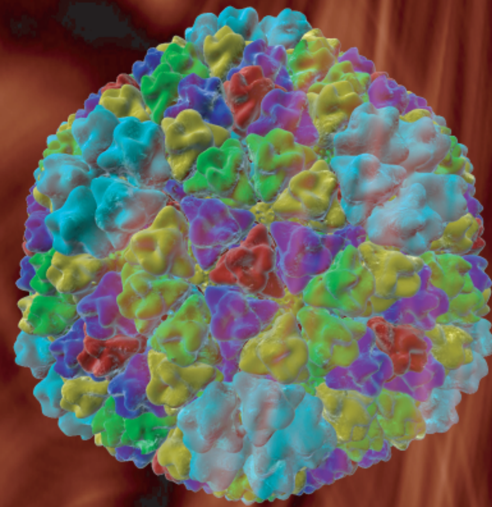
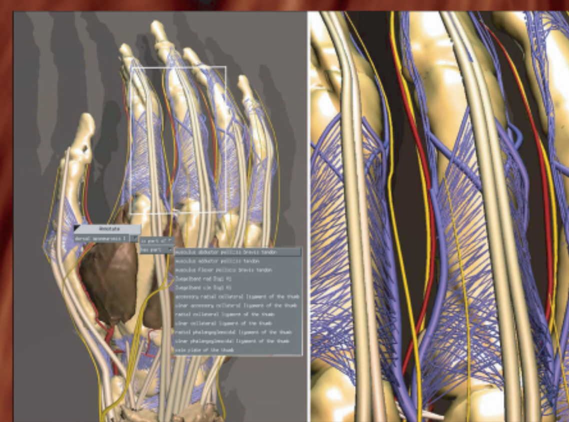
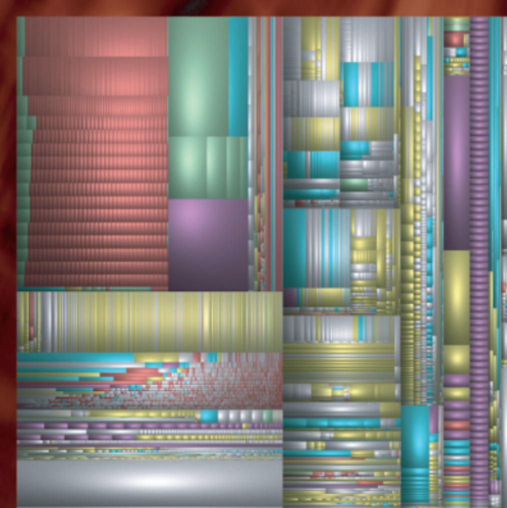


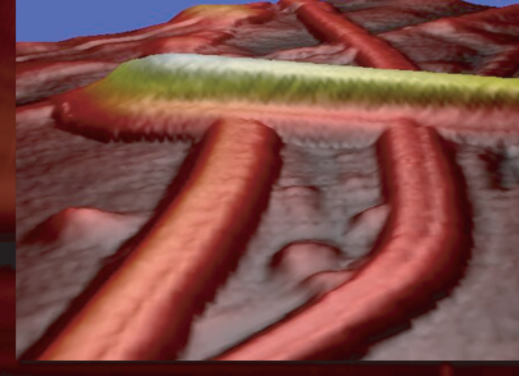




NIH/NSF

VISUALIZATION RESEARCH CHALLENGES

JANUARY 2006



NIH/NSF Visualization Research Challenges

January 2006

Chris Johnson
Robert Moorhead
Tamara Munzner
Hanspeter Pfister
Penny Rheingans
Terry S. Yoo

Trademarks

Published 2006
Printed by IEEE in the United States of America
ISBN: 0-7695-2733-7

Additional Copies of This Book

Additional copies may be ordered from the IEEE Computer Society:

IEEE Computer Society Order Number: R0235

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
Los Alamitos, California 90720 USA
Tel (toll-free): +1 800 CS BOOKS
Tel (direct): +1 714 821 8380
Fax: +1 714 821 4641
E-mail: csbooks@computer.org
CS Store: http://www.computer.org/portal/pages/store/b2005/r0235.xml

NIH/NSF Visualization Research Challenges Report

January 2006

Chris Johnson (Scientific Computing and Imaging Institute, University of Utah)
Robert Moorhead (Mississippi State University)
Tamara Munzner (University of British Columbia)
Hanspeter Pfister (Mitsubishi Electric Research Laboratories)
Penny Rheingans (University of Maryland Baltimore County)
Terry S. Yoo (National Library of Medicine)

Acknowledgments:

We are indebted to the expert and visionary panelists for sharing their considerable talents and insights, and to NIH and NSF for their sponsorship. Special thanks to Nathan Galli for the report design and Katharine Coles for her editing assistance.

This project was sponsored by NSF CISE, NSF Engineering, NIH NIBIB, and NIH NLM and received additional support from the Scientific Computing and Imaging Institute of the University of Utah, the IEEE Visualization and Graphics Technical Committee, Mitsubishi Electric Research Labs, and IEEE Press.

Workshop I Panelists:

Maneesh Agrawala, Microsoft Research; **Steve Cutchin**, San Diego Supercomputing Center; **David Ebert**, Purdue University; **Thomas Ertl**, University of Stuttgart; **Bob Galloway**, Vanderbilt University; **Mike Halle**, Harvard University, BWH SPL; **Chuck Hansen**, University of Utah; **Karl Heinz Hoehne**, University of Hamburg; **Ken Joy**, University of California Davis; **Arie Kaufman**, Stony Brook University (SUNY); **David Laidlaw**, Brown University; **Ming Lin**, University of North Carolina-Chapel Hill; **Bill Lorensen**, General Electric Global Research; **Alan MacEachren**, Pennsylvania State University; **Chris North**, Virginia Polytechnic Institute; **Art Olson**, Scripps Research Institute; **Catherine Plaisant**, University of Maryland College Park; **Jerry Prince**, Johns Hopkins University; **Will Schroeder**, Kitware, Inc.; **Jack Snoeyink**, University of North Carolina-Chapel Hill; **John Stasko**, Georgia Institute of Technology; **Barbara Tversky**, Stanford University; **Colin Ware**, University of New Hampshire.

Workshop I Guest Speakers:

Felice Frankel, Massachusetts Institute of Technology; **Alyssa Goodman**, Harvard University Astrophysics; **Leslie Loew**, University of Connecticut; **Wayne Loschen**, Johns Hopkins Applied Physics Laboratory; **Patrick Lynch**, Yale University School of Medicine.

Workshop II Panelists:

Liz Bullitt, University of North Carolina-Chapel Hill; **David Ebert**, Purdue University; **Thomas Ertl**, University of Stuttgart; **Steve Feiner**, Columbia University; **Michael Halle**, Harvard University, BWH SPL; **Pat Hanrahan**, Stanford University; **Chuck Hansen**, University of Utah; **Helwig Hauser**, VRVis Research Center; **Ken Joy**, University of California Davis; **Daniel Keim**, University of Konstanz; **Bill Lorensen**, General Electric Global Research; **Kwan Liu Ma**, University of California Davis; **Alan MacEachren**, Pennsylvania State University; **Will Schroeder**, Kitware Inc.; **Matthew O. Ward**, Worcester Polytechnic Institute; **Ross Whitaker,** University of Utah; **Turner Whitted**, Microsoft Research; **Jarke van Wijk**, University of Eindhoven.

Contents

EXECUTIVE SUMMARY 1

Nearly twenty years ago, the U.S. National Science Foundation (NSF) convened a panel to report on the potential of visualization as a new technology. Last year, the NSF and U.S. National Institutes of Health (NIH) convened the Visualization Research Challenges Executive Committee to write a new report. The goal of this new report is to evaluate the progress of the maturing field of visualization, to help focus and direct future research projects, and to provide guidance on how to apportion national resources as research challenges change rapidly in the fast-paced world of information technology. We describe some of the remarkable achievements visualization enables and discuss the major obstacles blocking the discipline's advancement.

Our findings and recommendations reflect not only information gathered from visualization and applications scientists during two workshops on Visualization Research Challenges but also input from the larger visualization community. We are indebted to the expert and visionary panelists for sharing their considerable talents and insights, and to the NSF and NIH for their sponsorship.

Advances in the science and technology of computing have engendered unprecedented improvements in scientific, biomedical, and engineering research, defense and national security, and industrial innovation. Continuing and accelerating these advancements will require people to comprehend vast amounts of data and information being produced from a multitude of sources. Visualization, namely helping people explore or explain data through software systems that provide a static or interactive visual representation, will be critical in achieving this goal. Visualization designers exploit the high-bandwidth channel of human visual perception to allow people to comprehend information orders of magnitude more quickly than they could through reading raw numbers or text.

Visualization is fundamental to understanding models of complex phenomena, such as multilevel models of human physiology from DNA to whole organs, multi-century climate shifts, international financial markets, or multidimensional simulations of airflow past a jet wing. Visualization reduces and refines data streams rapidly and economically, thus enabling us to winnow huge volumes of data in applications such as the surveillance of public health at a regional or national level in order to track the spread of infectious diseases. Visualizations of such application problems as hurricane dynamics and biomedical imaging are generating new knowledge that crosses traditional disciplinary boundaries. Visualization can provide industry with a competitive edge by transforming business and engineering practices.

Although well-designed visualizations have the power to help people enormously, naive attempts to create visualizations typically lead to "reinventing the wheel" at best, and all too often result in poorly designed visualizations that are ineffective or even actively misleading. Designing effective visualizations is a complex process that requires a sophisticated understanding of human information processing capabilities, both visual and cognitive, and a solid grounding in the considerable body of work that has already been introduced in the visualization field. Further research in visualization, and the technology transfer of effective visualization methodologies into the working practice of medicine, science, engineering, and business, will be critical in handling the ongoing information explosion. The insights provided by visualization will help specialists discover or create new theories, techniques, and methods, and improve the daily lives of the general public.

While visualization is itself a discipline, advances in visualization lead inevitably to advances in other disciplines. Just as knowledge of mathematics and statistics has become indispensable in subjects as diverse as the traditional sciences, economics, security, medicine, sociology, and public policy, so too is visualization becoming indispensable in enabling researchers in other fields to achieve their goals. Like statistics, visualization is concerned with the analysis and interpretation of information, both quantitative and qualitative, and with the presentation of data in a way which conveys their salient features most clearly. Both fields develop, understand, and abstract data analytic ideas and package them in the form of techniques, algorithms, and software for a multitude of application areas.

However, despite the importance of visualization to discovery, security, and competitiveness, support for research and development in this critical, multidisciplinary field has been inadequate. Unless we recommit ourselves to substantial support for visualization research, development, and technology transfer, we will see a decline in the progress of discovery in other important disciplines dependent on visualization. As these disciplines lose their ability to harness and make sense of information, the rate of discovery itself will decline. In the inevitable chain reaction, we will lose our competitive edge in business and industry.

Principal Finding: Visualization is indispensable to the solution of complex problems in every sector, from traditional medical, science and engineering domains to such key areas as financial markets, national security, and public health. Advances in visualization enable researchers to analyze and understand unprecedented amounts of experimental, simulated, and observational data and through this understanding to address problems previously deemed intractable or beyond imagination. Yet, despite the great opportunities created by and needs fulfilled by visualization, NSF and NIH (and other Federal government agencies) have not effectively recognized the strategic significance and importance of visualization in either their organizational structures or their research and educational planning. The recent and cogent book outlining the visual analytics research agenda, which dovetails closely with visualization, offers promise that visualization for the application area of national security will be well-funded in the near future. However, that effort encompasses only one sector and is short-term, whereas the field needs long-term support across many sectors. The current distribution of funding sources does not reflect the potential benefits of visualization research to specific application areas. These inadequacies compromise the future of U.S. scientific leadership, public health, and economic competitiveness.

Principal Agency Leadership Recommendation: NSF and NIH must make coordinated investments in visualization to address the 21st century's most important problems, which are predominantly collaborative, crossing disciplines, agencies, and sectors. Both NSF and NIH can and should provide leadership to other Federal funding partners, as well as to the research communities they support, by modifying their programmatic practices to better engage visualization capabilities across disciplines important to scientific and social progress and to encourage and reward interdisciplinary research, open practices, and reproducibility in technical and scientific developments. Such agency leadership is critical if we are to meet the needs of our nation in critical areas and to maintain U.S. competitiveness in a global environment.

Short-Term Policy Recommendation: Policy changes for both funding and publication review to encourage evaluation of visualization and collaboration between visualization and other fields can be implemented immediately, without requiring new funding initiatives. Within visualization, review protocols should reflect the importance of evaluation to determine the success and characterize the suitability of techniques. Methodological rigor should be expected when user studies are proposed or reviewed, and as necessary visualization researchers should collaborate with those trained in fields such as human-computer interaction, psychology, and statistics. Peer review of proposals and publications should also reward visualization driven by real-world data and tasks, in close collaboration with target users. In other fields, review protocols should encourage domain scientists who create partnerships with visualization researchers, just as engagement with statisticians is considered normal practice in many areas such as biomedical research.

Mid-Term Direction Recommendation: Pilot programs should be established to combine efforts and create collaborative development between visualization and other research domains. Funding for such programs should contribute proportionately to both the visualization research and the domain specialty. The purpose of this effort will be to improve the penetration of emerging technologies into new domains, increasing their facility to move data and share results through visualization. All of the awards in this area should be dedicated to open access of source code, availability of research data to the worldwide community, and reproducibility of the technical and scientific developments.

Long-term Investment Recommendation: We recommend a coordinated and sustained national investment be made in a spectrum of centralized and distributed research programs to promote foundational, transitional, and applied visualization research in support of science, medicine, business, and other socially important concerns. This investment is critical for the U.S. to remain competitive in a global research and development community that has increasing resources. In addition to funding transitional research, such programs should emphasize foundational research and integration of methodologies from other fields, and collaboration with domain specialists who provide driving problems in areas of national concern. A long-term funding commitment is required for the creation and maintenance of curated data collections, and open-source software, to promote open science. Characterizing how and why visualizations work, systematically exploring the design space of visual representations, developing new interaction approaches, and exploiting the possibilities of novel display hardware will be particularly important areas of emphasis.

THE VALUE OF VISUALIZATION 2

Visualization is a growing part of everyday life in every sector of society, from family life, to research and education, and even to homeland security and public health. We routinely depend on it every time a meteorologist puts a weather map into motion tracking the progress of hurricanes and alerting us to life-threatening weather conditions. Essential discoveries such as the structure of the molecules that control our lives and our genetic code are shared through images, physical models, and interactive displays that are generated today using visualization technologies. Such information is shared across the internet by students and scholars. Visualization plays a role in saving lives, accelerating discovery, and promoting education through improved understanding.

As a discipline, visualization focuses on helping people explore or explain data through software systems that provide static or interactive visual representations. Visualization designers exploit the high bandwidth channel of human visual perception to allow people to comprehend information orders of magnitude more quickly than they could through reading raw numbers or text. Visual representations of information have a rich and lengthy historical tradition stretching back to cave paintings and beyond, but the recent advent of computer graphics has created the ability to represent increasingly larger datasets, and has simultaneously introduced the ability for users to manipulate data interactively. Visualization is useful for detecting patterns, assessing situations, and prioritizing tasks[21]. Understanding, and, ultimately, knowledge cannot be delivered directly from computation. Visualization is the tool through which computation addresses an end user and allows the user to derive knowledge from data.

