Editor: Theresa-Marie Rhyne

A Visual Analytics Agenda

James J. Thomas and Kristin A. Cook

Pacific Northwest National Laboratory

Attacks on the World Trade Center and the Pentagon on September 11, 2001 gave the US a stunning wake-up call. In August 2005, Hurricane Katrina provided another reminder that unprecedented disasters can and do occur. When they do, people must be able to analyze large amounts of disparate data to make sense of complex situations and save lives.

This need to support penetrating analysis of massive data collections isn't limited to security, however. It exists in many fields—from systems biology to human health; from evaluations of product effectiveness to strategizing for competitive positioning to assessing marketing campaign results. Simply put, the ability to collect data far outstrips the ability to analyze the collected data.

Following the September 11 attacks, the US government initiated efforts to evaluate existing technologies and ongoing near-term research. Two National Academy of Sciences reports identified serious gaps in the technologies. *Making the Nation Safer*[1] describes how advances in science and technology can help protect the nation against terrorism. *Information Technology for Counterterrorism*[2] expands on this work, focusing specifically on information technology's opportunities to help counter and respond to terrorist attacks.

Researchers have made significant progress in disciplines such as scientific and information visualization, statistically based exploratory and confirmatory analysis, data and knowledge representations, and perceptual and cognitive sciences. However, the research community hasn't adequately addressed the integration of these subspecialties to advance analysts' ability to apply their expert human judgment to complex data in pressure-filled situations. Although some research is being done in this area, the pace at which new technologies and technical talents are becoming available is far too slow to meet the urgent need.

National Visualization and Analytics Center

The Department of Homeland Security (DHS) established the National Visualization and Analytics Center (NVAC) with the mission of stimulating next-generation technologies and talents to reduce the risk of terrorism. These core technologies will also have much broader applicability, with the potential to bring value to almost any situation in which information overload and complexity hamper effective rapid analysis and communication of results.

NVAC's goal is to advance the state of the science to enable analysts to detect the expected and discover the unexpected from massive and dynamic information streams and databases consisting of data of multiple types and from multiple sources, even though the data are often conflicting and incomplete. In addition, all research and development (R&D) must fully support the protection of personal privacy and maintenance of data security.

NVAC assembled a crosscutting panel of about 40 research leaders in related fields from academia, industry, government agencies, and government laboratories to establish an R&D agenda to meet these needs. (The panel members, along with additional contributors to the agenda, are listed in the Acknowledgments.) The panel developed an understanding of the core analysis challenges. Panel members received training on analysis, heard from command center operators about lessons learned from the stress and confusion of past emergency situations, and learned about the challenges of monitoring US borders to detect wrongful entry of goods and people into the country.

Nobody on the panel questioned the need for new technology to meet these challenges. Building on previous National Academy of Sciences reports,[1,2] the panel defined the most urgently needed R&D. The report, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*,[3] presents the panel's results.

The science of visual analytics

The panel defined visual analytics as the science of analytical reasoning facilitated by interactive visual interfaces. People use visual analytics tools and techniques to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data; detect the expected and discover the unexpected; provide timely, defensible, and understandable assessments; and communicate assessment effectively for action.

Visual analytics is a multidisciplinary field that includes the following focus areas:

- analytical reasoning techniques that let users obtain deep insights that directly support assessment, plan-

 Published by the IEEE Computer Society 0272-1716/06/$20.00 © 2006 IEEE

ning, and decision making;
- visual representations and interaction techniques that exploit the human eye's broad bandwidth pathway into the mind to let users see, explore, and understand large amounts of information simultaneously;
- data representations and transformations that convert all types of conflicting and dynamic data in ways that support visualization and analysis; and
- techniques to support production, presentation, and dissemination of analytical results to communicate information in the appropriate context to a variety of audiences.[3]

The R&D agenda for visual analytics addresses technical needs for each of these focus areas, as well as recommendations for speeding the movement of promising technologies into practice. This article provides only a concise summary of the R&D agenda. We encourage reading, discussion, and debate as well as active innovation toward the agenda for visual analytics.

Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces.

The science of analytical reasoning

The science of analytical reasoning provides the reasoning framework upon which visual analytics technologies are built. The analytical reasoning process is central to the analyst's task of applying human judgment to reach conclusions from a combination of evidence and assumptions. Through the analytical reasoning process, users obtain insights that directly support situation assessment, planning, and decision making. The more complex the analysis, the more likely it will also require collaborative effort.

The analytical process involves diverse tasks such as

- understanding historical and current situations, as well as the trends and events leading to current conditions;
- identifying possible alternative future scenarios and the signs that one or another of these scenarios is coming to pass;
- monitoring current events to identify both expected and unexpected events;
- determining indicators of the intent of an action or an individual; and
- supporting the decision maker in times of crisis.

Especially in emergencies, analysts often conduct these tasks under extreme time pressure.

Visual analytics strives to facilitate analytical reasoning by creating software that maximizes human capacity to perceive, understand, and reason about complex and dynamic data and situations. It must build upon an understanding of the reasoning process, as well as an understanding of underlying cognitive and perceptual principles, to provide task-appropriate interactions that let users have a true discourse with their information. The goal is to facilitate high-quality human judgment with a limited investment of the analysts' time.

