Velo: A Knowledge Management Platform for Modeling and Simulation

Molly Kessie

Pacific Northwest NATIONAL LABORATORY

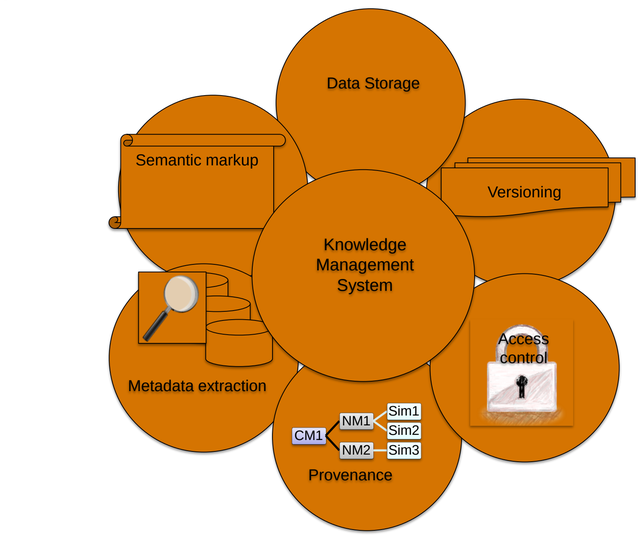
Proudly Operated by Battelle Since 1965

INTRODUCTION

Many scientific domains involve modeling and simulation and share a common need for data and knowledge management. The Data Intensive Scientific Computing Group at PNNL have created Velo, a software tool that can easily be customized for different scientific demands.

VELO

Velo is a knowledge management platform for models and simulations that integrates collaborative and content management technologies to create an accessible and flexible core platform to meet the needs of specific scientific domains. The collaborative nature of Velo is built upon MediaWiki which allows for its users to access all of the functionalities within Velo.



CAPABILITIES

- Support collaborative teams of modelers working on common problems from different geographic locations
- Acquire, organize, query, and share data and observations
- Use data and observations to develop computational models
- Metadata extraction
- Provides customization of model inputs for specific projects
- Permit simulations to be launched on distant computational platforms
- Integrate external tools to allow for model development, specific visualization and analysis capabilities

References:

[1] Ian Gorton, Chandrika Sivaramakrishnan, Gary Black, Signe White, Sumit Purohit, Michael Madison, and Karen Schuchardt. 2011. Velo: riding the knowledge management wave for simulation and modeling. In Proceeding of the 4th international workshop on Software engineering for computational science and engineering (SECSE '11). ACM, New York, NY, USA, 32-40.

*[2] **I. Gorton**, G. Black, K. Schuchardt, C. Sivaramakrishnan, S. Wurstner, P Hui. 2010. GS3: A Knowledge Management Architecture for collaborative Geologic Sequestration Modeling, in 43rd Hawaii International Conference on System Sciences, Knowledge Management Track, January 2010*

[3] Ian Gorton, Yan liu, Maria Vlachopoulou, Tim Seiple, and Nino Zuljevic. 2011. Keeping it Local: The Challenges of Regional Earth System Modeling.

Acknowledgements:

Ian Gorton (Mentor)

Maria Vlachopoulou (Programming Mentor)

For more information about the science you see here, please contact:

Molly Kessie or Ian Gorton

Pacific Northwest National Laboratory

P.O. Box 999, MS-NO

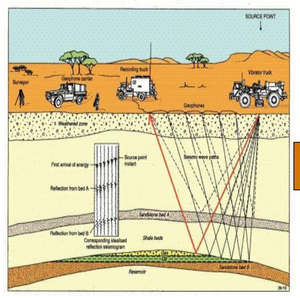
Richland, WA 99352

(509) 375-3850

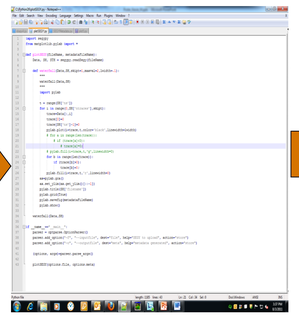
Molly.kessie@pnnl.gov or ian.gorton@pnnl.gov

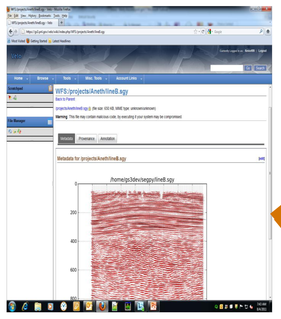
MY CONTRIBUTION

Velo is a software framework designed to support scientific modeling and simulation. Velo can be easily customized for distinct scientific domains, such as climate modeling and carbon sequestration, and can be extended to integrate tools and data types that are relevant to the scientists. In various modeling and simulation domains, data is collected in many different file formats and I was responsible for creating scripts in Python to handle distinctive data types and associated metadata. Specifically I worked with seismic data, which is important in the mapping of underground rock layers. This then gives the geoscientists the necessary tools to create 3D geologic models.



