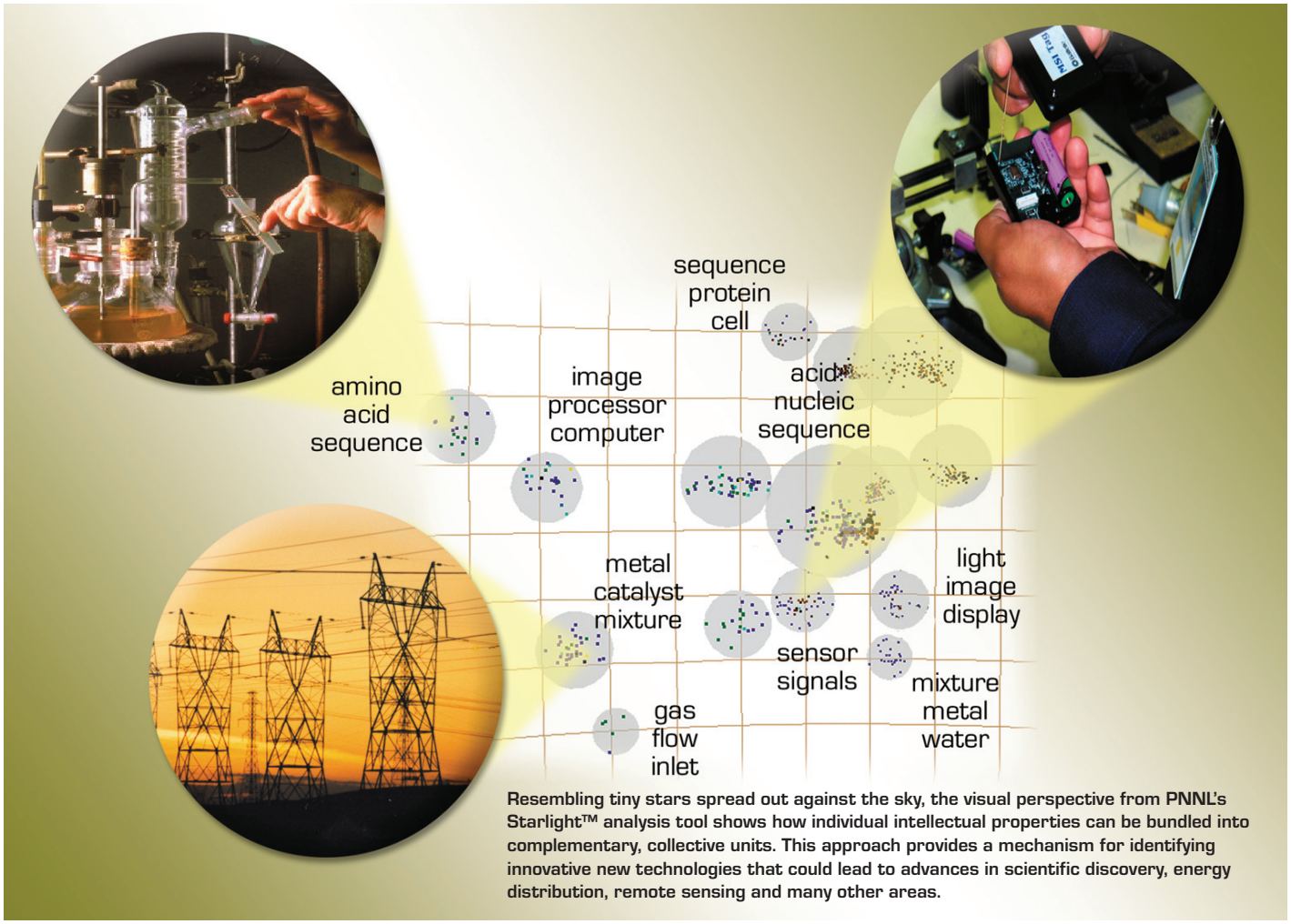


Resembling tiny stars spread out against the sky, the visual perspective from PNNL's Starlight™ analysis tool shows how individual intellectual properties can be bundled into complementary, collective units. This approach provides a mechanism for identifying innovative new technologies that could lead to advances in scientific discovery, energy distribution, remote sensing and many other areas.