People are biologically equipped to make spatial inferences and decisions, and experience refines their ability to do so. Visualizations can bootstrap this facility metaphorically, by mapping elements and spatial relations in the abstract domain onto elements and relations in a concrete visualization. Through such maps, the human ability to make spatial inferences can be transferred to abstract domains. However, human information processing capabilities, both visual and cognitive, are limited and systematically biased. Effective visualizations must take these facts into account, selecting and highlighting essential information, eliminating distracting clutter, and conveying ideas that are not inherently visual, such as transformations or causality, through visual channels. Although there are tools and methods for designing effective visualizations, too many naive designers fail to use them, and their failure results in poor visualizations.

While many areas of computer science aim to replace human judgment with automation, visualization systems are explicitly designed not to replace the human but to keep the human in the loop by extending human capabilities. Figure 2.1 shows a simplified view of the discovery process in which raw data is transformed into knowledge. In reality, the user is an active participant, interaction is common and flexible, and the process of exploration using visual display and interaction happens in many different ways throughout a complex process[16]. Nevertheless, this simplified view captures the main areas where visualization research needs to focus: perception/cognition, exploration/interaction, and specification/visualization.

The invention of abstractions, models, and mechanisms to explain the world around us is an inherently human endeavor. Ultimately, the practice of visualization should assist in the generation, evaluation, and exploration of hypotheses about the information under study, allowing the rapid consideration and possible rejection of old hypotheses and facilitating the creation of new hypotheses. Visualization leverages a combination of imagination, computer tools, and interactive interfaces to extend the power of human insight to aid in the discovery and synthesis of truth.

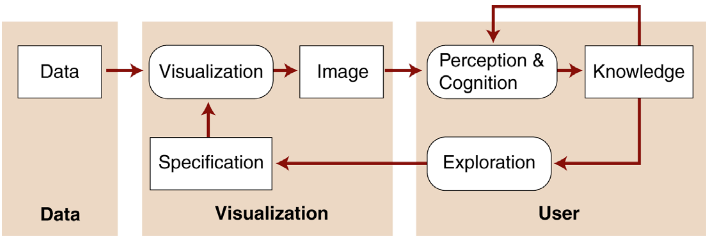


Figure 2.1: *The visualization discovery process. Data encompasses the range form a single bit, to time-varying 3D tensor fields, to multi-modal data sources requiring alignment, registration, and fusion, and to non-spatial information sources integrating broad areas of human knowledge. The visualization specification includes the hardware, the algorithms, and the specific parameters. Users adjust the specification of the visualization, requiring the introduction of interactive controls. The resulting image will often be an image in the usual sense, but it can also be an animation, or auditory or haptic feedback. Adapted from van Wijk*[50].

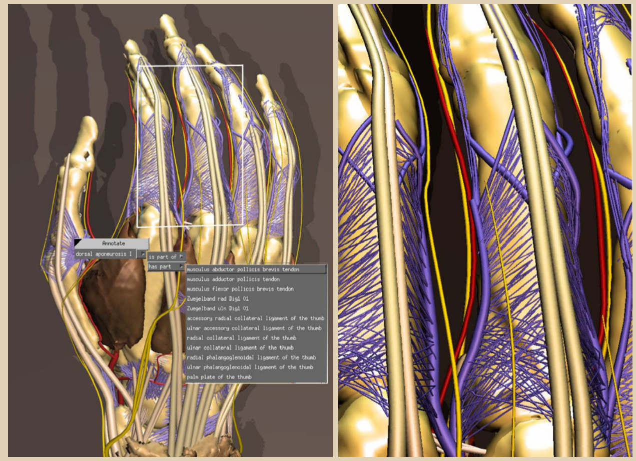
Just as knowledge of mathematics and statistics has become indispensable in subjects as diverse as the traditional sciences, economics, security, medicine, sociology, and public policy, so too is visualization becoming indispensable in enabling researchers in a vast range of fields achieve their goals. Like statistics, visualization is concerned with the analysis and interpretation of information, both quantitative and qualitative, and with the presentation of data in a way that conveys their salient features most clearly. Both fields develop, understand, and abstract data analytic ideas and package them in the form of techniques, algorithms, and software for a multitude of application areas.

Nearly twenty years ago, the NSF convened a panel to report on the potential of visualization as a new technology[28]. During the 17 years since the last NSF Visualization Report[28], the world has experienced an “information big bang,” an exponential explosion of data. New data produced in the two years since 2003 exceeds the information contained in all previously created documents. Of all this new data produced since 2003, more than 90% takes digital form, vastly exceeding information produced in paper and film forms[26]. This growth in data does not necessarily mean a corresponding proportional increase in useful information. Raw data is, in and of itself, of questionable value. We are continually challenged to make sense of the enormous growth and onslaught of information and use it in effective and efficient ways. The 1971 observations of Nobel Prize winning economist, Herbert Simon, are more true now than ever:

> What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.[44]

Among the greatest scientific challenges of the 21st century, then, is to effectively understand and make use of the vast amount of information being produced. Our primary problem is no longer acquiring sufficient information, but rather making use of it. Imagine that all relevant and irrelevant facts were available; suppose we knew everything. What would we do with this overwhelming resource? If we are to use it to make discoveries in science, engineering, medicine, art, and the humanities, we must create new theories, techniques, and methods for its management and analysis. By its very nature, visualization addresses the challenges created by such excess – too many data points, too many variables, too many timesteps, and too many potential explanations. Visualization harnesses the human perceptual and cognitive systems to tackle this abundance. Thus, as we work to tame the accelerating information explosion and employ it to advance scientific, biomedical, and engineering research, defense and national security, and industrial innovation, visualization will be among our most important tools.

Teaching Anatomy and Surgery



Detailed anatomic models of delicate organs such as the human hand are a prerequisite for both teaching the complex anatomy and the preoperative simulation of interventions. While classical anatomy atlases can provide sufficient anatomical detail in a set of static images, they do not allow choosing user defined views or performing surgical interaction. The above picture (with a magnification of the area bound by the rectangle in the left image) illustrates a novel computer-based anatomy model (“VOXEL-MAN”) that not only allows arbitrary viewing and dissection, but also the interrogation of the anatomic constituents by mouse click. The pictorial model was created from the Visible Human data set using volume visualization (bone, muscles) and surface modeling (blood vessels, nerves, ligaments, tendons). The pictorial model is linked to a knowledge base, describing the anatomic constituents and their relations. With its new features, it offers possibilities to both students and expert surgeons which are indispensable to cope with the complexity of state-of the-art microsurgical interventions. One of the challenges is to extend the knowledge base such that the system can warn the user of consequences of a surgical interaction for the patient.

K.H. Höhne, B. Pflesser, A. Pommert, M. Riemer, R. Schubert, T. Schiemann, U. Tiede, and U. Schumacher, A realistic model of human structure from the Visible Human data. *Meth. Inform. Med.*; Vol. 40 No. 2, pp. 83-89, 2001.

THE PROCESS OF VISUALIZATION 3

Visualization research can live up to its full potential only if it addresses the fundamental challenges of the field. These challenges demand an approach that moves beyond incremental improvements, incorporates evaluation of success as an integral part of research, freely shares research results and products, and includes research types spanning the range from foundational to applied.

3.1 Moving Beyond Moore's Law

In 1965, Gordon Moore, who would later be one of Intel's founders, observed that the number of components on a chip had doubled for each of the last three years and predicted that this trend would continue for ten more years. Loosely interpreted, Moore's Law is now taken to mean that processing power will double every couple of years without impact on cost. The beauty of Moore's Law is that certain problems will solve themselves if we just wait. For example, much research effort has been devoted to developing special-purpose hardware, specialized optimizations of specific algorithms, methods for out-of-core computation, parallel and distributed implementations for complex computations, and detail-reduction methods for meshes. Such research can make visualization faster and more useful for real-scale problems, but it does not often lead to breakthroughs with genuinely new possibilities.

Many extremely important areas of visualization research tackle problems not governed by Moore's law. Advances in these areas can yield new capabilities, new visions, new applications, and a firmer theoretical basis for visualization research and practice. The following are examples of areas that emphasize aspects of visualization that involve the human in the loop.

Collaborating with Application Domains To achieve greater penetration of visualization into application domains we must better integrate visualization capabilities with the requirements and environments of these domains. To achieve this integration, we must allow application goals, domain knowledge, and domain-specific conventions and metaphors to shape visualization methods. Visualization methods must address the characteristics of real, rather than ideal, data, addressing among others the challenges of heterogeneity, change over time, error and uncertainty, very large scale, and data provenance.

Finding: Visualization researchers should collaborate closely with domain experts who have driving tasks in data-rich fields to produce tools and techniques that solve clear real-world needs.

Integrating with Other Methodologies Visualization is rarely a stand-alone process: visualization is often necessary but not sufficient for solving problems. Visualization tools and methods should provide tighter integration with other analytic tools and techniques, such as statistics, data mining, and image processing, in order to facilitate analysis from both qualitative and quantitative perspectives. The newly-coined term Visual Analytics[48] is a good example of an explicitly cross-disciplinary approach.

Finding: To extend visualization's utility, we must integrate visualization with other techniques and tools from other disciplines.

Examining Why and How Visualizations Work Human perceptual and cognitive capacities are largely fixed, not subject to Moore's Law. Even our understanding of these capacities grows slowly rather than doubling in a matter of years. Addressing the human element in visualization may require not simply making the system faster, but rather making the system *different* in order to better leverage human characteristics, strengths, and limitations. To this end, visualization research must actively seek to identify perceptual and cognitive influences on visualization effectiveness in order for visual displays to best augment human reasoning. Many current design principles of visualization are based on the century of work characterizing human psychophysical responses to low-level visual stimuli. We would benefit immensely from a more thorough understanding of higher level phenomena such as spatial memory and environmental cognition. We can furthermore distinguish between the noun *visualization*, which refers to a display showing visual information, and the verb to *visualize*, which refers to the process of how a human uses that display. We need to identify more accurately when, why, and how visualization provides insight to enable analytic thinking and decision making in a world of changing data sources, input and display devices, and user needs.

Finding: Investigation of the nature, options, limitations, and effects of human perception, cognition, and the visual exploration experience will accelerate progress in visualization and visual communication.

Exploring New Visualization Techniques Systematically The set of current visualization techniques is rich and powerful but far from complete, especially as complex data sources continue to expand and create new challenges in data synthesis and representation. We have much work still to do on the discovery and design of new representations for complex, multivariate, heterogeneous, multiscale, and dynamic data. Thus, we must systematically explore the design space of possible visual representations.

Finding: In order to benefit both current and new application domains, we must engage in the systematic exploration of the design space of possible visualization techniques.

Designing Interaction Research in new interaction techniques will allow users to interactively manipulate and explore data and extract meaning from it. Fluid interaction requires that we create user interfaces that are less visible to the user, create fewer disruptive distractions, and allow faster interaction without sacrificing robustness. In addition to developing novel interaction metaphors, future visualization interfaces will need to respond to rapid innovation in visual display technology that is resulting in a range of hardware devices quite different from the current standard, including high resolution and lightweight projectors, flat panel displays, and touch-sensitive display surfaces. Haptic and tactile devices for both input and output are becoming commercially available, along with embedded and wireless technologies that make computing power ubiquitous. The challenge will be to characterize the strengths and weaknesses of these new kinds of hardware when they are used to support visualization for both single-user and collaborative systems.

Finding: We must design appropriate interaction metaphors for both current and future hardware in order to harness the full power of visualization systems.

Whale Tracks

This image shows a visualization revealing the path of a humpback whale over the course of several hours. The data were acquired from a tag attached to the whale via suction cups that recorded depth, angular acceleration and magnetic north. These data were used to construct the pseudo-track ribbon. The saw tooth patterns represent angular accelerations due to fluke strokes. Twists in the ribbon reveal rolling behavior. The ribbon plot makes patterns of behavior much more clearly evident. For example it shows that this particular whale always swam up and glided down. Also, a particular foraging behavior, side-rolls, believe to be in pursuit of a small fish species called sand lance, was revealed to be ubiquitous and highly stereotyped. The ribbon plot is a key feature of an interactive 3D application *TrackPlot* that was developed to allow ethologists to better interpret the underwater behavior of individual whales. Previous to its development, researchers had either "played back" the motion of the whale or constructed depth-time graphs. Neither of these techniques was as effective in revealing complex 3D patterns. Future challenges include visualizing the interactions of multiple tagged whales and visualizing whale-prey interactions.

C. Ware, R.A. Arsenault, D. Wiley, and M. Plumlee, Visualizing the underwater behavior of humpback whales, (submitted for publication).

3.2 Determining Success

As with all computer disciplines, visualization occasionally makes ground-breaking and innovative advances that provide obvious advantages and orders of magnitudes of improvement over previous techniques. More often, however, we must quantify advances and measure improvement through benchmarks and carefully designed evaluation studies. Evaluation allows a researcher to answer the question "Did this technique actually help human users solve their targeted problems?" or "How much does this new approach improve the confidence or accuracy of human insight?" To effectively answer these questions, a visualization researcher must have an active connection with a domain researcher with a driving problem, providing context for the measured improvements. The very act of measuring the performance and value of a visualization helps to guide the field and help it grow.

The field of visualization has unique evaluation challenges. While we can quantitatively measure the time and memory performance of an algorithm, such metrics do not shed light on the ultimate measure: human insight gained by computation or visualization.

We do have some methods for determining whether or not a visualization tool has helped a person solve a problem. A quantitative user study performed in a formal laboratory setting can measure the performance of users on an abstracted task using metrics such as task completion times or error rates. The human-computer interaction and psychology communities teach sound study design and statistical analysis in order to ensure good methodologies and accurate results.

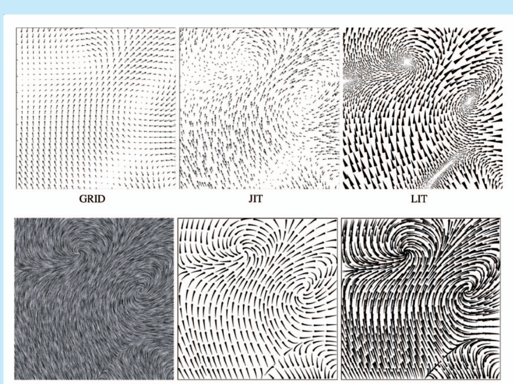
However, there are also many ways to qualitatively evaluate systems. Anecdotal evidence from satisfied real-world users that a visualization system is helpful can be useful in demonstrating that the system has succeeded in its design goal. These anecdotes include accounts of “eureka moments” in which something previously unknown was discovered. The size of the user community can also demonstrate a system’s usefulness, because voluntary adoption reflects a judgment from the users that a visualization tool is effective. Indeed, a powerful measure of success is provided when visualization tools become so pervasively deployed in an application domain that their use is considered unremarkable. Qualitative user studies, ranging from ethnographic analysis of target user work practices to longitudinal field studies to informal usability evaluation of a prototype system, also play an important role in both design and evaluation.