Visual analytics must enable analytical techniques that permit the creation of hypotheses and scenarios, and it must support the user in examining these hypotheses and scenarios in light of available evidence. Visual analytics must help the analyst discover unexpected and missing relationships that might lead to important insights. Visual analytics must also let users capture their progression through the analytical process and make this information available to others as a basis for future understanding.

To advance the science of analytical reasoning for visual analytics, the R&D community must build on theoretical foundations of reasoning, sensemaking, cognition, and perception, and combine these theoretical foundations with practical understanding of user tasks and processes to support collaborative analytic reasoning. The community must also address the challenges and seize the opportunities posed by the analytic problem's scale. Scale issues are manifested in the analytical task's complexity and urgency, the massive volume of diverse and dynamic data involved in the analysis, and the collaboration challenges.

Visual representations and interaction technologies

Visual representations and interaction technologies give users a gateway into their data, letting them see and understand large volumes of information at once. To facilitate analytical reasoning, visual analytics builds on the human mind's ability to understand complex information visually.

The R&D community needs to develop scientific principles for depicting information, identifying new interaction techniques, and mapping interaction approaches to analytical tasks. Visual representation and interaction techniques must be adaptable to tasks of varying complexity and urgency. These techniques must fully consider the impact of analysis at massive data scales. These scientific foundations provide the basis for new visual paradigms that can scale to support analytical reasoning in many situations.

Creating effective visual representations is a labor-intensive process requiring a solid understanding of the visualization pipeline, the characteristics of the data to be displayed, and the analytical tasks at hand. Most visualization software is developed with incomplete information about the data and tasks. New methods are needed for constructing visually based systems that simplify the development process and result in better-targeted applications.

To advance the science of visual representations and interactions for visual analytics, the R&D community must build on cognitive and perceptual design techniques and its wealth of visualization expertise to develop a new suite of visual paradigms that facilitate analytical reasoning. These visualizations must

- facilitate understanding of massive and continually

growing collections of data of multiple types;
- provide frameworks for analyzing spatial and temporal data;
- support the understanding of uncertain, incomplete, and often misleading information;
- provide user- and task-adaptable guided representations that enable full situation awareness while supporting development of detailed actions; and
- support multiple levels of data and information abstraction, including integration of different types of information into a single representation.

The R&D community must also develop a science of interactions to support analytical reasoning. This interaction science must provide a taxonomy of interaction techniques ranging from low-level to complex, and must address the challenge of scaling across different types of display environments and tasks. Interaction theory is less mature than visual design theory, so development of interaction theory is a priority.

In addition, the community must build upon the expanded sciences of visual representations and interaction techniques to create new methods for rapidly assembling and deploying visual analytics systems in various environments.

Data representations and transformations

Visualization is an essential aid to the analytical reasoning process. A visualization's quality is directly affected by the quality of the data representation underlying the visualization.

Visual analytics tools must transform data into representations that are suited to the analytical task and appropriately capture the important content of a large, complex, and dynamic collection. This R&D agenda defines a data transformation as a computational procedure that converts between data representations. Data transformations can augment data by deriving additional data. Some data transformations convert data into new, semantically meaningful forms. For example, we can use linguistic analysis to assign meaning to the words in a text document. Other data transformations can determine how to optimally display data, such as by creating a 2D or 3D representation of data with hundreds or thousands of dimensions.

Transforming and representing data are complex for many reasons. One issue is the sheer number of data types requiring analysis: short or long textual documents in many languages; numeric sensor data; structured data from relational databases; and audio, video, and image data. Each data type may require different transformations to facilitate visual analysis.

The data's massive scale and dynamic nature dictate that transformations must be fast, flexible, and capable of operating at many levels of abstraction. Data are of varying levels of certainty and reliability, and transformations must adequately reflect these assessments. An analysis often requires data of different types, so it is urgent to develop a data synthesis capability—a capability to bring different data types together in a single environment so analysts can concentrate on the data's meaning rather than on the form in which it was originally packaged.

To advance the science of data representations and transformations for visual analytics, the R&D community must

- develop both theory and practice for transforming data into new scalable representations that faithfully represent the underlying data's relevant content;
- create methods to synthesize different types of information from different sources into a unified data representation so users can focus on the data's meaning in the context of other relevant data, regardless of data type; and
- develop methods and principles for representing data quality, reliability, and certainty measures throughout the data transformation and analysis process.

A visualization's quality is directly affected by the quality of the data representation underlying the visualization.

Production, presentation, and dissemination

Production is the creation of materials that summarize the results of an analytical effort. *Presentation* is the packaging of those materials in a way that helps the audience understand the analytical results in context using terms that are meaningful to them. *Dissemination* is the process of sharing that information with the intended audience.

Production, presentation, and dissemination are often the most time-consuming parts of analysis and the only portions of the analytical process that are visible to decision makers. Because attention to this phase of analysis has been limited to date, visual analytics research has the potential to make great contributions. We need to substantially reduce the time required to prepare analysis results and share them with their audiences while dramatically improving the communication's effectiveness.

