



Pacific Northwest
NATIONAL
LABORATORY

Proudly Operated
by **Battelle** Since 1965

Benefits: Have Green provides real-time analysis of large graphs with up to one million nodes, on a modest desktop computer, and scalable to address larger needs. This broad architecture is tailored to client needs for different types of network graph-related problems.



National Visualization and Analytics Center™

VISUALIZING SEMANTIC GRAPHS

A semantic graph is a network of heterogeneous nodes and links annotated with a domain ontology. In information analysis, semantic graphs are generated and applied in a visual analysis approach known as link analysis. Through link analysis, investigators query, draw, lay out and link people, facts, locations, events, objects and data in hopes of discovering key trends, patterns and insights. In today's analysis environment, however, users are bombarded by massive amounts of information from a multitude of sources. The vast amounts of information being fed into semantic graphs may easily overwhelm an analyst's cognitive capacity.

The Pacific Northwest National Laboratory is developing a new visual analytics capability that interactively analyzes semantic graphs with up to one million nodes. Our objective is to develop graph-based tools and environments that will enhance analysts' natural analytical capabilities to create, comprehend and analyze large semantic graphs—allowing analysts to effectively and efficiently perform in an information world that grows more complex daily.

APPROACH

Have Green is a suite of visual analytics technologies developed at PNNL to support analysis of large semantic graphs.

Once a semantic graph is ingested, it enters one of two cycles: 1) the cycle of analytical discourse of query, navigation and visualization if the graph is ready in terms of size and format or 2) the cycle of computation discourse if the graph needs further processing. *Have Green* also has a knowledge base component to supplement information that is lacking in the graphs.

The design requirement of *Have Green* is enormous but manageable. We champion software reusability and practice modular design throughout the development stage. After the architecture is established, individual components are implemented separately so that we can pinpoint design weaknesses in the earliest stage. Each component undergoes multiple usability studies. Results collected from the studies and post-study interviews are used to further revise our designs. These individual components eventually become the foundation of *Have Green*.

IMPACT AND OUTCOMES

To date, we have developed four major system prototypes to support *Have Green* components.

Greenland is our first prototype designed to support the *Have Green* system. Signature vectors extracted from a graph are projected onto a low-dimensional scatterplot through the use of scaling. Brushing, linking and clustering are used extensively to cross-examine different visualizations created by different signatures.



U.S. DEPARTMENT OF
ENERGY

Early Development

Lab Prototype

Commercial Product

September 2009

While Greenland provides a way to browse a large graph and look for clues, Green-Sketch, our second component, provides a graphical interface for users to query graphs.

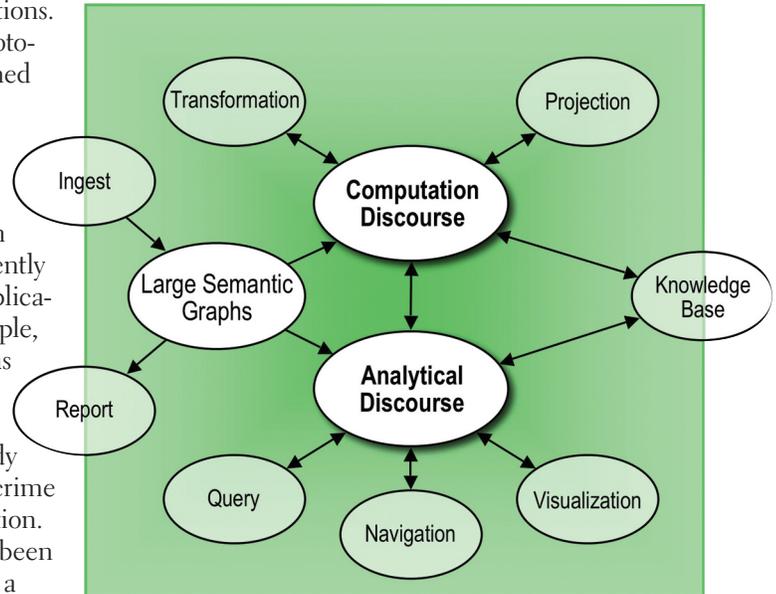
A hallmark signature of a semantic graph is the rich information associated with its individual nodes and links. This graph metadata ranges from a short phrase to a full sentence to an entire paragraph and beyond. We have developed a practical prototype, known as GreenArrow, which allows users to browse this metadata interactively.