http://mines.industry.qld.gov.au/geoscience/what-are-seismic-surveys.htm





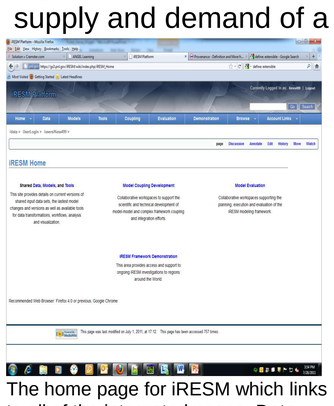
- Seismic data from site surveys is collected and stored as SEG Y files.
- Python scripts were written in order to read, query, and extract the data.
- The python scripts are then integrated into the Velo to handle the multi-step data extraction process.
- When a SEG Y file is uploaded to the Velo, the python script automatically runs and extracts the relevant data and metadata
- The python scripts then create a metadata page within the Velo with relevant information for the scientists.

FUTURE WORK

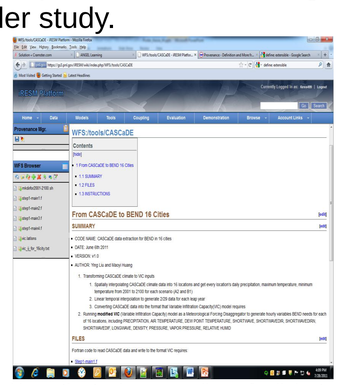
Field data is recorded and saved in various file formats. For modeling and simulation purposes it is important to be able to process all data types (data saved in various file formats). Velo will continue to be integrated with scripts in order to manage and read all data types. Integrating this functionality of reading and processing different data types is key for quick analysis of certain data

iRESM

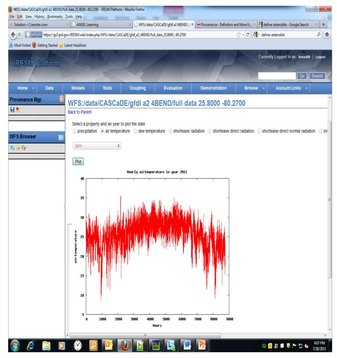
iRESM (integrated Regional Earth System Model) is a collection of combined models working with high resolution data that are able to represent the climate, geography, economy, and energy supply and demand of a region under study.



The home page for iRESM which links to all of the integrated pages: Data, Models, Tools, Model Coupling Development, Model Evaluation, and iRESM Framework Development.



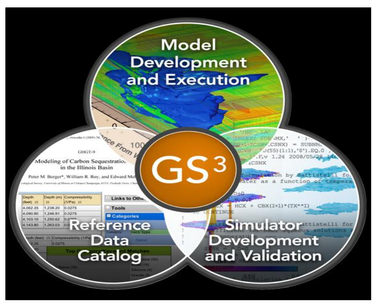
Here, the tools page, allows for links to all of the code for different models. In this example the scripts given are to transform CASCaDE data into VIC inputs.

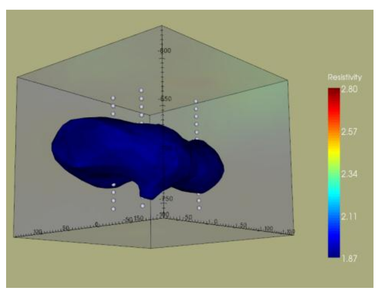


This page is created when a file is uploaded to the iRESM wiki. The python script extracts certain data and creates a Metadata page.

GS^3

GS^3 (Geologic Sequestration Suite) is a software that supports the study of geologic sequestration.





CAPABILITIES

- Gather and interpret field and experimental data
- Integration of off-the shelf technologies
- Collect, organize, and effectively query or mine data from diverse sources
- Support wide-scale model development
- Provenance tracking
- Metadata Extraction of multiple data types
- Job Launching (launching a simulation on a remote computer)

This material is based upon work supported by the S.D. Bechtel, Jr. Foundation and by the National Science Foundation under Grant No. 0952013 and Grant No. 0934785. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the S.D. Bechtel, Jr. Foundation or the National Science Foundation.

U.S. DEPARTMENT OF ENERGY · STAR · NSF · Washington State University World Class. Face to Face. · CESAME · S. D. Bechtel, Jr. Foundation Stephen Bechtel Fund

www.pnl.gov