The immediacy of the demand for analytic results will spur scientific advancements that incorporate expertise from the visualization, graphic design, and rhetoric fields. To advance the capabilities for production, presentation, and dissemination of analytic results, the R&D community must develop a methodology and tools that let users capture the analytic assessment, decision recommendations, and resulting actions into information packages that are tailored for each intended receiver and situation and permit expansion to show supporting information as needed. The community must also develop technologies that enable users to communicate what they know through the use of appropriate visual metaphor and accepted principles of reasoning and graphic representation. In addition, the community must create techniques that enable effective use of limited, mobile forms of technologies and support the seamless integration of tools so that data requests and acquisition, visual analysis, note taking, presentation composition, and dissemination all take place in a cohesive environment that supports around-the-clock operation and provides robust privacy and security control.

Moving research into practice

To have a lasting benefit to the analytical community, successful research results must be moved into practical use. The issues associated with moving research into practice are often omitted from R&D agendas, but the NVAC panel felt compelled to provide a framework for addressing four fundamental issues for accelerating the adoption of new technologies. Each of these issues has the potential to make or break the successful deployment of visual analytics technologies.

- New tools, algorithms, and techniques must be evaluated to ensure that they represent a significant advance over current practice and that they operate correctly.
- The technical community must address security and privacy issues from the start and throughout the research, development, and deployment process.
- The technical community must attend to software interoperability, architecture, and data handling to facilitate collaborative research, software evaluation, and software deployment into a wide variety of software environments.
- A concerted and sustained effort to insert the resulting technology into operational environments will be essential if the research results are to be of benefit.

To accelerate the transition of research into analytical practice, the R&D community must:

- develop an infrastructure to facilitate evaluation of new visual analytics technologies;
- create and use a common security and privacy infrastructure, incorporating privacy-supporting technologies such as data minimization and data anonymization;
- use a component-based software development approach for visual analytics software to facilitate evaluation of research results in integrated prototypes and deployment of promising components in diverse operational environments; and
- identify and publicize best practices for inserting visual analytics technologies into operational environments.

Positioning for enduring success

Achieving the agenda outlined here will require the sustained efforts of a multidisciplinary community of researchers. Establishing and sustaining an enduring visual analytics R&D community capable of meeting these challenges require the following educational efforts and partnerships:

- programs to support the research community's education about the drivers for visual analytics research;
- university-led centers of excellence and partnerships with government, industry, national laboratories, and selected international research entities to bring together the best talents to accomplish the visual analytics R&D agenda;
- special partnerships with users, and the IT professionals who support them, to facilitate adoption of new technologies into operational environments; and
- ongoing support for collaborations, internships, staff exchanges, and educational efforts to help build an understanding of user needs.

Call to action

NVAC's R&D agenda is a call to action on many fronts, and many are responding. For example, this R&D agenda is becoming a basis for DHS and other government agencies to establish programs addressing the need for new technologies, and many in academia and industry are directing their interests and resources toward the agenda. The international research community is also joining in the advancement of these technologies.

We must respond now with creative ideas and teams of interdisciplinary talents to move the visual analytics field forward rapidly. The breakthroughs will not only improve safety and security, but will support advancements in the sciences as well. ■

Acknowledgments

We thank the Department of Homeland Security for their sponsorship of this effort, particularly Joseph Kielman, Maureen McCarthy, Alexandra Landsberg, and David Shepherd.

We also acknowledge the following people, who helped create this agenda: Anthony Bartoletti, Mark Billinghurst, Stuart Card, Dan Carr, John Dill, Rae Earnshaw, Dave Ebert, Stephen Eick, Brian Fisher, Robert Grossman, Chuck Hansen, Beth Hetzler, Frank Hughes, Don Jones, Ken Joy, David Kasik, David Laidlaw, Sharon Laskowski, Alan MacEachren, David Moore, Lucy Nowell, Donna Peuquet, Catherine Plaisant, Bill Ribarsky, John Stasko, Maureen Stone, Alan Turner, Matt Ward, David White, Mark Whiting, Paul Whitney, Pak Chung Wong, David Woods, and Bill Wright.

Finally, we express special appreciation to panel members George Robertson, Pat Hanrahan, Nancy Chinchor, and Russ Rose for their extra efforts to refine the panel recommendations and supporting text.

References

1. B. Alberts and W.A. Wulf, *Making the Nation Safer: The Role of Science and Technology in Countering Terrorism*, Nat'l Academies Press, 2002; http://www.nap.edu/openbook/0309084814/html.
2. J.L. Hennessy, J.L. Patterson, and H.S. Lin, eds., *Information Technology for Counterterrorism: Immediate Actions and Future Possibilities*, Nat'l Academies Press, 2003; http://www.nap.edu/books/0309087368/html.
3. J.J. Thomas and K.A. Cook, eds., *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE CS Press, 2005; http://nvac.pnl.gov/agenda.stm.

Contact the authors at [kris.cook, jim.thomas]@pnl.gov. Contact editor Theresa-Marie Rhyne at tmrhyne@ncsu.edu.