Taming Data Through Visualization

Clustering information can unveil common themes and reveal hidden relationships. BY JUDITH GRAYBEAL

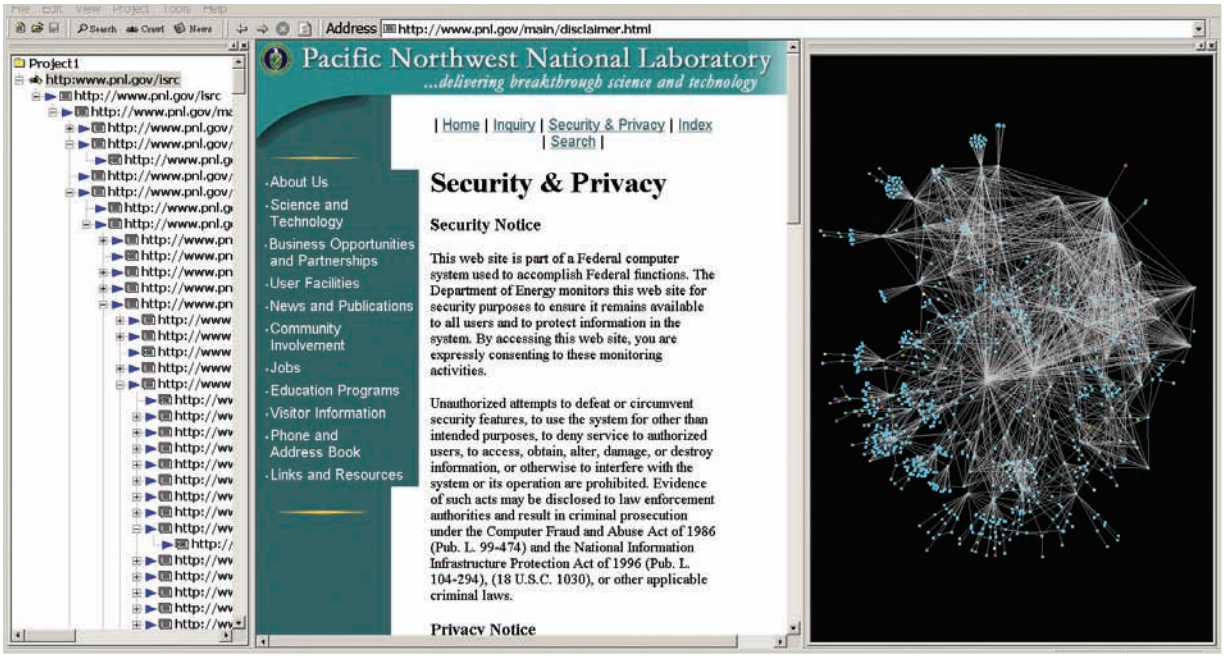
THE ABILITY TO EXTRACT INFORMATION from large quantities of data, process it quickly, and use the results is extremely important to government and industry. Computational science and information researchers at Pacific Northwest National Laboratory are addressing this challenge on several fronts, including visualization analytics, airline safety and cyber security.

Every day private and public analysts confront huge amounts of data from which to draw conclusions, find relationships and discover hidden trends. Visualization analytics tools let users "discover the unexpected," said Gary Morgan, a portfolio manager for PNNL's Technology Commercialization activities. "Large quantities of data can be quickly and easily analyzed to improve competitive performance."

Two technologies developed at PNNL are IN-SPIRE™ and Starlight™. IN-SPIRE quickly and automatically conveys the gist of large sets of text documents such as technical and patent literature, marketing and business documents, web data, accident and safety reports, newswire feeds and message traffic. By clustering similar documents together, this Windows-based software unveils common themes and reveals hidden relationships within the collection.

