OCIO Makes Collecting, Analyzing Data Easy with Spreadsheet Manager Tool

OPERATIONS

April 4, 2016

When it comes to working with data—collecting, categorizing, formatting, and analyzing data—there are two kinds of people. "There are people who have lost data, and there are people who won't admit they've lost data," said **Ed Wolfrum**, principal researcher and group manager with the Biomass Compositional Analysis team at NREL's National Bioenergy Center. In other words, keeping large amounts of complex data organized and centrally accessible is no easy feat.
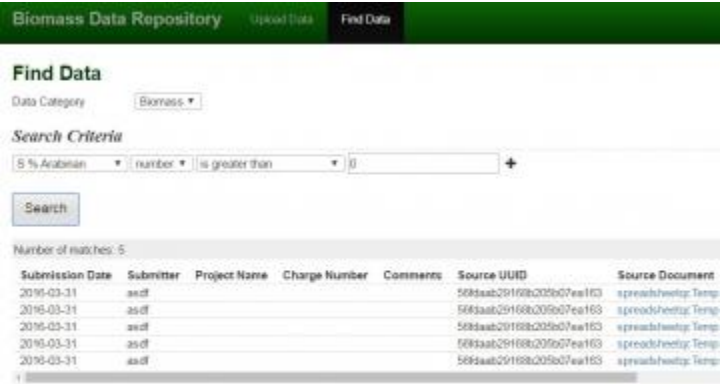
Last summer, the Biomass Compositional Analysis Team was drowning in spreadsheets that weren't searchable—many were in different formats, lacked the metadata to be efficiently located, or couldn't be easily saved. They struggled to create an application to fix the problem, but the tool soared in complexity to handle all of the ever increasing combinations of data formats and types of research data being collected. After eight years, the application was complex, written on a now aging platform, and had high maintenance costs.

Ed finally reached out to the Office of the CIO (OCIO) for help and found the OCIO Application Integrations team, led by **Sheila Hatfield**. After asking for help to modify the old code to add security features so that more individuals could use the system, the OCIO Application Integrations team came back with a better idea. It suggested starting from scratch—offering Ed a simpler design and much more efficient system using lightweight, NoSQL, open-source tools.

The tool can store an unlimited number of spreadsheets, along with metadata, to locate data within the spreadsheets and distinguish them by category, date, or researcher.

Image courtesy of Ed Wolfrum, NREL

Less Complexity, Better Analysis

The resulting Data Repository Web App can store, aggregate, and retrieve spreadsheets or .csv files—regardless of the structure of the data. With a precise focus on researchers' needs, the tool can store an unlimited number of spreadsheets, along with metadata, in order to locate data within the spreadsheets and distinguish them by category, date, or researcher. Consolidating common data from multiple spreadsheets for fast, streamlined analysis, the NoSQL tool set does not attempt to correlate data based on structure. Everything that's uploaded is automatically backed up online, and compared to the previous application, the new tool was developed on a much stronger technology stack.

"The OCIO team did a fantastic job looking at our existing code and providing the best possible solution," said Ed. "We were able to deliberately leave out the stuff we didn't need to stay within our budget. We could not have done this without Sheila and her team." A cost-sharing arrangement was agreed upon to balance resource needs for another small project on the OCIO side as well, enhancing the cost-effectiveness of the tool for Ed's team even more. Further, the application can easily be copied and configured to suit other data management needs.

Sharing Solutions and Furthering the Mission

Now that the tool is meeting Biomass needs, one of Ed's goals is to allow other NREL Centers to use it as well. "Anyone can benefit from what we've already created, using the application as a foundation and then configuring or even enhancing their version."

The tool offers an adaptable space to efficiently collect and organize historical data sets across any data domain, from project management, Environment Health and Safety inventory, or communications analytics. The NoSQL tool set doesn't care what structure the data has, it makes analysis correlations on the data that matches name and type. Infrastructure can be scaled to handle large data sets, as the application has been built to perform when scaled. There is no cost for NREL centers to license the application, other than finding infrastructure and storage to host the application.

"Spreadsheets are here to stay," said Sheila. "As long as the application is maintained, this kind of repository tool will last as long as anyone chooses to use it."

Innovations and insights often come from analyzing old data in different ways. The new spreadsheet management tool allows for this analysis without the monotonous task of re-formatting all the data to match structure or even finding it from colleagues long gone off to other adventures.

