

PNNL invests in a portfolio of internal research activities to deliver transformational science and technology, accelerate innovation, develop new partnerships for national and international impact, transform our scientific and technical workforce, and nurture and evolve our core capabilities. Our Laboratory Directed Research & Development (LDRD) program is the principal mechanism for renewing capabilities within PNNL and bringing forward novel ideas that will become the next generation of science and technology. Current and previous investments that create new analytic algorithms, tools, and methodologies include:

- » **Analysis In Motion (AIM) Initiative.** Technology advances have enabled the continuous capture and streaming of data from an increasingly diverse set of sources, from scientific instruments, to the web, to consumer electronics. Timely discovery, evaluation, and response to key insights from these streams require new computing capabilities and a rebalancing of human and computer effort. AIM is developing analysis and interaction technologies to enable humans to respond more quickly and accurately to data streams reflecting changing events.
- » **Asymmetric Resilient Cybersecurity (ARC) Initiative.** ARC is delivering the theory, processes, methodologies, and algorithms that will enable a resilient cyber infrastructure with an asymmetric advantage, thwarting adversaries who seek to infiltrate and damage our national security through digital means. ARC's science-based approach comes from research and development in the theory of resilience, models and metrics that inform the quality of resilience, and methods that provide validation of the approach.
- » **Signature Discovery Initiative (SDI).** SDI is designed to deliver a systematic domain-agnostic process to rapidly discover signatures. We hypothesize that it is possible to describe a generalized signature discovery process to efficiently and robustly construct candidate signatures, validate their reliability, measure their quality, and overcome challenges associated with the detection of those signatures despite dynamic conditions, measurement obfuscation, and "noisy" data environments. Resulting data has proven the methodology to discover signatures that are more reliable, more efficient, and/or more rigorous.

ABOUT PNNL

Interdisciplinary teams at PNNL address many of America's most pressing issues in energy, the environment, and national security through advances in basic and applied science. Founded in 1965, PNNL employs 4,400 staff and has an annual budget of nearly \$1 billion. It is managed by Battelle for the U.S. Department of Energy's Office of Science.



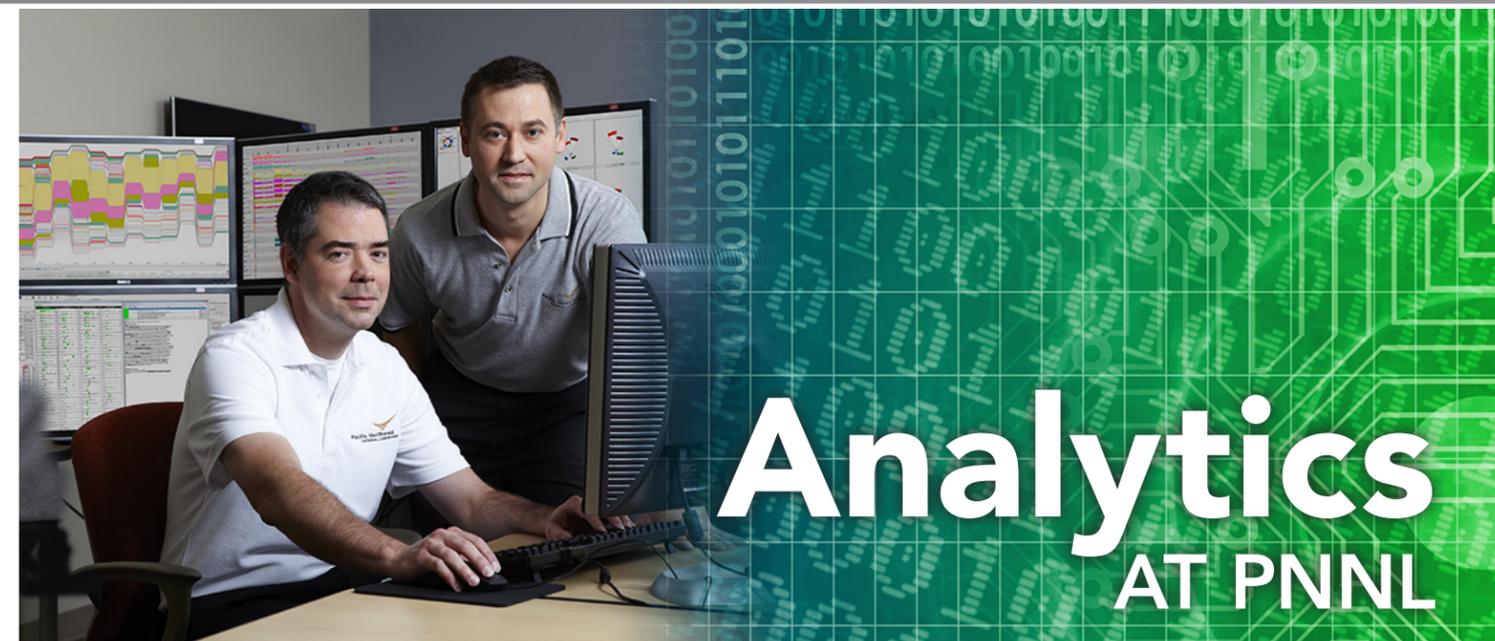
For more information, contact:

