

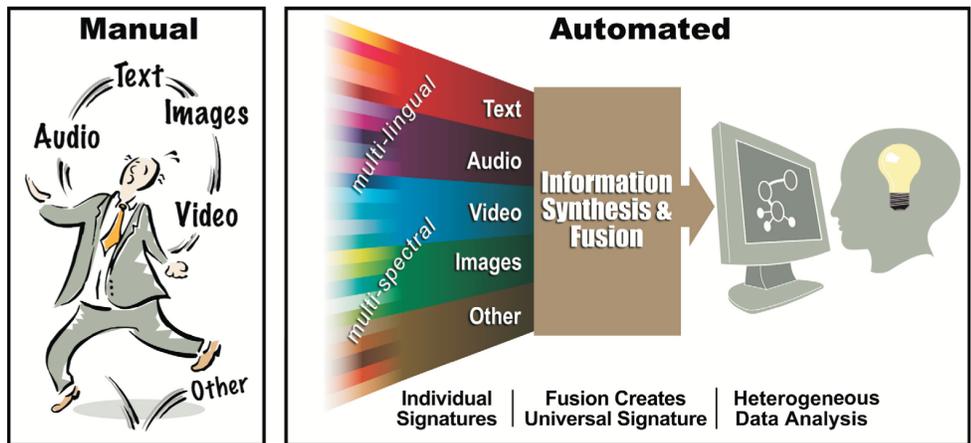
Benefits: To overcome the challenge of sifting through enormous amounts of heterogeneous data, information synthesis and fusion enables analysts to integrate diverse data types and focus on the content and meaning of data rather than on its structure. This effort provides foundational algorithms that will underlie all areas of information analysis.

National Visualization and Analytics Center™

INFORMATION SYNTHESIS AND FUSION

MISSION NEED

Analysts are challenged with not only the deluge of data they must sift through and refine but also the reconciliation of different types of data (such as text, imagery, audio, and maps). As a result, during the analysis process analysts must manually fuse these information types, mentally juggling all these different pieces and determining how they connect. This process can lead to loss of information and traceability in developing the analytic product. Our goal is to develop mathematical models and methodologies that merge information with different conceptual representations into a common semantic space. This will enable analysts to focus on exploring, discovering and analyzing the data, rather than integrating it. Furthermore, gaps in the information space can be filled to give a richer understanding of the content space. We believe that through the development of these models and methodologies it will be possible to not only identify the gaps in the current generation of semantic signatures but also suggest alternative pathways where research should be pursued. Our approach will help unify the various technologies developing signatures, which we believe is critical for interoperability. This work will strengthen the mathematical foundation of information fusion for visual analytics.



APPROACH

The fusion process reconciles the semantic signatures (feature characterizations) of different information types into a common basis. Signatures for different data types have different information content. For example, imagery signatures can be constructed in many different ways, based on either image features, color components,

and image compression or semantic content. However, semantic signatures are universal. Therefore every signature developed for various data types requires that it be characterized by the semantics of the information.

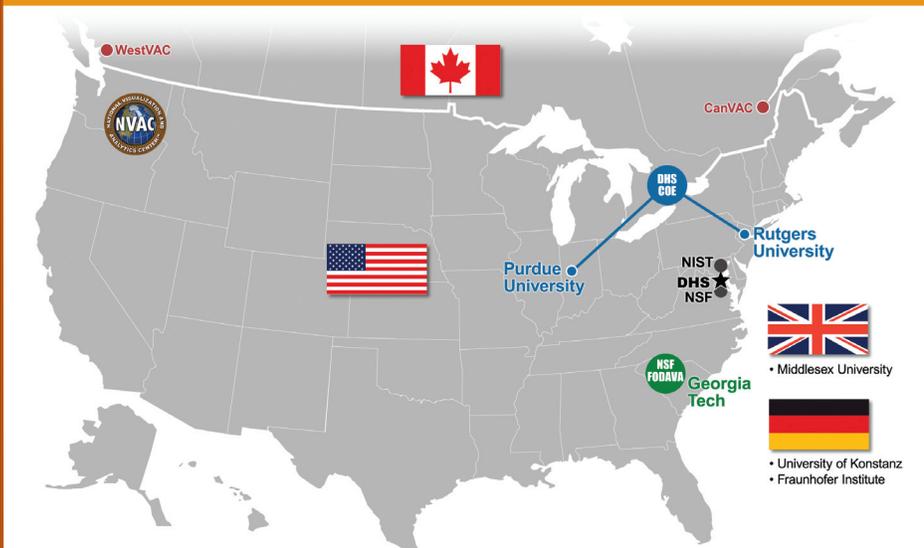
We then identify or generate semantic reference structures (i.e., semantic bridges), which enable us to fuse or reconcile both spaces, thereby creating a common space where semantics between the spaces hold (i.e., a joint context is formed between the spaces). From this common space, a new signature for each data type can

be generated that respects both its original space and the fused one. The resultant signatures can now be used to enhance other analytical tools and provide a better picture of how data fits together. Note that once the spaces have been fused, the other type-dependent characterizing features (e.g., affect in text or nonverbal noises in audio) that aren't semantically based can be used to support the process of synthesis—the generation of new data types. To date, we have developed the models and are verifying the various approaches to different data types.

IMPACT

Addressing this fundamental problem for heterogeneous data analytics will provide the foundation for a new area of science that will benefit all applications of information analysis. Analysts at the U.S. Department of Homeland Security, U.S. Immigration and Customs Enforcement, other government agencies, and industry all use and depend on heterogeneous data sources.

Visualization and Analytics Centers



VAC Consortium Members



For more information, contact
Shawn Bohn
Phone: (509) 375-2574
shawn.bohn@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