Finally, an analysis that relates design choices to a conceptual framework is a powerful evaluation method. Measuring the effectiveness of the design of a visualization requires the use of case studies, in which design choices are discussed and justified in the context of the theoretical foundations of the research field. The outgrowth of these studies is the ontological organization of visualization itself, organizing the very structure, utility, and expressiveness of visual tools along guidelines and design principles. The resulting frameworks help us move beyond simply asking *whether* something helps by offering tools to answer questions of *why* and *how* it helps. Several authors have presented such frameworks, including Bertin[6], Cleveland[11], Card and Mackinlay[8], Shneiderman[43], and Wilkinson[52].

Too often, visualization is considered the last step of a research project, in which the visualization specialist is engaged to present the results of an experiment already completed. However, visualization can help to frame questions, to guide an investigation, and to develop intuitions and insight about the problem under study. In order to foster these capabilities and empower the field of visualization as an equal partner with domain experts in the exploration of science and society, we need to encourage the formalization of visualization design and the rigorous development of evaluation metrics. When improved formalizations and quantitative performance metrics are established for visualization, the field will more effectively assist research in almost all areas of human endeavor.

Finding: In order to be most effective, visualization research must move toward completing the research cycle by examining and evaluating the effects of visualization techniques and approaches.

Characterizing Flow Visualization Methods



For decades researchers have been developing visualization techniques that advance the state of the art and are published in peer-reviewed journals. However, there are disproportionately few quantitative studies comparing visualization techniques, such as this characterization of the differences between flow visualization methods. The image shows six different methods for visualizing the same 2D vector field. Subjects who participated in the user study performed several tasks including identifying the type and location of critical points in visualizations. Assuming roughly equal importance for all tasks, the GSTR visualization performed best overall: on average, subjects were fastest and most accurate when using it. This study produced both quantitative results as well as a basis for comparing other visualization methods, for creating more effective methods, and for defining additional tasks to further understand tradeoffs among methods. A future challenge is to develop evaluation methods for more complex 3D time-varying flows.

D.H. Laidlaw, M. Kirby, C. Jackson, J. Davidson, T. Miller, M. DaSilva, W. Warren, and M. Tarr, Comparing 2D vector field visualization methods: A user study. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 11, No. 1, pp.59-70, January-February 2005.

3.3 Supporting Repositories and Open Standards

One of the basic requirements of science is that experiments be repeatable. For visualization, this requirement entails sharing data, models, and tasks to verify and benchmark new algorithms and techniques, comparing them to the results of previous work. Although scattered examples of shared data exist, such as the bluntfin test data from the Flow Analysis Software Toolkit (FAST) developed at NASA/Ames[5], and the data from the Visible Human project sponsored by NIH[3], there is great need for sustained and methodical creation of data, model, and task repositories.

Many of the arguments for open source software also hold for *open science*; that is, making the fruits of publicly funded science available to the community. Open data and task repositories are critical for continued progress in visualization[13]. However, the difficulty is that visualization practitioners are typically not themselves the primary source of the data. We must depend on the willingness of those who generate the data to share it. Thus, we can and must be advocates for data sharing whenever possible. The visualization community must consider this advocacy, and the curation of visualization-oriented data and task repositories, as part of our own contribution to open science.

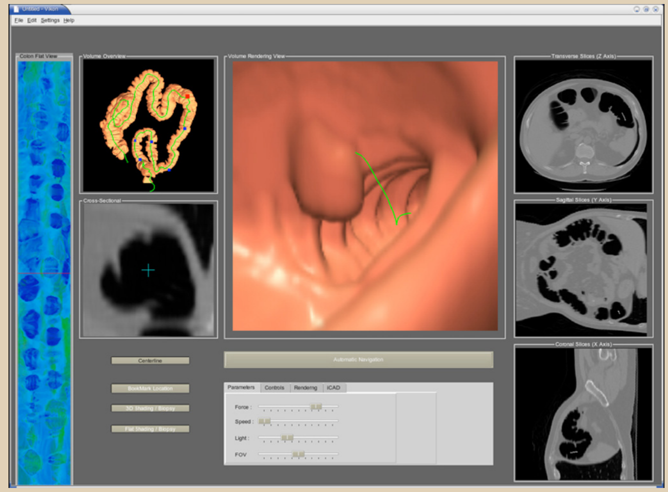
As noted in the PITAC Report[37], "The explosive growth in the number and resolution of sensors and scientific instruments has engendered unprecedented volumes of data, presenting historic opportunities for major scientific breakthroughs in the 21st century. Computational science now encompasses modeling and simulation using data from these and other sources, requiring data management, mining and interrogation." We agree with the PITAC Report that "The Federal government must provide long-term support for computational science community data repositories" and "The Government must require funded researchers to deposit their data and research software in these repositories or with access providers that respect any necessary or appropriate security and/or privacy requirements."

Finding: Since creating and maintaining open data and task repositories is critical for the health of the visualization field, we must support it through both policy and funding mechanisms.

3.4 Achieving Our Goals

Research, and therefore research funding, is often divided into three categories: basic, transitional, and applied. The progressive relationship among these divisions is often clear. However, visualization, as a field that creates techniques as much as it explores new phenomena, does not fit so neatly

Virtual Colonoscopy



This visualization shows the user interface for a *virtual colonoscopy* (VC) system. VC employs computed tomography (CT) scanning and volume visualization, and is poised to become the procedure of choice in lieu of the conventional optical colonoscopy for mass screening for colon polyps – the precursor of colorectal cancer. The patient's abdomen is imaged by a helical CT scanner during a single-breath-hold. A 3D model of the patient's colon is then reconstructed from the CT scan by automatically segmenting the colon out of the abdomen followed by *electronic cleansing* – computer-based removal of residual material in the colon. The system, running on a PC, allows physicians to interactively navigate through the colon and view the inner surface using volume rendering, with tools for measurements, *electronic biopsy*, to inspect suspicious regions, as well as painting already seen areas to help in visualizing 100% of the surface. The interface shown above provides multiple linked views: 2D axial, sagittal and coronal views (right); an oblique slice perpendicular to the colon centerline (middle left); an outside 3D colon model with current virtual position and orientation, bookmarks of suspicious regions, and the centerline in green (upper left); volume rendered endoscopic view with the centerline and a polyp (center); and a flattened volume rendered biopsy view (left). Unlike optical colonoscopy, VC is a patient friendly, fast, non-invasive, more accurate, inexpensive procedure. Grant-funded university research of VC at Stony Brook University (SUNY) led to a license to Viatronix Inc. that has installed the technology in numerous sites by which the lives of hundreds of patients have been saved. VC has been extended to 3D virtual endoscopy of other organs, such as the heart, arteries, lungs, stomach, and bladder. The primary future challenge in VC is in the development of computer-aided detection (CAD) of colonic polyps.

L. Hong, S. Muraki, A. Kaufman, D. Bartz, and T. He, "Virtual voyage: Interactive navigation in the human colon," *Proc SIGGRAPH* 1997, pp. 27-34.

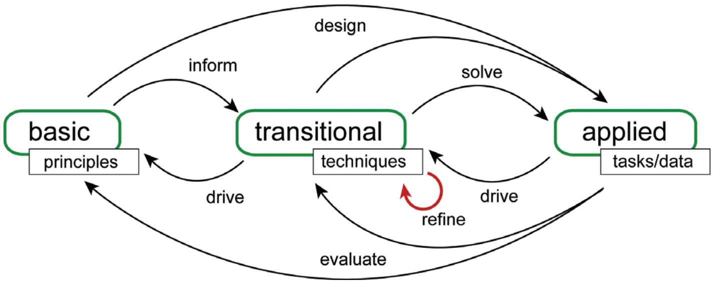


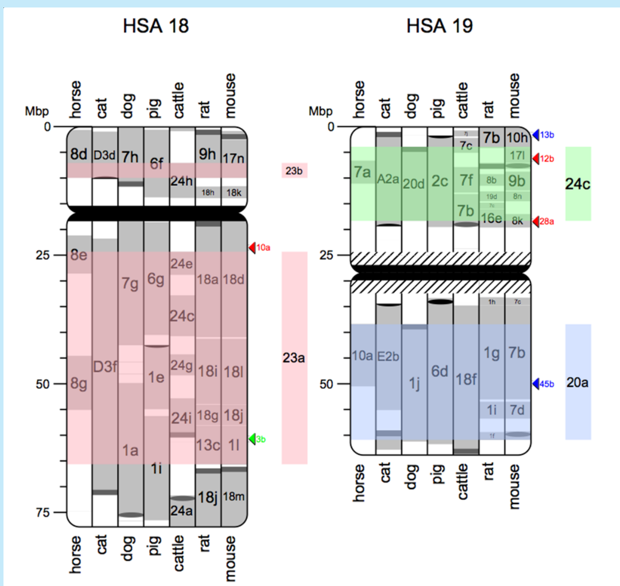
Figure 3.1: *Cycles at several levels interconnect the areas of basic, transitional, and applied research.*

into these three categories. Visualization research often gives rise to complex cyclic relationships between the divisions, and so it requires funding programs that encompass the entire breadth of these divisions.

Consider the diagram in Figure 3.1. In visualization, basic or foundational research areas include psychophysics, perception, and visual encoding; while, on the other end of the spectrum in application-driven research, data and domain tasks, including computational fluid dynamics, medical imaging, telecommunications networking, and software engineering, provide driving problems. The bulk of the work of the field exists not at either pole but between the two poles, in the transitional research of creating and refining techniques. The red arrow in the diagram depicts the local feedback process of research progress in the transitional area; all the other arrows in the diagram show important cycles of interaction between these three areas. We must support these cycles in all their phases, as well as their feedback interactions to keep the field of visualization vibrant.

Although transitional research is at the heart of any field, in some areas of visualization a disproportionate amount of attention is currently devoted to incremental refinement of a very narrow set of techniques.

The ideal research program of a visualization researcher or group would include all cycles in the above diagram, targeting basic, transitional, and application-driven research. The fruits of basic research are principles that inform transitional research, the needs of which in turn drive basic research. The fruits of transitional research are techniques that solve applied problems, the needs of which drive the development of new and better techniques. Application-driven research contains its own cycles; in one direction, basic and transitional



Visualization of Genomes

Several visualization tools and products have been developed to provide a visual means to compare genomes. The genomes of multiple mammalian species – in the case shown here, humans, horses, cats, dogs, pigs, cattle, rats, and mice – can be compared using Evolution Highway. This tool, built by NCSA, removes the burden of manually aligning these maps and allows researchers' cognitive skills to be used on something more valuable than preparation and transformation of data.

The visual chromosome metaphor presented by Evolution Highway allows comparison of the orthologous sequences that humans have with other species to see where chromosomal breakpoints occur. The breakpoints that show up in the same place across multiple distinct species, are thought to mark fragile places in the genome where rearrangements are more likely to occur. The tool helped show that the historical rate of chromosome evolution in mammals was different than previously thought and revealed provocative new features of chromosome breakpoints as they relate to cancer.

An Agilent software product "CGH Analytics" elucidates the genomic anomalies underlying many forms of cancer as well as developmental disorders. Many cancer researchers have used this tool to study the cancer-related genomic variation in diseased tissue. The tool allows viewing tens of thousands of array-based Comparative Genomic Hybridization (aCGH) measurements across many different tumor samples or cancer cell lines, allowing rapid identification of regions of common aberration at the global genome scale as well as at the single gene level.

W.J. Murphy et al., Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps, *Science*, Vol 309, Issue 5734, pp. 613-617, 22 July 2005. (http://evolutionhighway.ncsa.uiuc.edu/index.html)

R. Kincaid, A. Ben-Dor, and Z. Yakhini, Exploratory visualization of array-based comparative genomic hybridization, *Information Visualization* Vol. 4, No. 3, pp. 176-190, 2005.

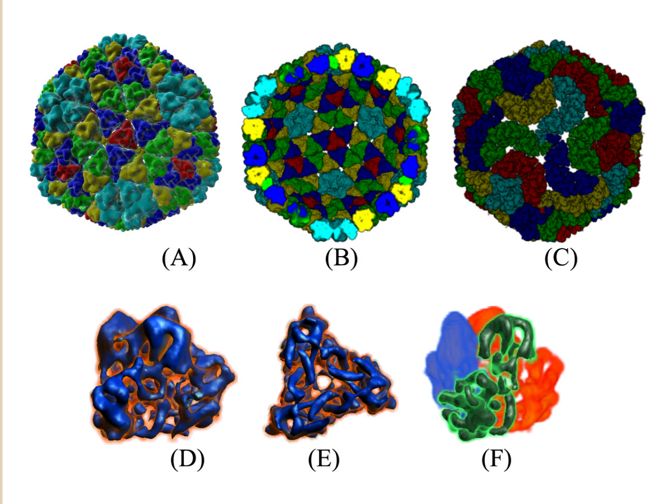
research provide design knowledge for application problems, while, in the other direction, evaluating the solutions to application problems closes the loop, allowing refinement of techniques and principles by analyzing the reasons why a particular approach did or did not solve the intended problem.

Designing and building systems that solve real-world problems is the best way to make significant progress in refining and adding rigor to both the techniques and the theoretical foundations of visualization. The iterative process of science is to make observations, construct theories to analyze and explain them, and continue the cycle by using the theory to guide the next set of observations. In visualization, we must build a working system before we can gather observations of its use. Building systems for real users with real tasks allows researchers to gather valid data and evaluate whether and how visualization techniques are effective for the intended task. These observations and explanations grounded in specific techniques create a foundation from which we can draw general theoretical conclusions about visualization. Another advantage of using real datasets is that researchers are then driven to create robust and scalable algorithms. Many visualization algorithms that work well for "toy" datasets do not scale to the large or noisy datasets of interest to real users.

We must close the loop, accelerating the maturation of basic research into applied research. Visualization already functions as a crossroads connecting foundation research with applications, integrating the capacity of computational techniques to simulate and model natural and societal phenomena and to predict and report results. However, we must refine the precision with which we balance the resources in research support, working to promote visualization solutions to realworld problems by providing end-to-end approaches for growing new science into practical answers to hard and important questions. With guidance, the discipline of visualization will become a powerful partner for those addressing scientific and social problems.

Finding: A disproportionate percentage of current visualization research is transitional. Although technique refinement is a core aspect of our field, a balanced visualization portfolio should also be driven by applied problems and grounded in basic research.

Virus Structure



Determining the three-dimensional structure model of a virus is the first step towards understanding its virulent function. A new experimental imaging approach utilizes cryo-Electron Microscopy (cryo-EM) for elucidating single particle macro-molecular structures (such as viruses) at the highest quasi-atomic resolution, typically around 10 A. In the above figure (A-C) are visualizations using combined surface and volume rendering of the same half-shell model of the Rice Dwarf Virus (RDV). The different colors show the nucleo-capsid shell from the outside (A,C) and the inside (B), elucidating the local and global complexity of the quasi-symmetric packing of the individual structural protein units. Each individual structural unit, exhibiting a trimeric fold, is further visualized in (D) and (E) from two different views, as well as with different colors (F), to show the three different conformations of the monomeric protein chain forming the trimeric structure unit. All of the structure units are automatically segmented from a reconstructed 3D electron density map. Previously structural biologists have largely attempted this 3D ultrastructure elucidation steps, manually. The quasi-atomic models of these viruses and other macromolecular machines, constructed via our structure elucidation pipeline provides micro-biologists and drug discovery scientists, crucial insights into molecular interactions within macro-assemblies. The large physical size and complexity of such complexes, combined with the very low signal to noise ratio of cryo-EM, still presents significant computational and experimental challenges in this research.