Dave Thurman
Director, Computing Programs
National Security Directorate
dave@pnnl.gov
(206) 528-3221

Visit us online at analytics.pnnl.gov



Proudly Operated by **Battelle** Since 1965



The Pacific Northwest National Laboratory (PNNL) is a leader in large-scale data analytics and visualization. Building on its' decades-long foundation of visualization tools for text analytics, PNNL continues to advance innovative new technologies in visual analytics, knowledge-based systems, data modeling and analysis, and high-performance computing. PNNL's analytic tools are used in a variety of application domains, including cyber systems defense, power grid control systems, intelligence analysis, climate change, and scientific exploration.

Applications in these domains are complex—they involve vast amounts of different kinds of data that must be collected and analyzed at high speeds. Commercial

tools often struggle to provide situational awareness in rapidly changing situations, to address the heterogeneity of data involved, or to incorporate human-in-the-loop decision making. Data and information analysis technologies developed at PNNL are robust and flexible enough to help calculate contingencies for power grid operations, identify malicious cyber activity in large networks, and visualize themes and relationships in vast collections of text, audio, and video. PNNL strives to understand its partners' most pressing mission needs and combine its deep domain knowledge to address them through the creative integration of advanced algorithms coupled with next-generation hardware and software technologies.



PNNL analytics activities leverage a broad set of technical capabilities that we apply to address government and industry needs and challenges. Our interdisciplinary teams of experts partner with other national laboratories, universities, and industry to develop innovative tools that meet the challenges of securing the nation and preventing terrorism.

APPLIED STATISTICS AND COMPUTATIONAL MATHEMATICS

We use novel data-analysis methods to extract hidden features, anomalies, and signatures from high-dimensional, large-volume, multimodal data in support of scientific discovery and quantifiable decision-making. Complex mathematical and stochastic models are developed to represent physical, chemical, biological, and nuclear phenomena, while experiments and sampling campaigns are used to increase confidence and explicitly manage and quantify uncertainty.

DATA INTENSIVE SCIENTIFIC COMPUTING

We create reusable software frameworks, platforms, and tools that enable management and analysis of large-volume data, set up complex models, execute large-scale simulations on high-performance computing platforms, and analyze the results. New innovations are created in distributed software architectures, workflow frameworks, knowledge management platforms for modeling and simulation, and system-level tools for high-performance data capture and processing.

HIGH PERFORMANCE COMPUTING

We develop algorithms that are scalable, resource-efficient, and load-balanced, and manage computational complexity and exploit space-time locality. We also create programming models, numerical libraries, communication libraries, compilers, and debuggers that support data decomposition, low communication overhead, and portability. Our efforts increase the ease of use and availability of high-performance computing for non-specialists through domain-specific problem-solving environments.

DATA SCIENCES AND ANALYTICS

Our expertise in analytical reasoning, social and behavioral science, natural language processing, semantic technologies, and human-information interactions is fundamental to our “data sciences and analytics” capability. At the core of our work is to identify, extract, represent, organize, and synthesize data to create and manage knowledge. Knowledge management is fundamental to the research process, enabling our staff and associated projects to span a wide array of technical disciplines and application areas.

SCALABLE DATA MANAGEMENT SOLUTIONS

We emphasize architecting integrated solutions from data capture to knowledge generation, closely embedded into the scientific research processes. These solutions are facilitated through data quality management, metadata, provenance, semantic technologies and knowledge management systems. The solutions offer data capture, annotation, storage, access, assessment, analysis, multi-source integration, data sharing, publication and curation.

SOFTWARE ENGINEERING AND ARCHITECTURES

PNNL uses a risk-based, “right-sized” approach to deliver customized solutions involving complex information system architectures that enable innovation in scientific research, analysis systems and business processes. We are the recognized leaders in the research, construction and

deployment of complex software-centric, complex operational systems. Our expertise in combining adaptive architectures with solid engineering practices provides next generation research capabilities and cutting edge sustainable solutions to our clients across the entire product lifecycle.

VISUAL ANALYTICS

We make sense out of data. We’re a team of visual analytics developers and researchers who love making massive, dynamic, ambiguous, and often conflicting information useful through great visual design, compelling interaction, support for sound analytic methods, and solid engineering. We help people discover patterns, trends, relationships, and events in complex data. We care about the sensemaking process – how new analytic workflows and reasoning strategies help people make sound decisions from their data. We invent new visual metaphors, create analysis algorithms, and deliver software products that put useful analytic capabilities into our users’ hands.

