Northwest Power and Conservation Council

October 1, 2013

MEMORANDUM

TO: Power Committee Members

FROM: Ben Kujala

SUBJECT: Pilot on Web-based Data Management and Publishing

The Council produces a variety of data that are used by staff and other regional planners. Most of these data are stored in Microsoft Excel spreadsheets. Managing these data can be challenging. Staff are exploring new technologies to help manage the data more efficiently and effectively.

The Council staff is proposing to contract with Squishymedia for an amount not to exceed $25,000 to develop a prototype of an online system to store data. The proposed system will allow for viewing data online as well as connecting Excel to the online repository.

The attached proposal provides further details about the proposed pilot effort. I will brief the Committee on this topic at the meeting on October 8.

851 S.W. Sixth Avenue, Suite 1100
Portland, Oregon 97204-1348
www.nwcouncil.org

Steve Crow
Executive Director

503-222-5161
800-452-5161
Fax: 503-820-2370

Data Workflow & Display Pilot Project

Northwest Power and Conservation Council

DRAFT II

August 30, 2013

Prepared by Squishymedia, Inc
116 NE 6th Ave, Ste 300
Portland OR 97232
503.488.5951

For more information, please contact

Eric Miller, Principal
eric@squishymedia.com

Greg Lund-Chaix, Director of Technology
gchaix@squishymedia.com

Document Overview

This document proposes an engagement between the Northwest Power and Conservation Council ("Council") and Squishymedia. The proposed effort is a limited-duration Phase I prototyping effort designed to act as a technical proof-of-concept for streamlining the Council's existing Excel-based workflow and facilitating data management and sharing online.

Project Rationale

Currently, Council staff and partners work with multiple versions of datasets stored in Excel files. These files change over time and need to be shared across departments and organizations. A largely manual workflow exists, primarily based on emailed spreadsheets, but this has three main shortcomings: versioning errors are introduced when multiple versions of the datasets are in circulation, it's labor-intensive for staff, and there's no central record of how the datasets are being used or managed over time.

By centralizing the datasets and management in a freestanding web service, the Council can eliminate versioning issues, reduce staff time needed to coordinate data usage and management since data consumption will become a one or two click process in Excel, and will serve as a foundation for automated data sharing with outside partners.

Summary of Project Scope

This initial engagement will focus on developing a functional prototype to demonstrate technical viability of an upgraded data flow connecting existing Excel documents and workflow to the nwcouncil.org website. The project is also intended to explore options for improved and automated dissemination of datasets online.

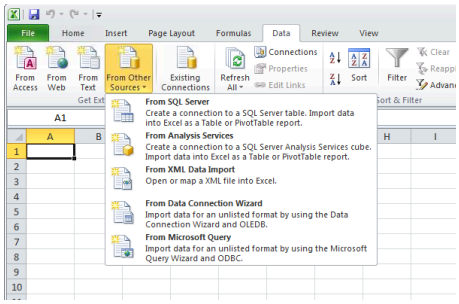
The data set to be used for the prototype is currently posted here:
http://www.nwcouncil.org/energy/powersupply/map

The final deliverable from Squishymedia for this phase will be code required for an interactive prototype displaying the Council's Power Generation Map with extended content and functionality. It will look similar (if not identical) to the current implementation but the improved backend functionality will be sufficient for the Council to evaluate the viability and cost/benefit of a broader implementation of the technical approach.

End User Experiences and Benefits (as envisioned by the prototype)

John is a visitor to the Council web site. John can browse the power sources map, click on an entity on the map, and see a detailed page describing additional information about the power source; this page and information is not available on the current Council website.

Jane is a Council staffer working on a project to generate an annual report. She needs to integrate a list of power sources into the report. She uses the prototype website to browse a list of data sources, including multiple versions of the power source database. She selects a data set from a specific date relevant to the time period covered by the annual report and copies the provided URL. She opens her personal copy of her project Excel workbook and creates a "power sources" worksheet, then selects the "data" tab. She selects to pull data from an XML source via the "Other Sources" tab and pastes in the URL. The data is automatically populated into Excel. She can then work in other worksheets in her workbook to consume and build on the power sources dataset.



Jack is a Council staffer who needs to update information about an entity in the power sources dataset. He uses the online editor[1] to open up the power sources data table and make the edit. Jack is given the option to publish the updated dataset as a new version or a minor undated revision; he chooses to save it as a minor update to the current dataset. As a result, John sees the updates on the map. Jane, however, is working from a historical data set and does not receive the update.