Z. Yu, and C. Bajaj, Automatic ultra-structure segmentation of reconstructed cryo-EM maps of icosahedral viruses, *IEEE Transactions on Image Processing: Special Issue on Molecular and Cellular Bioimaging,* Vol. 14, No. 9, pp. 1324-1337. 2005.

THE POWER OF VISUALIZATION 4

Visualization is poised to break through from an important niche area to a pervasive commodity, much as the Internet did after decades of Federal funding. Visualization has already had significant impact in medicine, science, engineering, and business, and has the promise to dramatically transform these and other important social endeavors in the future. Visualization has the potential to transform virtually every aspect of our society, helping to make us healthier, safer, better informed, and more economically competitive. This chapter lays out a vision for how visualization can transform health care, science and engineering, and daily life.

4.1 Transforming Health Care

The escalating cost of health care in the nation concerns everyone, from national policymakers to health care professionals to individual consumers. Visualization has the potential to transform health care, both nationally and globally, by lowering cost, improving quality, accelerating research, and empowering individual consumers.

Bioinformatics Visualization Now that scientists have mapped the human genome, we are faced with the challenges of transforming this knowledge into medical tools and procedures that will eventually combat disease and improve human health on a global scale. Visualization will play a critical role as we journey towards understanding how to use this information effectively for health care. For instance, the need to develop new pharmaceuticals will require a deep understanding of the complex chain of chemical reactions governed by enzymes that are themselves regulated by genetic sequences. Within this context, effective visualizations will illuminate dynamic processes such as degenerative arthritis, embryological development and differentiation in stem cells, and the imbalances among adult stems cells that lead to conditions such as osteoporosis.

As we work to understand these processes, visualization will play an essential role in mapping gene expression, displaying the causes of failure of healthy physiological regulation, and aiding in the development and monitoring of restorative and management therapies. Alliances between biochemists, physiologists, pharmacologists, and technical visualization tool developers will arm researchers with the capacity to tap their intuition, rapidly prototype new drugs, and accelerate and facilitate pioneering research in treatments that do not simply address symptoms but target the basic biology governing human health.

Surgical Support Visualization is already successfully helping surgeons more quickly and easily comprehend medical imaging data drawn from scanners such as MRI and CT. Using the imaging data drawn from these scanners, visualization specialists have had some early successes in creating tools for surgical planning, designing procedures, and predicting outcomes. Visualization support during surgery using augmented reality offers great promise in merging information acquired preoperatively with the surgeon's view of patient anatomy as it is revealed in real time[2]. New surgical techniques being developed through partnerships in visualization and medicine combine the ability to see through the skin with minimally invasive tools to perform surgery with precision and with almost no disruption of the body itself.

So far, this visualization capability has been limited to extensively planned demonstrations at a few cutting-edge hospitals. The challenge is to make it available for routine use by surgeons at all hospitals in the country. This will require significant further development of visualization techniques and software, better augmented reality systems, and the successful integration of these systems in the operating room. In addition, surgical planning using these technologies has so far been limited to static images. We must increase the efficacy of these technologies by extending their use to deformable adaptive models that track procedures and therapies during an operation, helping to plan for contingencies and avoid complications. Evaluation of the effectiveness of these visualization techniques will be critical in this endeavor.

Prevention and Policy Historically, the most effective way to lower the cost of health care is to prevent problems before they start. Vaccines cost less and do more to further public health than any devised intervention for treating illness. However, to fight such diseases as influenza through vaccines, public health officials must integrate the best information available to predict the virus strains most likely to emerge in the coming season. Visualization can not only help improve the predictive power of current methods but also facilitate the incorporation of additional factors and information into the decision process.

We know too little about how exploratory data visualization tools can be most effectively employed by epidemiologists, bio-statisticians, or others to explore health and related data, generate and evaluate hypotheses, summarize findings, and advance public health. Continued development in geoscience visualization (see Section 4.2) will benefit the health sciences as well. It is important to stimulate collaborative work between geo/information visualization (and exploratory data analysis) researchers and health researchers so that advances in visualization fit the conceptual framework for analysis of those with knowledge in the domain.

It is also essential to focus research on how visual methods are used or not used, and this research cannot be confined to the laboratory. Field-based participant observation, knowledge elicitation, cognitive task analysis, etc., are important strategies for building the understanding of real-world problem solving with visual tools in public health and other essential areas. In addition, limited progress has been made in integration of visual, statistical, and computational methods into usable methods for work in public health research. Observations of prevention and policy decision processes and exploration of integrated methods are essential for targeting visualization research on relevant problems and for the construction of usable systems that empower society to take command of emerging issues in public health.

Image Guided Surgery

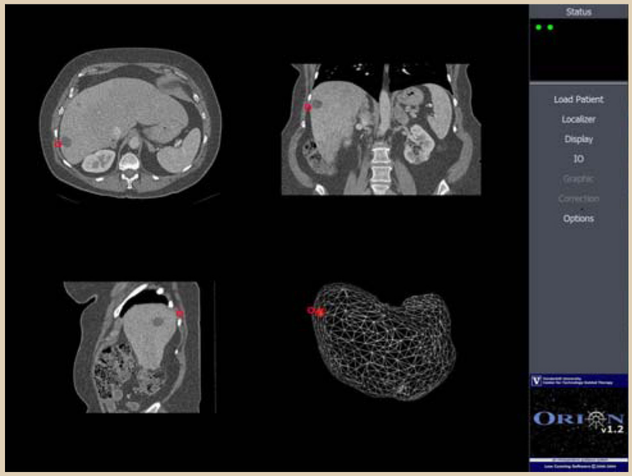


Image-Guided Surgery was made possible by the rapid rise in visualization technology. Large multislice image sets can be manipulated in real time providing instantaneous feedback to the surgeon as to their surgical location relative to targets they wish to hit and structures they wish to avoid. In the display screen by the surgeon, the tumor can be seen as a dark spot in the liver tissue and a “floating” red ball in the three-dimensional wireframe.

JD Stefansic, A.J. Herline, Y. Shyr, W.C. Chapman, J.M. Fitzpatrick, and R.L. Galloway, Registration of physical space to laparoscopic image space for use in minimally invasive hepatic surgery. *IEEE Transactions on Medical Imaging,* Vol. 19, No. 10, pp. 1012-1023. October 2000.

Biological Imaging The field of biological imaging is exploding. We have developed new medical imaging modalities and increased the performance of many existing ones. We have also found new scientific and biological applications for such data. These new applications do not necessarily coincide with the conventional clinical paradigm, where a highly trained radiologist looks at individual patients and makes careful, specific diagnoses. Instead, biological imaging deals with populations of subjects, and the goal is not necessarily diagnosis but quantifying some aspect of the data in order to test a particular hypothesis. These new trends in biological imaging require new methods for processing large 3D datasets. For instance, the strategy of discriminating tissue types based on homogeneous intensity values is not as applicable in these new biological domains.

Personalized Medicine Medical care today involves a coordinated interchange of specialties and information management. Modern, rapid, accurate diagnosis requires the integration of a spectrum of data about each patient from an increasing number of available laboratory tests. The response to any diagnosis may incorporate patient-specific characteristics including gender, age, body-mass, personal genetics, medical history, individual anatomy, diet, family history, or idiosyncratic responses to pharmaceutical agents. Future routine medical care may utilize time in the waiting room by automating data collection, acquiring a patient’s immediate vital statistics such as temperature, pulse, blood pressure, and weight, and even perhaps performing 3D medical scans and creating customized models of the patient’s current condition merged with her medical history. The resulting adaptive 3D digital patient record will be the foundation of a new generation of medical tools, a generation that incorporates visualization at its core.

As we automate and aggregate volumes of information about a patient, health care professionals will need increasingly powerful tools to view the patient's condition at a glance. For instance, while a lab technician may immediately be able to interpret a raw column of numbers, another specialist may benefit from seeing a visual representation of those test results in the context of the normal range of values, or the range for all patients with this particular condition, or how they have changed over months or years for this particular patient. The LifeLines system[38] is a first step towards an interactive visual representation of complex medical histories.

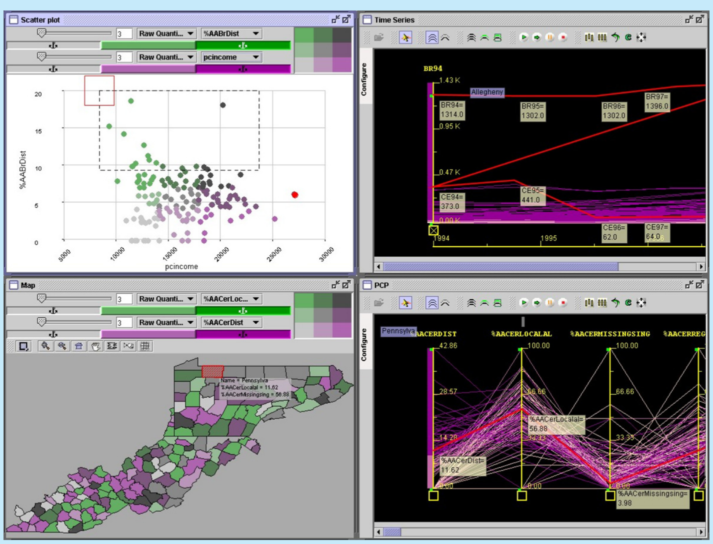
Beyond these first efforts, applying visualization techniques to the process of healing will help us coordinate complex data and fit a particular patient into the spectrum of medical literature and experience to pinpoint the problem and find the right treatment as quickly as possible. The promise is not only for reducing errors in medication and other forms of treatment–enormous problems in medicine today–but for the development of a truly personalized medicine.

4.2 Transforming Science and Engineering

In its early years, of the scientist's job was to observe immediate surroundings through direct perception and draw conclusions from that observation, understanding the laws of gravity, for example, by measuring and recording the time required for an object to fall a given distance. Today we routinely record information across a wide range of spatial scales from femtometers to parsecs, at time increments as small as attoseconds, and with instruments that monitor frequencies far above and below the range of visible light. Scientists and engineers must explore the resulting enormous datasets, which outstrip the capabilities of today's visualization algorithms. We need to develop techniques to work with these time-varying, unstructured, irregular, multifield, multidimensional, massive datasets. New algorithms for the visual presentation of information need to be integrated with techniques for information extraction and abstraction.

Physical Sciences Many areas of the physical sciences are experiencing a flood of data arising in part from the development of instruments that acquire information on an unprecedented scale. Some of the most celebrated examples are the Sloan Digital Sky Survey[47] and the COMPLETE project[1], generating terabytes of astrophysics data each day. Although some interesting results have come from technology transfer of volume visualization methods originally developed for medical imaging[7], much work remains to develop visualization systems tuned for both the scale and domain requirements of astrophysics.

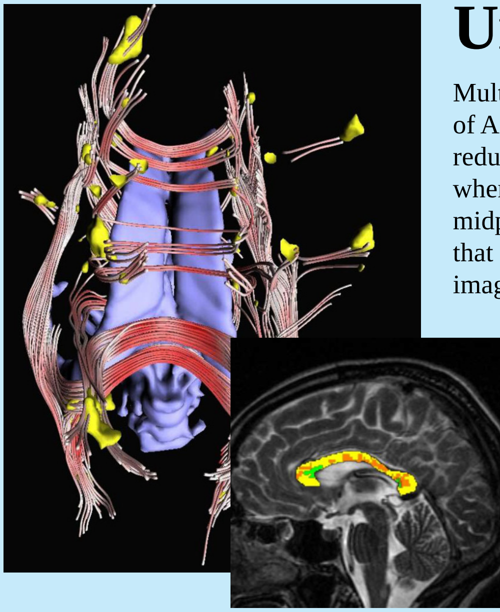
Similar problems with large data exist in physics. For example, the Large Hadron Collider (LHC) at CERN is being



Health Demographics

Many problems in public health management have complex interdependencies with factors such as education, poverty, environmental quality, safe water, clean air, climate, season, and even animal migration patterns. The Exploratory Spatio-Temporal Analysis Toolkit figure, was designed under contract from the National Cancer Institute and implemented using the GeoVISTA *Studio* system. The four linked views support the flexible visual exploration and analysis of geo-spatial health data and covariates across space and time. The data displays in the example show possible correlations between stage of diagnosis for cervical cancer (comparing "local stage", meaning the cancer has not spread, and "distant stage," meaning the cancer is found in other organs) in women across the economically disadvantaged Appalachian counties in PA, WV, and KY. The scatterplot in the upper left section shows income against distant stage diagnosis for breast cancer. Existing visualization tools have already empowered public health officials with the capability to trace such factors and study their relationships, but they do not go far enough. Innovative responses in public health management require that we develop integrated visualization systems that will enable officials to pull together traditional and non-traditional information to explore the relationships between disease factors and to create policies and programs that respond to the essential causes of health problems and not just the effects.

A.C. Robinson, J. Chen, H.G.. Meyer, and A.M. MacEachren. Human-centered design of geovisualization tools for cancer epidemiology. In *Proc. GIScience*, pp. 314–316. 2004.



Understanding Multiple Sclerosis

Multiple Sclerosis (MS) is a neurodegenerative disorder that affects hundreds of thousands of Americans; understanding its progression better could lead to new treatments that could reduce the devastating effects of the disease. The image on the bottom shows orange spots where neural “wires” or fibers-at-risk (FAR) for the brain visualized on the top pierce the midplane of the brain. Fibers at risk are those that pass through lesions. We hypothesize that these piercing points anticipate where new pathology will develop – in the bottom image existing pathology is shown in green. The view on the top is from above; fluid-filled regions are shown in blue, lesions are shown in yellow, and fibers that pass through lesions in shades of red. Continued research is needed to test our hypothesis and to extend and refine this fibers-at-risk approach to better understand MS and other diseases.

J.H. Simon, S. Zhang, D.H. Laidlaw, D.E. Miller, M. Brown, J. Corboy, D. Singel, J. Bennett, Strategy for detecting neuronal fibers at risk for neurodegeneration in earliest MS by streamtube tractography at 3T. In *Proceedings of ISMRM*, Miami, FL, May 2005.

developed to investigate how the universe began, the origin of mass, and the nature of antimatter[51]. It will detect and record 100 interesting collision events a second, corresponding to a data rate of about 100 MB per second, and produce 1 petabyte or the equivalent of 20 million CDROMs a year.

While the proposed grid technologies tackle issues of computation and storage[14], developing scalable visualization solutions for this kind of data will require handling orders of magnitude more data than current methods allow. One of the challenges is to develop appropriate abstractions that can show overviews of the data without aggregating away the fine-grained details of possible interest to physicists.