GreenMonster is our latest Have Green addition that addresses the scalability issue of our large semantic graphs. The requirement is to provide a capability to visualize semantic graphs with up to one million nodes adaptively and interactively on both desktop computers and PDAs. GreenMonster supports analytic discourse through exploratory

navigation. GreenMonster is currently undergoing evaluation.

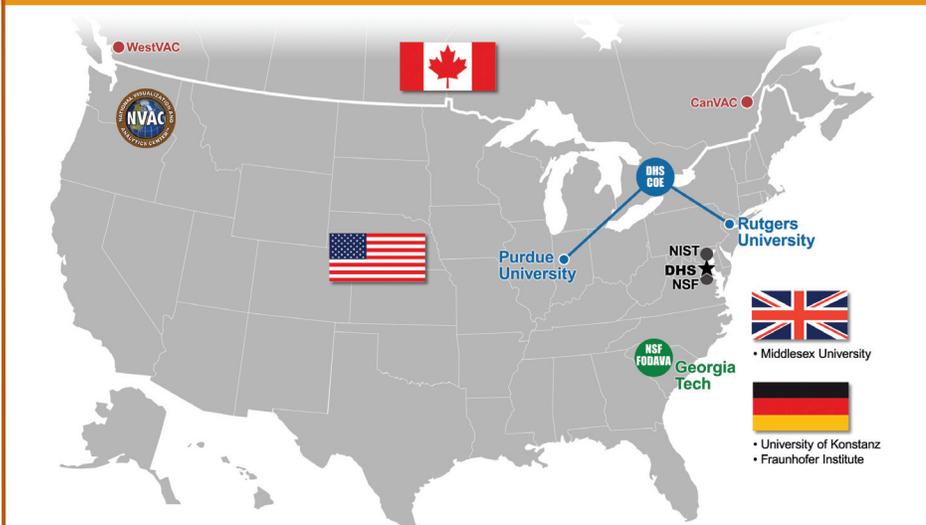
The design of individual Have Green prototypes has been carefully studied and system usability has been thoroughly investigated and reported in various publications. While these prototypes are designed to eventually become part of the Have Green system, some have been used independently for different applications. For example, GreenArrow has been used to support a story-telling case study that involves a crime scene investigation. Greenland has been used to support a

PNNL study on threaded dialogues. Together these components form an innovative graph analytical platform that allows developers to uniquely customize graph analytical tools for different applications.



A framework overview of Have Green

Visualization and Analytics Centers



VAC Consortium Members



For more information, contact
 Pak Chung Wong
 Phone: (509) 372-4764
 pak.wong@pnl.gov

Bill Pike
 Phone: (509) 375-2689
 william.pike@pnl.gov

Project Funding
 U.S. Department of Homeland Security

ABOUT PNNL

The Pacific Northwest National Laboratory, located in southeastern Washington State, is a U.S. Department of Energy Office of Science laboratory that solves complex problems in energy, national security and the environment, and advances scientific frontiers in the chemical, biological, materials, environmental and computational sciences. The Laboratory employs more than 4,200 staff members, has a \$918 million annual budget, and has been managed by Ohio-based Battelle since 1965.

Contact

Richard May
 National Visualization and Analytics Center Director
 Pacific Northwest National Laboratory
 P.O. Box 999, MSIN J4-32
 Richland, WA 99352
 Phone: (509) 375-6976
 richard.may@pnl.gov
 nvac.pnl.gov



Pacific Northwest
 NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965