Jeff works for a planning group in another organization. He needs to use the dataset for a project. A Council staffer creates an account for Jeff [2], grants access to the relevant dataset, and sends him the URL and his login information. Jeff is then able to access the dataset in Excel the same way as Jane.

Jane, Jack, and Jerry (another Council staffer) are all working on a project that uses the current power source data set. Jerry finds an incorrect power generation capacity for one entity. Jack uses the online editor to fix the capacity entry and publishes the update to the "_current" data set. Jane, Jack, and Jerry then open Excel and click "Refresh" on the Data tab. The current data set (including the newly fixed capacity entry) is refreshed in Excel for all three of them, with the raw data worksheet being replaced and all linked calculations in other worksheets being automatically updated by Excel.

---

[1] The online editor will be available but may or may not be customized to Council workflows; for purposes of this prototype we may use more generic database management tools.

[2] Only one user account will be supported by the demo, but the prototype's functionality will be designed with an eventual multiuser / multigroup deployment in mind.

Technical Roadmap

Data input: The power sources XLS Workbook (link) manually imported into an online database.

Data output: An upgraded and extended display of the power sources data set using the existing power sources map on the Council website code; a browse interface to allow users to view different versions of the data set; and a web service point for Excel to consume the data set as a worksheet for integration into existing workflows.

Core technical activities:

- Build and deploy an Excel-friendly web service containing the power sources data set
- Edit existing power sources map code to consume the web-based code directly
- Expand the power sources map to display additional data
- Stand up password-protected review server for Council to review prototype

Additional functional elements, time permitting:

- User can view a list of data set files uploaded on file system, browsing by date and comments
- User can preview the map with a given data set
- Support historical data display (such as change over time) within the power source entity detail page
- Online editing functions for the data sets enabled; straight database table edits will already be possible, but a more streamlined approach with Council-specific tools will be useful

It should be understood that the Phase I prototype deliverables are geared primarily as a proof-of-concept deliverable specific to the Power Generation dataset. The scope does not include production-grade development activities applicable to other data sets. It also will not include key functional components required by a production system, such as the ability to create and manage individual user accounts with varying abilities to access and/or edit data sets. That said, certain technical aspects of the subsequent development phase will be very useful in any subsequent production phases, so we'll be working closely with the Council's staff to ensure that we make appropriate technology decisions.

Deliverables

The deliverables at the end of Phase I will be sufficient for detailed review and subsequent decision-making by the Council. The exact technical composition of the final deliverables will be established by mutual agreement between the Council and Squishymedia, but is currently planned to consist of a guided 'tour' of a live web service, with subsequent access provided for review and analysis by Council stakeholders.

Desired Outcomes

By streamlining, standardizing, and modernizing the technical workflow supporting the Council's data sharing efforts, the Council will improve internal staff efficiency and will realize enhanced data sharing and dissemination impact with stakeholders and the public.

- Streamlines data management across the organization by establishing a canonical set of data resources complete with key metadata
- Establishes successful proof of concept for streamlined Web-to-Excel data transfer
- Provides the Council with a functional demo useful for making decisions on subsequent broader adoption of the approach embodied in the proof of concept

After This Effort

Following the delivery of this functional prototype, the Council will have a functional software tool intended to be useful for evaluation and planning purposes. If the Council elects to move forward with the technical approach demonstrated by the prototype, a larger Phase II effort would allow the Council to adapt and expand the technical approach across a broader range of data sets, web-based public data visualization, and data sharing mechanisms. If the Council did not elect to move forward, the prototype could be used as-is to continue updating the power sources dataset online.

Budgets & Schedules

Project activity window: 10 week engagement (approximately)

Estimated project start date: Work can begin upon client acceptance of this estimate and written acceptance of engagement. The kickoff meeting is to be conducted within two weeks of acceptance.

Core objective, Phase I: development of a functional single-purpose prototype to demonstrate technical feasibility of a streamlined data management and display workflow.

Level of Effort: Targeted at 165 hours, not to exceed 180

The following table is an estimation of hourly breakdown and cost. Adjustments may be made by mutual agreement in order to emphasize highest-value activities.

Deliverable	Estimated Timeframe	Hours
Technical audit and requirements	Workweeks 1-2	25
Prototype development	Workweeks 3-10	110
Project management	Continuous	30

Squishymedia proposes to execute the project as described for a proposed Phase I project budget of **$23,100**. Outside expenses (such as licensing or stock photography) are billed at cost.