Attend a Demo

Sheila Hatfield (mailto:sheila.hatfield@nrel.gov), Ed Wolfrum (mailto:ed.wolfum@nrel.gov), and Craig Laurin (mailto:craig.laurin@nrel.gov) will host a demo on the Data Repository Web App **2:30-3:30 p.m. Wednesday, April 27, in the Integrated Biorefinery Research Facility Aspen conference room**. You can also contact them directly with questions.

2 Likes   Give this thumbs up (/flag/flag/thumbs_up/64030?destination=node/64030&

token=K_Ix3aFC7PCnbQFVQIZxTMStgnqLlios4La1OI6Qylk)

Comments

Submitted by shatfiel on April 4, 2016 - 11:34am

A quick comment that this spreadsheet application was re-written by Mike Brown (mike.nrel.related@superbrown.com (mailto:mike.nrel.related@superbrown.com)) who did a majority of the design and development. Unfortunately our budget did not allow for him to stay after this project wrapped up, but he can be reached at the personal email address above.

reply (/comment/reply/64030/626)

Add comment

Guidelines for commenting (/commenting-guidelines)

Your name shatfiel

Comment *

☑ Notify me when new comments are posted

Save comment   Preview

ADVANCING THE MISSION » (/APRIL-4-2016/SECTION/14)

- Exec Connect: Strategic Investment—NREL Links Personal Energy Choices with Optimized Systems (/2016/04-04-16/science-and-technology/exec-connect-strategic-investment%E2%80%94nrel-links-personal-energy)

- NREL Raises Rooftop PV Technical Potential Estimate (/2016/04-04-16/analysis/nrel-raises-rooftop-pv-technical-potential-estimate)

- Bioenergy Highlight: High Titer and Yields Achieved with Novel, Low-Severity Pretreatment Strategy (/2016/04-04-16/science-and-technology/bioenergy-highlight-high-titer-and-yields-achieved-novel-low)

- Transportation Highlight: New Analysis Methods Estimate a Critical Property of Ethanol Fuel Blends (/2016/04-04-16/science-and-technology/transportation-highlight-new-analysis-methods-estimate-critical)

- Cities-LEAP Tool Can Inform City Leaders' Energy Decisions (/2016/04-04-16/analysis/cities-leap-tool-can-inform-city-leaders%E2%80%99-energy-decisions)

- NREL's Capabilities Boost a Wide Range of Innovative ARPA-E Research (/2016/04-04-16/science-and-technology/nrels-capabilities-boost-wide-range-innovative-arpa-e-research)

- NREL Publications Released March 2016 (/2016/04-04-16/operations/nrel-publications-released-march-2016)

- OCIO Makes Collecting, Analyzing Data Easy with Spreadsheet Manager Tool (/2016/04-04-16/operations/ocio-makes-collecting-analyzing-data-easy-spreadsheet-manager-tool)

- Lab-Corps Program Announces Second Round of Training: NREL Sends Two Teams (/2016/04-04-16/commercialization-and-deployment/lab-corps-program-announces-second-round-training)

- Analysis Coffee Break: Spatial Analysis & Data Visualization Web Tools (/2016/04-04-16/analysis/analysis-coffee-break-spatial-analysis-data-visualization-web-tools)

- Lessons Learned: The "Why" and "How To" (/2016/04-04-16/operations/lessons-learned-%E2%80%9Cwhy%E2%80%9D-and-%E2%80%9Chow-%E2%80%9D)

OTHER SECTIONS

NRELIANS » (/APRIL-4-2016/SECTION/15)

ON CAMPUS » (/APRIL-4-2016/SECTION/16)

WORK, LIFE, PLAY » (/APRIL-4-2016/SECTION/17)

ANNOUNCEMENTS » (/APRIL-4-2016/SECTION/18)

CONTRIBUTE TO NREL NOW

- Submit story or story idea (/node/add/story-or-idea?advagg=-1). Read Submission Guidelines (/submission-guidelines) for more information.

- Create a classified ad (/node/add/classified).

CLASSIFIEDS » (/CLASSIFIEDS)

SUBMISSION GUIDELINES » (/SUBMISSION-GUIDELINES)

ARCHIVES » (/ARCHIVES)