CI-Data for WC-WAVE

Title

Data Management Platform and Adapter Development Component of the WC-WAVE project

Outcome

By automating the data acquisition, upload and documentation process the Virtual Watershed Platform (VWP) minimizes the data management effort of modelers when they execute their models. The outcome of this is that the time it takes for the products of a given model run to be available for further research (subsequent model runs, visualization of results, sharing with collaborators or more broadly) can be reduced from days or weeks to minutes, limited primarily by the time it takes for the data to be transferred to and from the VWP to the modeling system.

Image Credits:
William Warby. "Stopwatch". https://flic.kr/p/62n3eY. CC-BY 2.0
Dafne Cholet. "Calendar*" https://flic.kr/p/9bUbH3 . CC-BY 2.0
U.S. Department of Agriculture, USDA photo by Scott Bauer "k5060-2".
  https://flic.kr/p/dWVv1o. CC-BY 2.0

Impact / Benefits

By minimizing the dependency upon intermediate effort from researchers the data management loop is shortened to the least time possible, yielding more efficient and impactful data intensive research.

Explanation

The effective and efficient management of data throughout the research process is essential for maximizing the utility of those data within a given research project and the impact of those data beyond the life of a project. In data intensive research access to required data, documentation of research processes, and accelerated sharing of generated data and associated documentation all potentially pose barriers, that when lowered, can increase the efficiency of research and the impact of research data. The CI-Data (Data Management Platform and Adapter Development Component of the WC-WAVE project) has made significant progress in lowering these barriers through the development of a loosely-coupled data access and ingest system that automates the process of acquiring the data required to execute a watershed-scale model (iSNOBAL), running the model, and transferring the model outputs to a web-accessible data management, discovery and access platform – the Virtual Watershed Platform (VWP) developed by the project team.

The VWP provides a robust shared data management, discovery and access system that enables use of data access, documentation, and data processing services by any authorized computer that supports interaction through standard web protocols and open standards. Through the use of custom *adapters* that communicate with the VWP's web services, any model may obtain data from the VWP to execute a model run and push model data and associated documentation back into the VWP for near-real-time discovery and access by other models, visualization systems (as are being developed by the WC-WAVE CI-Vis Development team), and external data 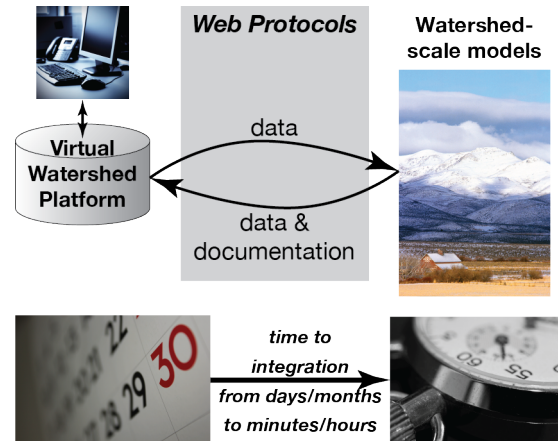network (such as the DataONE and Data.gov systems).