GeoSciences Almost one-third of the U.S. Gross Domestic Product ($3 trillion) is derived from weather-dependent industries. Although weather prediction has progressed to the point that a 3-5 day forecast is generally close enough to be considered correct, longer-range and more accurate forecasts are needed. Improving weather forecast accuracy by just one degree Fahrenheit would save $1 billion a year in energy costs by increasing our ability to optimize energy generation and distribution[24]. Many of these improvements will be driven by new visualization methods to examine the interaction between systems and fronts. Visualization researchers must work with meteorologists to develop visualization tools that show the necessary phenomena and interactions in ways that are most meaningful to scientists. For example, better storm prediction can save lives and reduce costs, as well provide a better quality of life. Hurricane analysts want to see multiple co-located values (wind speed, pressure, temperature, percentage dissolved water vapor, etc.) at multiple locations to predict hurricane strength and path. We need to develop new visualization techniques that can show this multitude of variables in a comprehensible way.

Because almost every area in the nation is vulnerable to earthquake, accurate damage predictions have the potential to save millions of dollars, not to mention lives, over the long run. To date, researchers have developed good models that simulate ground motion and, separately, good models that simulate structural damage based on ground motion. However, because of differences in model design, grid structures, time scales, and prognostic variables, the two classes of models remain largely uncoupled. So far, the primary coupling mechanism has been visualization of the two model outputs in the same view volume–a mechanism of limited use. In addition, though we have deployed sensor networks to measure seismic activity and structural deformations, we have not yet developed a methodology to assimilate the resulting information into a coherent picture. In order to address the enormous problems posed by earthquakes, we need visualization systems that allow researchers to manipulate these enormous data sets, to combine computational models from several disciplines to shed new light on the consequences of earthquakes, and to develop new technology by which we can better analyze seismic hazards.

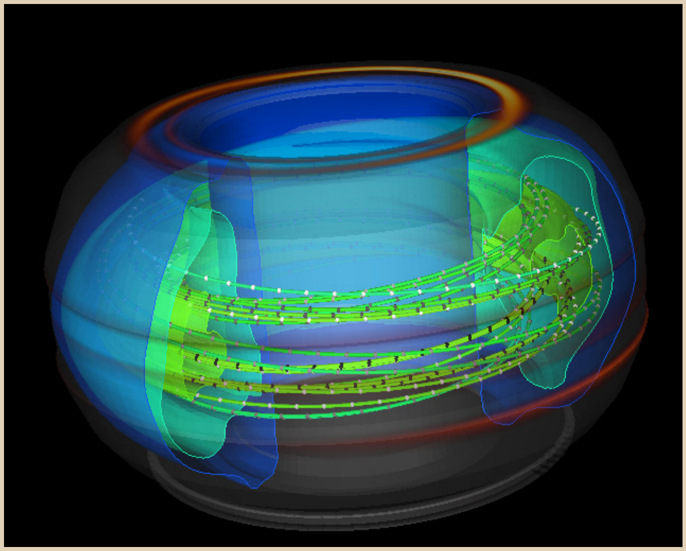
We are seeing a dramatic increase in the deployment of sensor, video, and other remote environmental data collection networks (for air borne and water borne chemicals, to monitor water quality around landfills, etc.). The technology is improving quickly enough to make this a trend with exponential growth potential. A big visualization challenge will be to transform our thinking and methods to deal with continually changing, streaming data inputs with uneven and changing quality. Visualization is used in this case both to understand the phenomenon in real time (based on data of varying quality) and to understand the functioning of the sensor network, to allow system managers to figure out where new resources are needed.

Engineering We will always face engineering grand challenges, from learning to manipulate single atoms to developing the ability to build longer bridges and taller buildings. Engineers need assistance in making complex decisions and analysis, especially with tasks involving large amounts of data. Often engineers have to deal with overspecified situations, and the greatest challenge is to filter out the irrelevant data. Visual analysis systems are needed that allow "what if" scenarios, that allow data to be examined under multiple perspectives and assumptions, to seek connections between any number of attributes, and to understand the reliability of any conclusions reached.

In order to meet construction challenges, material scientists are working to develop stronger but lighter materials for manufactured parts. To accomplish this, they must analyze the interaction of many correlated properties, including tensile strength, elasticity, conductivity, toxicity, and reflectivity, some of which are likely fail to reveal valid new materials. We need to develop better methods to show the interaction of the various parameters and their effect on material properties.

In order to better see fluid-structure interactions, computational fluid dynamic researchers need more robust and more accurate methods to find and show critical features like shedding vortices. If designers are able to observe the effect of moving or reshaping part of a large manufactured system in real time, they will be able to test competing designs and therefore achieve an optimal design more quickly and with less working memory, decreasing design and development costs. In order to achieve these results, we must develop new visual data representations and better ways of representing object hierarchies so that all necessary information is visible and associated data can be quickly shown when needed.

Fusion Plasma Physics



During tokamak experimental operation, events occasionally occur that rapidly terminate the plasma discharge. For future experiments, such as the International Thermonuclear Experimental Reactor (ITER), the stored energy will be approximately 100 times greater than in present day devices and these disruptions have the potential to severely damage the material wall, especially if the heat flux is highly localized. Shown here are the results of a NIMROD simulation of a particular disruption in the DIII-D tokamak. The temperature isosurfaces, the magnetic field lines, and contours of the heat flux on the wall are visualized. The heat flux contours show toroidal and poloidal localization as a result of the topology of the magnetic field lines. The localization results from the plasma instability compressing the flux surface in the core of the plasma and then transporting the resulting "hot spots" to the wall. Visualizing the data in 3D shows where the plasma wants to preferentially bulge, enabling the development of improved disruption mitigation techniques.

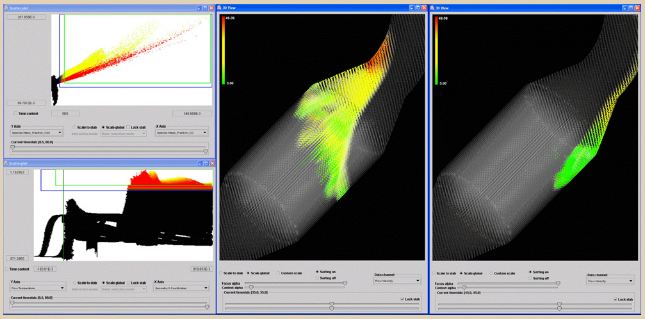
S.E. Kruger, D.D. Schnack, and C.R. Sovinec, "Dynamics of the major disruption of a DIII-D plasma", *Physics of Plasmas* Vol. 12, No. 056113. 2005.

Social Sciences Visualization has thus far had less impact on the social sciences than the physical sciences, in part because of a dearth of funding for such efforts, but it holds the promise of effecting similar transformations. For example, advances in visualization tools for archaeologists could allow interactive exploration of rich interlinked spatiotemporal information, such as the spatial location at which artifacts were found and the conjectured temporal and usage relationships between them, that is currently difficult to track simultaneously. Developing new visualization techniques for helping people understand the relationships between large networks of entities could help a wide variety of social scientists, including sociologists who seek to understand human social networks, librarians and others who use bibliometrics for co-citation analysis of document databases, and linguists working on automating natural language parsing and translation.

4.3 Transforming Life

Visualization is not used only by doctors, engineers, or scientists – it is both produced and consumed daily by millions of

Visual Engineering Analysis



A diesel particulate filter (DPF), which collects soot from automotive exhaust, needs to be cleaned periodically to avoid becoming clogged. The filter regeneration process (soot oxidation) needs to be quick, to be as complete as possible, and produce minimum pollution, resulting in a multi-parameter optimization design problem. Interactive visual analysis of large and complex simulation data allows one to understand complex relations within datasets like the gas flow through a DPF during filter regeneration. The interactive analysis here involved the joint investigation of 10 data dimensions (x, y, z, time, velocity, temperature, carbon monoxide, carbon dioxide, oxygen, and soot). Multiple linked views were used to drive the analysis concurrently in attribute space as well as in the space and time of the simulation. The image shows how the oxidation front in the DPF has been first selected by means of interactive brushing in two views of the attribute space of the simulation (on the left side) and then visualized at two different points in time on the right side (color encodes velocity). With this kind of interactive visual analysis it was possible to show that (due to a particular bend of the exhaustion system before the DPF) not enough oxygen was transported to one side in the back of the DPF and that therefore the oxidation dies off there before burning all the soot, requiring an improved design of the exhaustion system. One main research challenge for this type of engineering problem is the improved integration of data semantics within the process of interactive visual analysis, i.e., which features to show in the context of which others.

H. Doleisch, M. Mayer, M. Gasser, R. Wanker, and H. Hauser, Case study: Visual analysis of complex, time-dependent simulation results of a diesel exhaust system. *Proceedings of the 6th Joint IEEE TCVG—EUROGRAPHICS Symposium on Visualization* (VisSym 2004), May 2004, pp. 91-96.

ordinary people. As producers, we might use online services to generate location specific maps and driving directions. Or we might generate charts to analyze the cash flow in bank accounts or stock market portfolios using personal finance software. As consumers, we have learned the sophisticated symbology of weather maps, and we enjoy colorful explanatory visualizations in print or electronic media.

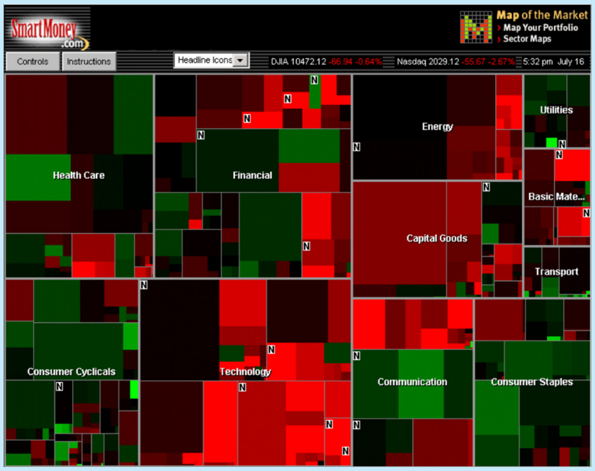
However, the real power of visualization has yet to be tapped in ordinary life. The information big bang has reached our homes, requiring us to manage large email, photo, and music collections, hundreds of TV channels, and the plethora of information available online in the form of public databases, web pages, blogs, podcasts, and streaming video. The rapid and seemingly boundless growth of Google alone is testament to the importance of data mining tools for the public. Yet, the most effective information management solutions we currently offer the ordinary person faced with these enormous challenges are text-based search engines. The visualization research community must work to expand the area's reach and benefits to the mass market.

Mass Market Visualization A big challenge is to create a kind of "Visualization Photoshop" or "Visual Google," a system that, while clearly not comprehensive and all-powerful, does help to enable non-experts to perform tasks otherwise beyond their capabilities in any reasonable time frame. These systems must allow ordinary people to experiment with "what if" scenarios, to examine data under multiple perspectives and assumptions, to seek connections among any number of data attributes, and to understand the reliability of conclusions reached by analyzing data. The goal is to make visualization a ubiquitous tool that enables ordinary folks to think visually in everyday activities.

Another challenge is to remove visualization from the desktop computer environment and move it into information appliances. For example, my refrigerator may generate a graphical analysis of the contents to visually depict the odds of having a good meal this weekend if I skip the trip to the grocery store. Or it may show estimates (based on live feeds about current sales) of the relative advantages of shopping at my three closes stores based on what I need to restock. Visualization researchers need to address the range of devices available in people's homes and businesses. As users move between cell phones, PDAs, laptops, desktops, and wall sized displays, visualizations should adapt to the unique input and output characteristics of the devices.

We need to develop specific techniques to address the needs of older adults or users with low cognitive abilities who are simply overwhelmed by the display complexities that seem trivial to designers. Making visualization tools accessible to users regardless of their background, technical disadvantages, or personal disabilities remains a huge challenge. For example, visually impaired users may need to use automatically generated text-based alternatives to visual displays. Encouraging results have been found with the sonification of simple graphs, scattergrams, and tables. Spatial sound might help sonify more complex data representations. High-resolution tactile displays, which may provide appropriate solutions in certain cases, are already appearing.

Mapping the Market



The Map of the Market (www.smartmoney.com/marketmap), launched in 1998 to show stock market data on the web, is an example of visualization for the mass market rather than for specialists in a specific scientific domain. Many people and companies are extremely interested in understanding this large, publicly available dataset with complex structure at multiple levels of detail. Each colored rectangle in the map represents an individual company, the rectangle's size reflects the company's market capitalization and the color shows price performance. This treemap technique was originally introduced in 1991, and has been steadily refined by many researchers to increase its effectiveness and scalability. It has been successfully applied to many kinds of hierarchical data including digital image collections, baseball statistics, gene ontologies, and election results. One of the remaining visualization challenges is to better communicate the time-varying aspect of the data; the image above shows the change over a single time period, whereas the full dataset contains many time steps. Showing the combination of hierarchical structure and dynamic patterns over multiple time scales will require the development of new techniques.

M. Wattenberg, Visualizing the stock market, *CHI Extended Abstracts*, pp 188-189, 1999.

Security Security is the foundation for a civilized and free society. Yet it has been evident since the attacks of September 11, 2001 that our national and global security is tenuous. The U.S. Department of Homeland Security chartered the National Visualization and Analytics Center (NVAC) in 2004 with the goal of helping to counter future terrorist attacks in the U.S. and around the globe. A major objective for NVAC is to define a five-year research and development agenda for visual analytics to address the most pressing needs in R&D to facilitate advanced analytical insight.

We must develop tools to support analysts who are trying to do multi-source analyses relating many qualitatively different kinds of data (images, sensor produced data, financial transactions, maps, etc.). Our visualizations need to help analysts cope with data of uncertain quality as well as uncertain relevance, and to find low frequency patterns in large, messy data sets. Future visualization tools must be able to extract the pertinent information automatically and assimilate and fuse these data into a coherent picture.

The visualization community has taken on these and other challenges of critical importance to our national security. In 2005, many leaders in our field came together in a series of meetings to discuss these issues. Their findings are enumerated in detail in the recently published NVAC report[48]. Jim Thomas, the director of NVAC, has coined the term "visual analytics" to refer to "the science of analytical reasoning facilitated by interactive visual interfaces." Visual analytics has been wholeheartedly embraced by the visualization community. The first symposium Visual Analytics Science and Technology (VAST) symposium is scheduled for 2006, co-located with the IEEE Visualization conference in Washington D.C.

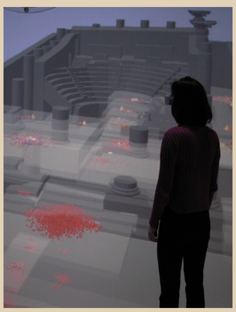
Business The information big bang has hit business and finance especially hard: companies and individuals spend a great deal of money and time collecting and curating information in hopes that it will give them a competitive advantage. However, far less Federal funding has been devoted to the problem of creating visualization systems to meet these needs than those of the physical and life sciences. Some first steps have been taken to meet the need for visualization tools to help people understand exploding business and financial information. Recently, a visualization of book buying trend data motivated a reorganization by O'Reilly Books that resulted in it being the only computer book publisher to increase market share after the dot-com crash[31].