The Starlight software analyzes large multi-media datasets and displays information so users can recognize subtle relationships. In Starlight, maps, timelines and images can be integrated into the visualization to assist in recognizing relationships among multiple datasets. Potential applications include patent analysis, financial analysis, competitive intelligence research, sales data for consumer acceptance analysis, and legal research.

For example, a major manufacturer employs Starlight to analyze data for product performance and consumer satisfaction. In one instance, analysts used both structured company data and unstructured web-based textual data to reveal why



Within minutes, PNNL's Mozart provides users with a listing of all web pages found under the URL (left), a closeup of any web page requiring further scrutiny (middle) and a hyperlinked 3-D graphic representation (right) of all inbound and outbound web page links, with the selected page highlighted.

product performances were lagging in a specific region of the country. Starlight enabled analysts to rapidly determine the cause so the business unit could fix the situation.

Other applications allow businesses to identify safety and market trends and compare their internal data compilations with data and responses generated externally. The ability to compare multiple types of information helps companies to better understand their customers and to develop strategies in line with customer preferences. According to Morgan, this advantage has the potential for generating millions of dollars in sales and for saving more millions from being spent in the wrong areas.

PNNL researchers are constantly creating and updating visualization technologies to provide ever more efficient information-gathering and analysis. "This generation of tools will lead to the next variations, so that we can continually provide our clients with the best possible information technology available," Morgan said.

Many information technologies began as research projects in the areas of safety and security. For example, a proactive monitoring tool called The Morning Report is being used to improve aviation safety. Although the typical black box technology does a great job of analyzing small portions of flight data from a catastrophic incident, it is not an accident prevention tool.

With the report's visualization analytics technology, airline safety professionals monitor multiple phases of thousands of flights daily. Data analysis isolates "atypicalities" that might not be outright safety problems, but that could help aviation experts identify potential problems before a significant safety concern emerges.

The technology also can track and analyze previously unnoticed trends and identify individual abnormalities that, when viewed in combination with other flight elements, might indicate a pattern of concern.

CYBER SECURITY

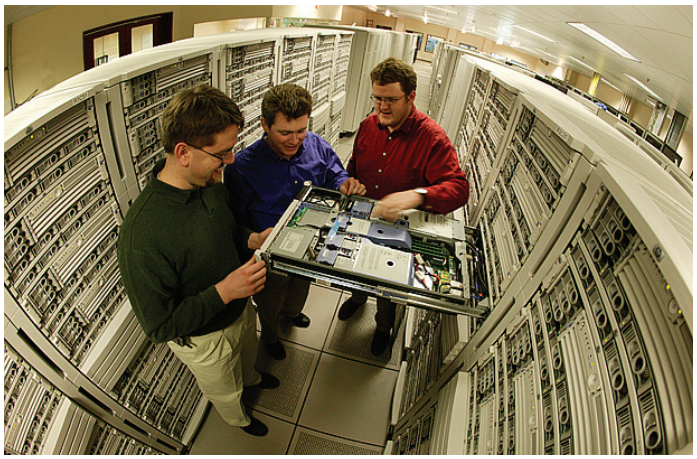
Web sites are an excellent means of presenting information to the public. However, these portals also create a potential weakness in that the size and breadth of the information posted to sites can be difficult to manage and can reveal information the organization does not want made public. Mozart™ is a PNNL software utility that scans websites to assess content clusters or to review for inadvertent release of sensitive information, such as trade secrets or sensitive, proprietary or classified information.

MOVING FORWARD

PNNL also plays a key role in building strategic relationships in the information technology community. The laboratory is working to improve communication and collaboration between government, which funds information technology research, and industry, which often has use for these technologies once they become available for real-world use.

McQuerry said PNNL researchers are only beginning to explore the possibilities of information technologies. "There are so many potential applications," he said. "We're researching how to apply information technology tools to other languages, to translate information for future analyses—the possibilities are endless."

JUDITH GRAYBEAL IS WITH PNNL COMMUNICATIONS.



PNNL's Linux-based supercomputer is composed of nearly 2,000 processors.