex digital objects; sustainability and risk assessment, including digital repository auditing frameworks and certifications; as well as the extension of traditional archival concepts like authenticity, integrity, quality control, and provenance to a digital realm.

we make information work

Last Update: Oct 27th, An Yan



Open Data Literacy (ODL) is a collaborative project led by the Information School (iSchool) at the University of Washington (UW) in partnership with public institutions that create, manage, and publish open data. Current partners include the Seattle Public Library, Washington State Historical Society, Washington State Department of Transportation, and the Washington State Office of the Chief Information Officer.



The overarching goal of ODL is to improve public accessibility and use of open data by increasing the data capabilities of both data producers and data consumers. To meet this goal, ODL will advance data literacy through new curriculum for iSchool students, internships and fieldwork at institutions with open data initiatives, and community workshops and open education resources for public librarians and information professionals.

Curriculum: ODL curriculum will prepare students and practicing information professionals to:

- Curate collections of open data of value to local communities
- Build infrastructures and preservation environments to sustain open data collections,
- Collaborate with open data providers on advocacy and outreach activities.

Student Internships and Fieldwork: UW iSchool students with data curation and data science expertise will be placed in partner organizations, to apply their classroom knowledge to real-world open data projects. Placement options include paid summer internships for intensive 8-week projects and directed fieldwork placements for longer 2-6 month projects. Areas of focus for ODL student projects include:

Data Collections on Civic Concerns	Professional Best Practices
<ul style="list-style-type: none"> ● the environment ● natural resources and hazards ● regional cultural heritage 	<ul style="list-style-type: none"> ● open access and use ● interoperability ● digital preservation

Community Workshops: In collaboration with project partners and UW personnel, ODL will host open data workshops in 2017 and 2018. These public events will engage particular communities in using open data through hands-on instruction and small group activities.

Open Education Resources: Over the course of the project, ODL will offer continuing education opportunities for information professionals to learn how to manage, curate, or preserve open data through webinars, an instructional blog, and use of open-source software.

To find out more about ODL, please visit:

<http://odl.ischool.uw.edu>

For information on institutional partnerships and education and outreach activities, contact:

odluw@uw.edu.

OPEN DATA LITERACY

STUDENT PROJECT DESCRIPTION TEMPLATE

Date:



Organization:

Name of Contact:

Contact Email:

1. General description of the project (100 words or less)

2. What is the topic of theme of the project? *

Natural resources, hazards, or public safety

Cultural Heritage

Community Engagement

Others

Please specify:

3. Has there been any previous work on this project

4. Is there any specific software or data that will be used for this project

5. How long will the project take to complete (estimate in hours):

6. Is there a deadline for the project to be completed?

7. What are the (anticipated) deliverables?

8. Who will be the mentor for this project (name and title):

9. How many students could the project support?

10. What skills will be necessary to complete the project (e.g. Knowledge of SQL databases; Data Management Planning, etc.)?

11. What skills will be learned during the project?

12. Does your organization have funding to support this project?

13. Any conditions or restrictions that may be applicable (team size, restrictions on public display of institution or project info, work authorization restrictions, intellectual property, etc)?

|| STUDENT FIELDWORK OPTIONS

A major part of our grant work will be to broker field experiences between you as project partners and students at the UW iSchool.

1. Capstone Projects

“Students serve as consultants to an organization and identify an information problem or need. They then develop the approach and methods needed to address the problem, conduct the research and present the findings in both oral and written formats. Capstone projects are developed during winter quarter (January-March) and are implemented during spring quarter (April through June).” (For more info see: <https://ischool.uw.edu/capstone>)

Capstone projects are posted on iCareers. Students are required to establish their projects **by Dec. 15, 2016**. Sponsor Networking Events are held in October, and students are looking for projects in Oct. through Dec. on iCareers. Sponsor’s website <https://ischool.uw.edu/capstone/sponsorship> . Official Capstone guide can be found in : <https://www.dropbox.com/s/pc8png1mmuxuv00/2016%20Capstone%20Proposal%20Guide.pdf?dl=0>

For ODL records, we need a **completed project proposal form for funded or unfunded work (we very much welcome feedback on this form)**: https://docs.google.com/a/uw.edu/forms/d/12_MjktvCSDZmmik4F5b68kfErAWRvFZFX920NI4Vsz4/edit

2. Directed Fieldwork

“Students work under the formal guidance from a qualified host site supervisor (mentor) during their DFW experience. Qualified host site supervisors have an MLIS, MLS, MSIM, library media endorsement (for teacher librarians), or other relevant degree as well as the depth of experience and strong professional identity of an information professional. Also, the supervisor should be more knowledgeable than the student for most of the DFW tasks.” (See: <https://ischool.uw.edu/current/mlis/fieldwork-internships/host-site-info>)

Students may pursue fieldwork according to these credits/hours:

- 2 credits = 100 hours
- 3 credits = 150 hours
- 4 credits = 200 hours
- 5 credits = 250 hours

Supervisor responsibilities include:

- Providing appropriate supervision and guidance, including regular (weekly) meetings between the supervisor and student to discuss activities, monitor progress, and discuss any challenges of problems
- Providing relevant tasks, projects and activities that constitute beginning professional level work and involve skills and knowledge that the student has already begun to develop through coursework but for which he/she has not yet developed a high level of expertise
- Working with the student on the following paperwork:
 - The initial Learning Outcomes Agreement
 - The Midterm Report
 - The DFW Supervisor's Final Evaluation

3. Internships

ODL internship will be a new, competitive field experience option for summer placements. Students will work with data mentors at their field site on projects designed to directly benefit partner organizations and local service communities. Paid interns will have additional responsibilities to work with their mentors and ODL team to create professional development resources to be disseminated and used by practicing librarians (e.g. case studies, technical documentation, and best practices guidelines). ODL funding will cover:

- 5 students in Years 1 & 2 - 10 total
- 8 Weeks (40 hours week)

Last Update: Oct 27th, An Yan

we make information work