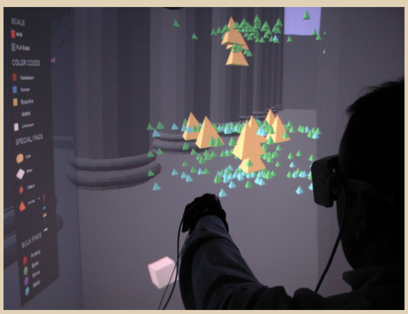
Further research in visualization must focus on helping companies make decisions or take action on the information they have gathered. The challenge of creating visualization systems to unlock the information in massive databases and data warehouses is considerable, both in terms of scaling up to handle a large total number of records and in terms of addressing the large number of dimensions of data contained in each record. Unlike many of the databases in the physical and geosciences, business these datasets often have no inherent spatial characteristics. In addition, the kinds of question users seek to answer in relation to these often differ in kind from those addressed by scientific communities. AT&T maintains a database of all calls made over its long-distance backbone for a year, a huge network of 250 million phones on which hundreds of millions of calls are made each day. Visualization and statistical processing have already been used to hunt down call fraud, but the sheer size of this dataset will require more advanced techniques to be developed to fully exploit this resource. Likewise, companies ranging from small businesses to behemoths like Walmart would benefit from more powerful visualization techniques to carry out market-basket analyses on their transaction logs.

Education The continued success and advancement of any society is inextricably bound to the presence of an effective and relevant education system. Like professionals in many other fields, educators are awash in a sea of data, including standardized tests, quarterly assessments, schoolbased assessments, and chapter tests. Each score is a potentially valuable clue to how a child is doing and what might be done to help them succeed. Visualization can show the patterns and relationships, providing teachers and administrators with valuable understanding.

Visualization also has a place in the K-12 classroom, helping students explore and understand quantitative topics of all sorts. Visual presentations of data and concepts can benefit students working in subjects ranging from science to math to social science. Visual presentations particularly benefit students who do not learn easily from verbal presentations, thus especially helping students who may be struggling. Interactive graphics also increase student engagement and thus time spent on task, both of which improve learning and comprehension.

Key research challenges in the application of visualization to education include the development of intuitive and flexible representations for the discovery of pattern in data about individual students and subgroups, the identification of the mechanisms by which different types of interactive visualizations enable learning, and the investigation of how that process varies across individuals and stages of cognitive development.



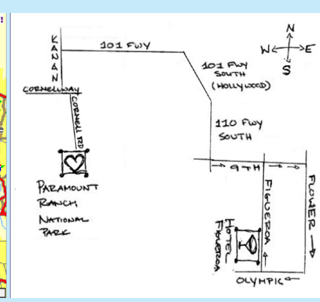


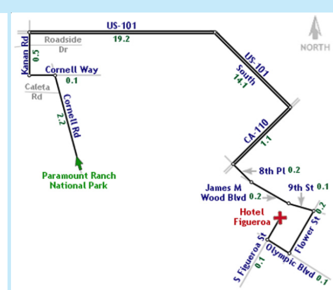
Virtual Archaeology

Archaeologists collect vast amounts of data that require the exploration and understanding of complex spatial relationships among the artifacts and surrounding architectural ruins. Common analysis practices involved the use of drawings, photographs, hand-written reports, and relational databases, making it very difficult to visualize and comprehend those three dimensional correlations. These two images show ARCHAVE, a virtual reality environment for archaeological research in which archaeologists perform queries and analysis of data collected on-site. By immersing the user in the three dimensional structure of the data and the excavation site, spatial relationships become immediately apparent, and complex comparisons among different types of information can be performed through simple interactions with the application. On the left, an archaeologist explores the excavation trenches and collected artifacts at the Great Temple of Petra site in Jordan. On the right, a user explores one of the trenches in which remains from Byzantine lamps (large yellow pyramids) appear to concentrate among other minor metal and bone remains (cyan and green small objects respectively). The addition of interactive annotation, along with more complex types of data and meta-data, and adequate and adapted on-site data gathering techniques remain as important and exciting challenges in this application field.

Eileen Vote, Daniel Acevedo Feliz, David H. Laidlaw, and Martha Sharp Joukowsky, Discovering petra: Archaeological analysis in VR. *IEEE Computer Graphics and Applications*, Vol. 22, No. 5, pages 38-50, September/October 2002.

Rendering Effective Route Maps





Three route maps for the same 35 mile route rendered by (left) a standard computer-mapping system, (middle) a person, and (right) LineDrive, an automated route map rendering system. The standard computer-generated map is difficult to use because its large, constant scale factor causes the short roads to vanish and because it is cluttered with extraneous details such as city names, parks, and roads that are far away from the route. In contrast, the hand-drawn map emphasizes the most essential information for following the route. Hand-drawn maps exaggerate the lengths of short roads, regularize turning angles and simplify road shape to ensure that all the roads and turning points are visible. LineDrive is based on these cognitive design principles and as a result it produces maps that similarly emphasize the most important information for following the route.

M. Agrawala and C. Stolte, Rendering effective route maps: Improving usability through generalization, *SIGGRAPH 2001*, pp. 241-250. 2001.

ROADMAP 5

Visualization as a discipline is at a critical juncture, poised to become an invaluable asset in areas across society. There are powerful lessons to be learned by comparing networking and visualization research, after the global internet with its many advantages was made possible only through long-term development and research support among Federal sponsors and commercial concerns. Penetration of internet technologies into almost all sectors of science and society has accompanied and enabled explosive economic growth worldwide. Like high performance communications, visualization today has undergone eighteen years of development and is poised to induce radical changes in our approach to research. Though associated mostly with the gaming and entertainment industries, the computer graphics technology for visualization was originally developed along with high performance computing and communications for science and engineering domains. The penetration of these technologies across modern computing assures us of the install-base necessary to make profound changes in scientific and other forms of exploration. We need only to take advantage of the advanced graphics capabilities already deployed with today's computer platforms to transform medicine, science, and lifestyles.

In response to the findings of this report, we lay out the following roadmap, broken down by the schedule of actions suggested in the short-term, mid-term, and long-term timeframes. We have listed the three divisions in progressive order of cost and effort we expect will be necessary to achieve them.

All of these recommendations have the potential to make a lasting impact on national research and development. We refrain from making specific assignments of these efforts to agencies, limiting our recommendations to suggestions. However, the overall findings and the suggested responses apply to all endeavors for advancing human knowledge. The common ground of visualization and the likely shared benefit among the interests represented by the separate Federal agencies suggests that a cross-disciplinary collaboration across the government may provide an effective means of addressing these needs.

5.1 Principal Agency Leadership Recommendation

NSF and NIH must make coordinated investments in visualization to address the 21st century's most important problems, which are predominantly collaborative, crossing disciplines, agencies, and sectors. Both NSF and NIH can and should provide leadership to other Federal funding partners, as well as to the research communities they support, by modifying their programmatic practices to better engage visualization capabilities across disciplines important to scientific and social progress and to encourage and reward interdisciplinary research, open practices, and reproducibility in technical and scientific developments. Such agency leadership is critical if we are to meet the needs of our nation in critical areas and to maintain U.S. competitiveness in a global environment.

5.2 Short Term: A Question of Policy

The extension of the visualization effort to close the research loop and better engage basic and application research may be advanced in the short term through changes in review policy with only limited new investment in sponsored research. We can achieve this by structuring grants in the basic sciences to reward those who include more visualization capabilities in their programs and who engage in more collaborative research with visualization researchers.

Short-Term Policy Recommendation: Policy changes for both funding and publication review to encourage evaluation of visualization and collaboration between visualization and other fields can be implemented immediately, without requiring new funding initiatives. Within visualization, review protocols should reflect the importance of evaluation to determine the success and characterize the suitability of techniques. Methodological rigor should be expected when user studies are proposed or reviewed, and as necessary visualization researchers should collaborate with those trained in fields such as human-computer interaction, psychology, and statistics. Peer review of proposals and publications should also reward visualization driven by real-world data and tasks, in close collaboration with target users. In other fields, review protocols should encourage domain scientists who create partnerships with visualization researchers, just as engagement with statisticians is considered normal practice in many areas such as biomedical research.

5.3 Mid Term: A Question of Direction

In achieving reproducible results, digital research domains have significant advantages over other areas. The data can be easily disseminated and experiments or analysis can be recreated in geographically distributed locations. Given open access to earlier methods, new methods can be built on the foundations of previous work, providing sufficient transparency. Visualization research can play a key role in bringing these advantages to other sciences and specialties.

Transparency and the advantages of the digital communication of information should be extended more widely into more domains supported by Federal investment - cross disciplinary efforts that help to bring together doctors, economists, scientists, meteorologists, and other experts with visualization researchers to enhance and enable open science. We advocate new mid-term funding for moderate efforts to increase the penetration of visualization, augmented reality, interactive and collaborative displays, and visual communication in the support of critical areas such as computer assisted surgery, interactive modeling of environmental and biological sciences, exploration of large astrophysical data, cataloging and investigation of genomic, financial, or demographic information, as well as other essential topics of national concern.

Mid-Term Direction Recommendation: Pilot programs should be established to combine efforts and create collaborative development between visualization and other research domains. Funding for such programs should contribute proportionately to both the visualization research and the domain specialty. The purpose of this effort will be to improve the penetration of emerging technologies into new domains, increasing their facility to move data and share results through visualization. All of the awards in this area should be dedicated to open access of source code, availability of research data to the worldwide community, and reproducibility of the technical and scientific developments.

5.4 Long Term: A Question of Investment

Meeting the future needs of our nation in critical domains affecting science and society will require the support of new foundation funds for exploratory research that moves beyond the simple acceleration or scaling of solutions, reaching for new and innovative methods rather the mere refinement of existing techniques. We must cultivate new technologies in advanced displays, portable and augmented data visualization, and data communication, exploration, and interaction. To achieve this goal we must establish a national infrastructure of data repositories, validation centers, technology development programs, research initiatives for interaction, abstraction, modeling and portrayal of complex information. We recognize that the NIH Roadmap initiatives have begun to address these and related problems, and we suggest that this effort be expanded and emulated in other arenas.

The Internet emerged as an international phenomenon and economic driver only after over twenty years of Federally funded research and development. Similarly, developing and validating realistic computational science software systems has required multiple cycles of development, computational experimentation, and analysis spanning multiple decades. Developing leading-edge computational visualization applications is a complex process that often involves multiple people with different backgrounds, and the effort often must be sustained for several years to yield the full fruits of investment. Some examples of successful long-term visualization projects that have successfully supported multiple national and international research programs include VTK, ITK, and SCIRun.

The usual three- to five-year length of a grant is often long enough only to explore a few new ideas and perform a proof of concept. Too often, project lifetimes do not extend past these initial efforts and we are left with one-off demos with no discernible reusable or extendible representations of the research. Unfortunately, software development, unlike the development of hardware and scientific instruments, has not succeeded in procuring the long-term investment needed to create effective tools. Federal research agencies usually stop short of providing needed support to extend the initial research ideas into usable software that can be beneficially shared with both the research and industry communities to drive further research and provide economic leverage. The ITK system discussed above is a welcome exception. Furthermore, there does not exist long term support for either basic visualization research or development of novel visualization software and hardware systems. Current Federal agency support tends to come in the form of a series of one-off programs (NSF Large-Scale Data Visualization, NSF ITR, NSF KDD, and the NIH R21/R33 Phased Innovation awards) that energize the community but fail to provide continued support to reap the rewards of the special program. Only sustained, coordinated investment in people, software, hardware, and data, based on strategic planning, will enable the U.S. to realize the promise of visualization to revolutionize scientific discovery and increase economic competitiveness.

Long-term Investment Recommendation: We recommend a coordinated and sustained national investment be made in a spectrum of centralized and distributed research programs to promote foundational, transitional, and applied visualization research in support of science, medicine, business, and other socially important concerns. This investment is critical for the U.S. to remain competitive in a global research and development community that has increasing resources. In addition to funding transitional research, such programs should emphasize foundational research and integration of methodologies from other fields, and collaboration with domain specialists who provide driving problems in areas of national concern. A long-term funding commitment is required for the creation and maintenance of curated data collections, and open-source software, to promote open science. Characterizing how and why visualizations work, systematically exploring the design space of visual representations, developing new interaction approaches, and exploiting the possibilities of novel display hardware will be particularly important areas of emphasis.

STATE OF THE FIELD 6

Many of the grand challenges set before the visualization community have been engaged and substantial progress has been achieved. We credit these advances to the industriousness of the researchers and the support of the National Science Foundation and other Federal funding agencies. The earlier sections of this report have presented a vision for how visualization can accelerate discovery and progress, but developments in visualization have already had a profound impact throughout the sciences and beyond. Throughout this report we are presenting some successes of visualization research and application as side bars. In this chapter we discuss other related reports, the current state of the national infrastructure for visualization, and past funding patterns.

6.1 Other Reports

There have been several previous national reports on the state of visualization and the need for significant investment in the creation and development of visually-based knowledge discovery techniques. The widely cited 1987 NSF Visualization report[28] is regarded by many as marking the birth of modern computer-supported visualization as a field, and certainly had a strong impact on funding priorities. That report noted that "Significantly more complexity can be comprehended through Visualization in Scientific Computing techniques than through classical ones" or through the "gigabit bandwidth of the eye/visual cortex system". The panel recommended a new initiative to get visualization tools into "the hands and minds" of scientists and noted the need for visualization researchers to team up with scientists and engineers to solve problems.

A more recent report sponsored by DOE and NSF focused on data manipulation and visualization of large-scale datasets[45]. This report, at least in part, resulted in the DOE Advanced Simulation and Computer (ASCI) Visual Interactive Environment for Weapons Simulation (VIEWS) program, which has significantly advanced visualization tool development and large dataset visualization methodology.

The National Visualization and Analytics Center (NVAC), sponsored by the Department of Homeland Security (DHS), has produced a major new book-length report defining the area of *visual analytics*: the science of analytical reasoning facilitated by interactive visual interfaces[48]. Visual analytics has a strong overlap with visualization, and we strongly endorse their findings. This report is complementary to theirs; their driving application area is national security, whereas we discuss the broad spectrum of application domains that can benefit from visualization in health, science, and engineering.

Many of the findings and recommendations in this report echo those of past reports. The 1999 Data Visualization Workshop at the University of Memphis[27], sponsored by the NSF and the Office of Naval Research (ONR), also argued strongly for the need for curated data and characterized tasks; the need for taxonomies and general principles to guide visualization design; and the need for visualization practitioners to collaborate with cognition and perception researchers. Another issue discussed in this report is the tension between application-specific and general-purpose visualization design.

The PITAC report on computational science[37], emphasizing the importance of long-term multidisciplinary and multi-agency efforts, cautions that

> despite the great opportunities and needs, universities and the Federal government have not effectively recognized the strategic significance of computational science in either their organizational structures or their research and educational planning. These inadequacies compromise U.S. scientific leadership, economic competitiveness, and national security.

Many past reports document the explosion of information that must be handled in science[29], computational science[19], information technology[18], and general society[26]. In addition to the reports, there have been several articles that discuss visualization research challenges[9, 17, 23, 30, 40, 41].

6.2 National Infrastructure

One of the key emphases of the 1987 NSF Report was the need for a national infrastructure to enable visualization research and application. Many of the specific needs discussed have been satisfied in the intervening years, but others have remained a challenge. This section summarizes the current state of hardware, networking, and software support for visualization.

Annotating Reality

An important problem in the automated design of visualizations is how to label and annotate real and virtual objects effectively. This can be especially challenging when the objects to be annotated and the viewer all reside in a dynamic 3D world. Naïve approaches can result in ambiguous labels and important objects being obscured by less important ones. This image shows the output of an interactive annotation system, photographed through a see-through head-worn display from the perspective of one user in a test-bed collaborative augmented reality environment. Two users are sitting across from each other discussing a virtual campus model located between them. All campus buildings have been labeled with their names; each label is scaled within a user-selectable range and positioned automatically. A label is placed either directly within a visible portion of its building's projection, or if the visible parts of the projection are deemed too small to accommodate a legible label, the label is placed near the building and connected to it with an arrow, while avoiding overlap with other objects. Additional annotations include a meeting agenda (left), and a building model and information sheet (right). All annotations have been constrained to avoid other objects, including the visible user's head, to allow a direct line of sight between the two users. The result is that annotations remain legible and clearly associated with the objects to which they are related, even as the users move. Challenges include supporting a richer range of spatial constraints, better addressing graphic design considerations, and developing systems that can choose constraints to fulfill on the fly as users' tasks change.

B. Bell, S. Feiner, and T. Höllerer, View management for virtual and augmented reality, *Proceedings of UIST 2001* (*ACM Symposium on User Interface Software and Technology*), Orlando, FL, November 11–14, 2001 (*CHI Letters*, vol. 3, no. 2), 101–110.

6.2.1 Visualization Hardware

Many of the hardware concerns from the original NSF report have been allayed by the passage of time and Moore's Law. Processors with what used to be considered supercomputer-class power are now available in commodity desktop PCs that cost a few thousand dollars. Graphics performance that used to require special-purpose workstations costing tens or hundreds of thousands of dollars is now available as a commodity graphics card for desktop PCs that cost a few hundred dollars. The good news is that fast and cheap hardware aimed at the business and entertainment mass markets allows unprecedented access to computational and graphics power for visualization, a boon for both visualization researchers and end users. The flexibility of the latest generation of programmable graphics pipelines on these cards has sparked an explosion of sophisticated rendering techniques that are feasible in real time for the first time[34], which also benefits visualization users by providing real-time interaction when exploring large datasets.

Great advances have also been made in visualization-specific hardware. The VolumePro card for hardware-accelerated volume rendering is a major technology transfer success story. Volume rendering is extremely computationally intensive and had been considered a clear candidate for hardware support for many years. A great deal of academic and industrial research in software volume rendering algorithms brought the field to maturity and finally made hardware creation feasible. Grant-funded university research that began at the State University of New York (SUNY)[36] led to the development of an actual product through the industrial lab Mitsubishi Electric Research Lab (MERL)[35], culminating the successful spin-off company TeraRecon.

In contrast, display technology improvements have historically lagged far behind the Moore's Law curve. In the past 20 years, cathode ray tube (CRT) displays have little more than doubled in physical display size and resolution and have retained the same weight and form factor. In the past, our ability to design user interfaces has been constrained by fact that a monitor is a relatively heavy and expensive object. However, recent breakthroughs in flatpanel and projector technology have broken the strangle-hold of the CRT. The combination of low cost, high resolution, and freedom from the weight and bulk constraints of CRTs will lead to an explosion of computer-driven displays in many new contexts, far beyond simply replacing the bulky CRT on a user's desk with a sleek flat panel display that has a smaller footprint[39].

Pixels are currently a scarce resource. The primary limitation in interactive visualization interfaces is the number of available pixels: we are pixel-bound, not CPU-bound or even render-bound. High-resolution displays will allow us to investigate new and exciting parts of the interface design space as displays approach the resolution of paper. Large wall-sized displays with a resolution of dozens or even hundreds of megapixels can be created by tiling the output of many projectors. Although active surfaces will still be relatively expensive in the near term, a longer-term vision is that gigapixel displays will eventually be as cheap, lightweight, and ubiquitous as wallpaper. Physically large displays that encompass the entire field of view of an observer allow applications that use peripheral vision as well as the foveal vision that we use with medium-sized desktop displays. Small gadget displays will have the one megapixel resolution that we currently associate with desktop displays. Small handhelds have high availability because they can be carried around, and when networked can be used as control panels for a shared large display.

Finding: Fast and cheap commodity hardware meets most of the CPU and graphics needs of visualization today. The current availability of commercial volume rendering hardware is a success story for the field. New advances in display hardware will have a major impact on visualization.

6.2.2 Networking

As of 2005, we have reaped vast benefits from the expansion and commoditization of the Internet. However, as the sizes of data sets continue to grow, it is still difficult, and sometimes prohibitive, to move large-scale data sets across even the fastest networks to visualize data locally. As such, there is a need for advances both in networking and in remote and collaborative visualization algorithms, such as view dependent algorithms, image based rendering, multiresolution techniques, importance based methods, and adaptive resource aware algorithms.

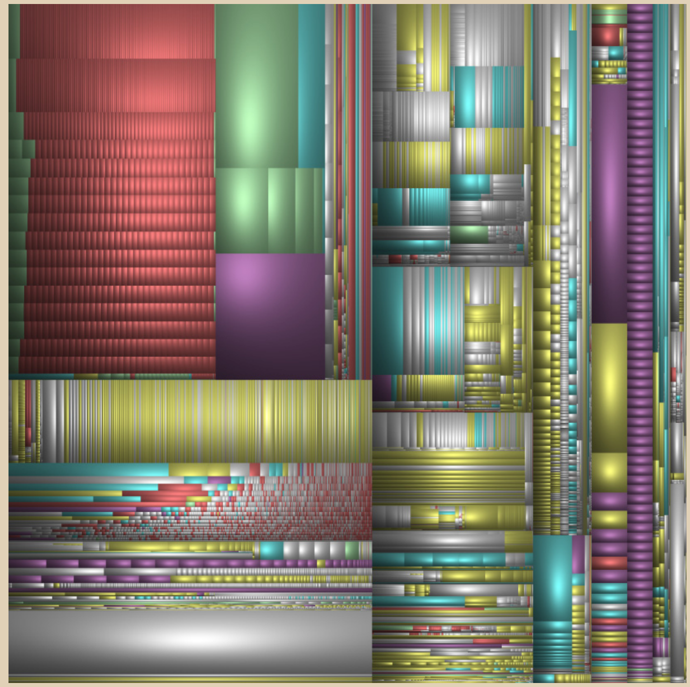
Finding: Advances in networking and network-aware visualization algorithms could greatly facilitate collaboration and other remote visualization opportunities.

6.2.3 Visualization Software

The 1989 introduction of the AVS dataflow toolkit[49] heralded the first generation of general-purpose software for visualization. Other systems of that generation include IBM's DataExplorer[25], now the open-source OpenDX system; IRIS Explorer from SGI and then NAG[15]; and the University of Wisconsin Vis5D/VisAD systems[20]. The open-source VTK[42] system stands out as the most widely used of the next generation of systems. Others currently in use include ParaView[4], Amira[46], and the InfoVis Toolkit[12], with the continuing presence of OpenDX[25] and AVS[49]. Many packages that focus on application-specific needs have been developed, including Ensight[10], Fieldview[22], SCIRun[32], and ITK[53].

The movement known as open source is the current incarnation of an idea that has been active for decades in the academic community, namely that there is great value in providing free

Resource Allocation



Why is my disk full? This question haunts many of us. SequoiaView is a tool that provides visual answers. It uses the treemap method enhanced with shaded cushions. The image is subdivided into rectangles, where each rectangle represents a folder or, on the lowest level, a file. The hierarchical cushions help to show the hierarchical structure. The color and size of the smallest rectangles show the type and size of the file. For instance, the red rectangles in the upper left corner represent image files, the big purple rectangle a large archive, the area with small grey and red rectangles is the cache of an internet browser. Guided by the images, large files or large collections of files can be found, and the user can clean up his disk efficiently and easily. SequoiaView has been downloaded more than 450,000 times and distributed on many CD's. Users have reported that its use saved them from buying a new hard disk. Also, it has led to a spin-off company, called MagnaView. The challenge is to extend these methods for the presentation of tabular data, ubiquitous in business and many other branches of our daily lives, such that viewers can, for instance, understand which products perform well or not for which reason, or to see which factors influence the scores of high school students.

J.J. van Wijk, F. van Ham, and H.M.M. van de Wetering, Rendering hierarchical data, *Comm. ACM*, Vol. 46, No. 9, pp. 257-263, September 2003.

software. One of the new aspects of the movement is formalizing the value of open source for industry as a business model. We note that there is bidirectional technical transfer with open-source software. In some cases, open-source government-funded research prototypes later evolve into commercial products. In other cases, commercial projects change to open source because the business model is more appealing.

There is a tradeoff between quickly creating a one-off prototype that suffices for a research paper but is too brittle to be used by anybody but the authors and devoting the time to create releasable code at the expense of making progress on the next research project. One of the benefits for researchers of releasing code to a user community is that real-world use typically spawns new research challenges strongly tied to real problems. Such ties are extremely important for our field as it matures. One benefit of the open-source model is that the user community itself sometimes takes over some or all of the support burden. Releasing software does not have to be a gargantuan task; often people who find that a particular piece of research software closely matches their needs are happy to use software that is less polished than a commercial product.

The VTK system[42] began as an open-source initiative within the General Electric Global Research division, and has rapidly moved into mainstream use in universities, national laboratories, and industrial research labs worldwide. It continues to accelerate development by providing reusable software, relieving programmers from reinventing necessary infrastructure. The spinoff company Kitware is built around an open-source business model, where customers can pay for support and customization while development of the free codebase continues. Similarly, ITK was an NIH open-source software initiative intended to support a worldwide community in image processing and data analysis. It is designed to interface openly with visualization platforms such as VTK and SCIRun. The University of Utah's SCIRun visualization system has also made the move to open-source infrastructure software to ease its integration into public and private research.

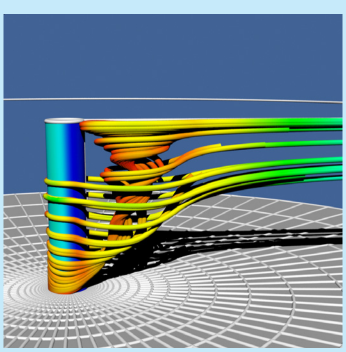
Finding: Both commercial and open-source visualization software systems are thriving as the field matures. The open source model offers many benefits to both academia and industry, and to both researchers and end users.

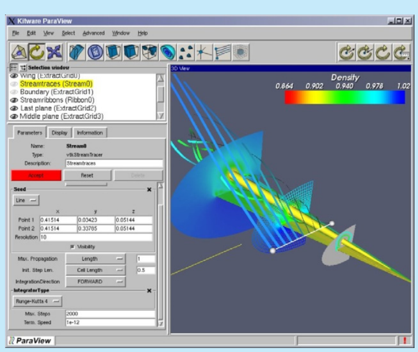
The Visualization Toolkit

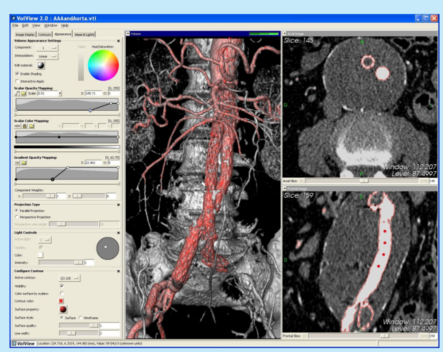
In 1993 three visualization researchers from GE Corporate R&D Center began to develop an open source visualization system. This system, which came to be known as the Visualization Toolkit (VTK), was initially envisioned as a teaching and research collaboration tool (hence its release under open source license). The software gained rapid acceptance, in part due to the sophistication of its object-oriented design and software process, but also because of the community of users that formed around it. VTK is now in world-wide usage, and has helped spawn several small companies and derivative products. For example, Kitware Inc. was formed in 1998 to support VTK, subsequently creating products based on the toolkit including the open source ParaView parallel visualization system and the proprietary volume rendering application VolView. VTK continues to evolve with contributions from researchers in academia, the US National Labs, and businesses, and is used in dozens of commercial software applications. The left figure uses the LOx Post dataset, which simulates the flow of liquid oxygen across a flat plate with a cylindrical post perpendicular to the flow. This analysis models the flow in a rocket engine, where the post promotes mixing of the liquid oxygen. The middle figure demonstrates CFD visualization using ParaView (middle), while the right figure demonstrates volume rendering using VolView.

W.J. Schroeder, K. Martin, and W. Lorensen, *The Visualization Toolkit: An Object Oriented Approach to Computer Graphics, Third Edition*, Kitware, Inc., ISBN-1-930934-12-2 (2004).

S. E. Rogers, D. Kwak, and U. K. Kaul, A numerical study of three-dimensional incompressible flow around multiple post. In *Proceedings of AIAA Aerospace Sciences Conference*. AIAA Paper 86-0353. Reno, Nevada, 1986.





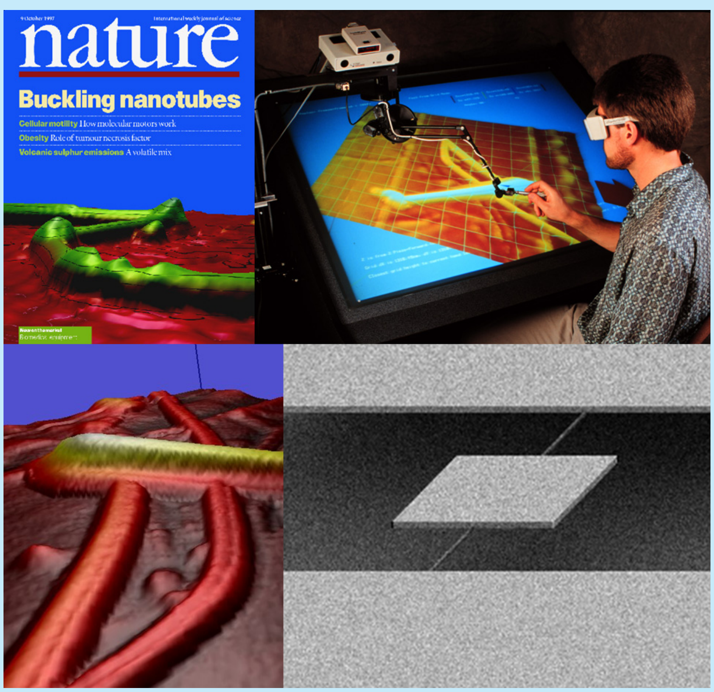


Funding for research in visualization has come from a number of sources. Looking at funding acknowledgments in papers appearing in the IEEE Visualization Conference during the period 1998-2004, approximately 34% of papers cite support from NSF (virtually all from the CISE directorate), 19% from non-U.S. governments, 18% from industry, 14% from DOE, 14% from U.S. military sources (including NRO, ARO, DARPA, and ARDA), 8% from NASA, 7% from NIH, and 5% from other sources (including other U.S. government agencies and private foundations). Approximately 30% of papers have no acknowledgment of financial support. Papers appearing in the IEEE Information Visualization Symposium for that same period have a higher percentage of support coming from industry. Specifically, approximately 23% of papers cite support from industry, 12% from non-U.S. governments, 10% from NSF (virtually all from CISE), 9% from U.S. military sources, 6% from DOE, 3% from NIH, 1% from NASA, and 3% from other sources. Approximately 41% of papers have no acknowledgment of financial support. In both cases, industry figures include authors employed by industry, even if no explicit acknowledgment of support is given.

We note with particular concern the recent article by ACM President David Patterson[33] that documents the decline in both industrial and military funding for basic research. Moreover, an increasingly low percentage of NSF proposals are being funded, and there appears to be a bias in proposal reviewing towards low-risk incremental work rather than attacking grand challenges.

One important exception to this downturn is the new NVAC initiative, which will be spending several hundred million dollars over the next five years on visual analytics. Much of this funding will be focused on the domain of national security. While a significant amount of fundamental research will arise from this short-term program, NIH and NSF must ensure that the domain areas of health and science are sufficiently funded, and that long-term research continues to put the field on a solid scientific foundation.

Finding: Visualization research is not being funded at a level that enables continued discovery. The distribution of funding sources does not really reflect the potential benefits of visualization research to specific application areas.



Nanoscale Science

Visualization played a large part in a series of carbon nanotube studies. The ability to rapidly explore hypotheses with immediate visual analysis of results led to fundamental new understanding in nanoscale bending and buckling and to the demonstration of atoms acting as gear teeth, atomic-lattice interlocking controlling how electrons flow between nanoscale parts, and nanoscale torsional coupling. Coupling the visualization into a direct-manipulation control system allows the performance of pilot experiments in minutes that used to take days.

The nanoManipulator system shown at the upper right enables one to directly see and touch nanometer-scale objects and get real-time qualitative analysis coupled to offline quantitative analysis. This tool and others like it have enabled rapid progress in biomedical and materials science.

This collaborative effort led to fifteen publications in physics on top of those in visualization. The multidisciplinary author lists indicate the level of intellectual involvement of the entire team.

S. Paulson, A. Helser, M. Buongiorno Nardelli, R.M. Taylor II, M. Falvo, R. Superfine, and S. Washburn, “Tunable resistance of a carbon nanotube-graphite interface,” *Science*, Vol. 290, pp. 1742-1744, 2000.
“Nanometre-scale rolling and sliding of carbon nanotubes,” *Nature*, Vol. 397, No. 6716, pp. 236-238, 1999.
“Bending and buckling of carbon nanotubes under large strain,” *Nature*, Vol. 389, No. 6651, pp. 582-584, October 9, 1997.

Bibliography

[1] COMPLETE – the COordinated Molecular Probe Line Extinction Thermal Emission survey of star forming regions. http://cfa-www.harvard.edu/COMPLETE/index.html.

[2] Surgical Planning Lab, Brigham and Women's Hospital. http://splweb.bwh.harvard.edu:8000/pages/aboutspl/about.html.

[3] Michael J. Ackerman. The visible human project. *Proceeedings of the IEEE*, Vol. 86, No. 3, pp. 504–511, 1998.

[4] Jim Ahrens, Berk Geveci, and Charles Law. Paraview: An end-user tool for large-data visualization. In Charles D. Hansen and Chris R. Johnson, editors, *The Visualization Handbook*, pp. 717–731. Elsevier, 2004.

[5] Gordon Bancroft, Fergus Merritt, Todd Plessel, Paul Kelaita, Robert McCabe, and Al Globus. FAST: A multi-processed environment for visualization of computational fluid dynamics, AIAA paper 91-0793. In *29th Aerospace Sciences Meeting, Reno, NV*, January 1991.

[6] Jacques Bertin. *The Semiology of Graphics*. University of Wisconsin Press, 1983. (First edition 1967).

[7] Michelle A. Borkin, Naomi A. Ridge, Alyssa A. Goodman, and Michael Halle. Demonstration of the applicability of "3D Slicer" to astronomical data using 13CO and C18O observations of ic348. Master's thesis, Harvard University, May 2005.

[8] Stuart K. Card and Jock Mackinlay. The structure of the information visualization design space. In Proc. *IEEE Symposium on Information Visualization*, pp. 92–99, 1997.

[9] Chaomei Chen. Top 10 unsolved information visualization problems, *IEEE Computer Graphics and Applications*, Vol. 25, No. 4, pp. 12-16, July/August, 2005.

[10] Computational Engineering International (CEI). Ensight. http://www.ensight.com.

[11] William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, Vol. 79, No. 387, pp. 531–554, September 1984.

[12] Jean-Daniel Fekete, The InfoVis Toolkit, Proc. IEEE Symposium on Information Visualization (InfoVis) 2004, pages 167-174.

[13] Jean-Daniel Fekete, Catherine Plaisant, and Georges Grinstein. IEEE InfoVis 2004 Contest: The History of InfoVis, 2004. http://www.cs.umd.edu/hcil/iv04contest/.

[14] Ian Foster and Carl Kesselm. *The Grid 2: Blueprint for a New Computing Infrasturcture*. Morgan Kaufmann, 2003.

[15] David Foulser. IRIS Explorer: A framework for investigation. *Computer Graphics*, Vol. 29, No. 2, pp. 13–16, 1995.

[16] Mark Gahegan. *Exploring Geovisualization*, chapter Beyond Tools: Visual Support for the Entire Process of GIScience, pp. 83–99. Elsevier Science, 2005.

[17] Bill Hibbard. The top five problems that motivated my work, *IEEE Computer Graphics and Applications*, Vol. 24, No. 6, pp. 9-13, November/December, 2004.

[18] Information Science Technologies Advisory Group. ISTAG report on grand challenges in the evolution of the information society. Technical report, European Commission, September 2004.

[19] Bernd Hamann, E. Wes Bethel, Horst Simon, and Juan Meza. NERSC Visualization Greenbook: Future visualization needs of the DOE computational science community hosted at NERSC. Technical report, Lawrence Berkeley National Laboratory, 2002.

[20] William L. Hibbard, Brian E. Paul, David A. Santek, Charles R. Dyer, Andre L. Battaiola, and Marie-Francoise Voidrot-Martinez. Interactive visualization of earth and space science computations. *IEEE Computer*, Vol. 27, No. 7, pp. 65–72, 1994.

[21] Jim Hollan and Pat Hanrahan. Visualizing information. DARPA/DOD Information Science and Technology (ISAT) Study, 1997.

[22] Intelligent Light. Fieldview. http://www.ilight.com.

[23] Chris R. Johnson. Top scientific visualization research problems, In *IEEE Computer Graphics and Applications: Visualization Viewpoints*, Vol. 24, No. 4, pp. 13-17. July/August, 2004.

[24] Rich Kostro. Making earth observations that matter. *Earth Imaging Journal*, Vol. 2, No. 4, page 8, July – August 2005.

[25] Bruce Lucas, Gregory D. Adams, Nancy S. Collins, David A. Epstein, Donna L. Gresh, and Kevin P. McAuliffe. An architecture for a scientific visualization system. In *Proc. IEEE Visualization*, pp. 107–114, 1992.

[26] Peter Lyman and Hal R. Varian. How much information, 2003. http://www.sims.berkeley.edu/how-much-info-2003.

[27] Jonathan I. Maletic and Priti Shah. Workshop on data visualization final report. Technical report, University of Memphis, 1999. Sponsored by the NSF and ONR.

[28] Bruce H. McCormick, Thomas A. DeFanti, and Maxine D. Brown. *Visualization in Scientific Computing*. National Science Foundation, 1987.

[29] Richard P. Mount. The Office of Science data-management challenge: Report from the DOE Office of Science data-management workshops. Technical report, March–May 2004.

[30] Gregory Nielson. Challenges in visualization research, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 2, No. 2, June 1996.

[31] Tim O'Reilly. Personal communication, 2005.

[32] Steve G. Parker and Christopher R. Johnson. SCIRun: A scientific programming environment for computational steering. In *Proc. Supercomputing*, pp. 2–19, 1995.

[33] David A. Patterson. The state of funding for new computer science initiatives in computer science and engineering. *Communications of the ACM*, Vol. 48, No. 4, page 21, 2005.

[34] Hanspeter Pfister. Hardware-accelerated volume rendering. In C.D. Hansen and C.R. Johnson, editors, *The Visualization Handbook*, pp. 229–258. Elsevier, 2005.

[35] Hanspeter Pfister, Jan Hardenbergh, Jim Knittel, Hugh Lauer, and Larry Seiler. The VolumePro real-time ray-casting system. In Proc. SIGGRAPH, pp. 251-260, 1999.

[36] Hanspeter Pfister and Arie Kaufman. Cube-4 - A scalable architecture for real-time volume rendering. In *ACM / IEEE Symposium on Volume Visualization*, pp. 47–54, San Francisco, CA, October 1996.

[37] Dan Reed (Chair) PITAC Subcommittee on Computational Science. President's Information Technology Advisory Commitee: Report on computational science, 2005. http://www.itrd.gov/pitac.

[38] Catherine Plaisant, Richard Mushlin, Aaron Snyder, Jia Li, Dan Heller, and Ben Shneiderman. Lifelines: Using visualization to enhance navigation and analysis of patient records. In *American Medical Informatic Association Annual Fall Symposium (Orlando, Nov. 9-11, 1998) AMIA, Bethesda MD*, pp. 76–80, 1998.

[39] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proc. SIGGRAPH*, pp. 179–188, 1998.

[40] Lawrence J. Rosenblum et al. (eds). Scientific visualization: Advances and challenges, Academic Press Ltd., London, UK, 1994.

[41] Lawrence J. Rosenblum. Research issues in scientific visualization, *IEEE Computer Graphics and Applications*, Vol. 14, No. 2, pp. 61-85, March 1994.

[42] Will Schroeder, Ken Martin, and Bill Lorensen. *The Visualization Toolkit: An Object-Oriented Approach to Computer Graphics (3rd Ed.)*. Kitware, Inc., 2003.

[43] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualization. In *Proc. IEEE Visual Languages*, pp. 336–343, 1996.

[44] Herbert Simon. Designing organizations for an information-rich world. In Martin Greenberg, editor, *Computers, Communications, and the Public Interest*, pp. 40–41. Johns Hopkins Press, 1971.

[45] Paul H. Smith and John van Rosendale. Data and visualization corridors. Technical Report CACR-164, California Institute of Technology, 1998.

[46] Detlev Stalling, Malte Westerhoff, and Hans-Christian Hege. Amira: A highly interactive system for visual data analysis. In Charles D. Hansen and Chris R. Johnson, editors, *The Visualization Handbook*, pp. 749–767. Elsevier, 2004.

[47] A. Szalay, J. Gray, P. Kunszt, and A. Thakar. Designing and mining multi-terabyte astronomy archives: The Sloan digital sky survey. In *Proc. ACM SIGMOD*, pp. 451–462, 2000.

[48] James J. Thomas and Kristin A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.

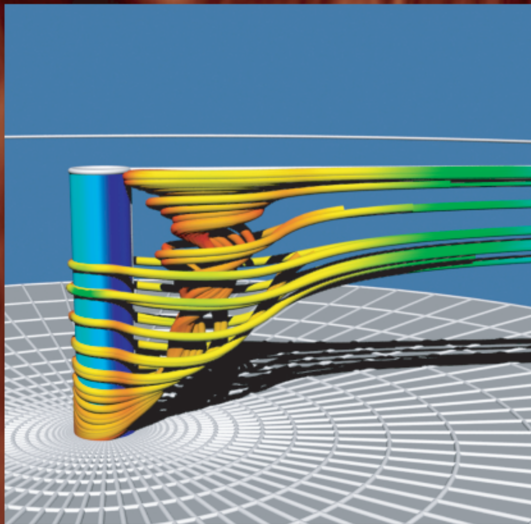
[49] Craig Upson, Thomas Faulhaber Jr., David Kamins, David Laidlaw, David Schlegel, Jeffrey Vroom, Robert Gurwitz, and Andries van Dam. The Application Visualization System: a computational environment for scientific visualization. *IEEE Computer Graphics and Applications*, Vol. 9, No. 4, pp. 30–42, 1989.

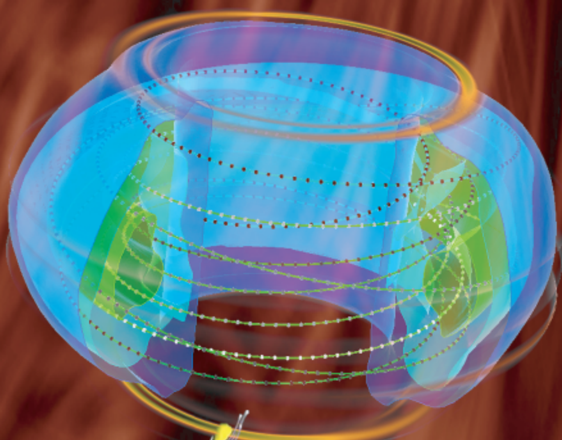
[50] Jarke van Wijk. The value of visualization. In *Proc. IEEE Visualization*, 2005.

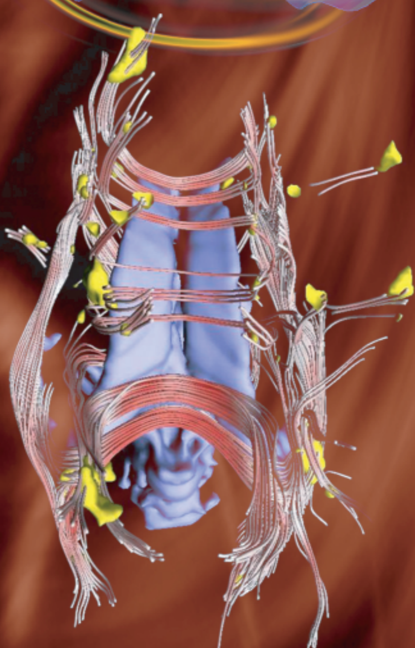
[51] Wolfgang von Rüden. CERN and grid computing, or extreme science. Vanguard Meeting on Extreme Interfaces, September 2005.

[52] Leland Wilkinson. *The Grammar of Graphics*. Springer-Verlag, 1999.

[53] Terry S. Yoo. The Insight Toolkit: An open-source initiative in data segmentation and registration. In C.D. Hansen and C.R. Johnson, editors, *The Visualization Handbook*, pp. 733–748. Elsevier, 2005.







NATIONAL INSTITUTES OF HEALTH

NSF

MITSUBISHI ELECTRIC
Mitsubishi Electric Research Laboratories

IEEE COMPUTER SOCIETY

IEEE

SCI INSTITUTE

IEEE vgtc