



MEMO

August 20, 2008

To: Bruce Hendrickson, Mark Foehse, Suzanne Rountree, Bill Cook
From: Mark Huey and Ann Miksovic
Subject: NGC Tracking Report

The Networks Grand Challenge (NGC) has retained Perspectives to track pertinent developments in the marketplace. Perspectives performed an initial market survey in early 2008. That effort made a number of things clear:

- There is significant activity – from industry, from academia, and from government agencies – that should be of interest to NGC.
- There is too much activity for a comprehensive ongoing tracking effort.

Perspectives has put together this document as a starting point to guide future efforts in tracking. We believe that a bimonthly (i.e., every other month) series of slim reports (20-30 page each) is the right general approach and should be very valuable to the NGC team as it works in FY09 to refine its differentiation and develop its long-term business model.

Each issue will have two parts:

- Identify key developments of interest to NGC around a core set of priority topics.
- Provide more in-depth information on a “special topic.”

The reports will focus on collecting capsules of information (most items will be less than a paragraph; links back to the original information source are always provided for the interested reader). Developments to be tracked on a systematic / ongoing basis:

- News about network analysis applied to intelligence analysis and terrorism, nonproliferation, cyber security, and the big data problem.
- Funding opportunities (BAAs).
- News about some of the more relevant / interesting commercial vendors.

“Special topics” that we believe may be of value to NGC:

- Funding by US government agencies (such as NSF).
- International funding (e.g., Europe, Japan).
- International research activities (e.g., what is happening in the UK, Japan, or Germany that appears similar to NGC).
- Patents and patent applications of relevance to NGC, and analysis of the collected set of intellectual property.
- Identification of data sets being use for network analysis.
- Profiles of key cybersecurity software vendors and reviews of product capabilities.
- Profiles of key text-mining software vendors and reviews of product capabilities.
- Network analysis for other applications (this may actually be a number of special reports, e.g., reports on epidemiology, social networks, business applications, or other areas).
- Scientific literature reviews of specific topics (e.g., novel approaches to network visualization).
- Chatter from the blogging community on network analysis.
- Other topics, as determined by discussion with the NGC team.

This inaugural “tracker” issue is longer than what we anticipate for ensuing issues (in part because it covers nearly four months of information). Subsequent issues would be in the range of 20 pages. We would also like to explore alternative means of getting the relevant content out to the broader NGC team (e.g., by publishing information from the reports onto the NGC wiki).

We look forward to working with NGC management to better understand where the effort can be focused for maximum value to the project.



This issue of the Networks News and Science Tracking Report from Perspectives covers material primarily from late April through early August 2008, although some important material from other periods is included. For example, older material uncovered as part of the tracking research may be included if it has not been discussed in previous reports.

This document contains items abstracted (edited, direct text from or summaries of source material) from the news or other sources. Occasionally, Perspectives' comments are included (indicated by italic type). Emphasis is ours and is indicated by bolding or underlining. Items of particular interest to us are indicated by a highlighted star.

The full text of any source item is available. Contact Ann Miksovic: ann@perspectivesweb.com or (505) 881-0370.

Table of Contents

I. Priority Applications	4
A. Terrorism / Intelligence Analysis.....	4
B. Cybersecurity.....	5
C. The Big Data Problem	10
D. Funding Opportunities	11
II. Other Applications and Items of Interest	14
A. Medical / Life Sciences / Pharma	14
B. Business Analytics.....	15
C. Other Items of Interest.....	16
III. Company Activities	17
A. In-Depth	17
B. Company News in Brief	19
IV. Resources / Overviews	23
A. Major Publications	23
B. Reviews and Overview Articles	24
C. Resources on the Web	24
D. Conferences	25

I. Priority Applications

A. Terrorism / Intelligence Analysis

- **University of Arizona – Dark Web:** for the **University of Arizona's AI Lab**, the "Dark Web" refers to a research project where the social phenomena of terrorism is studied via various techniques including social network analysis, content analysis, link analysis, web metrics, video analysis, data and text mining, sentiment and affect analysis, and authorship analysis. Through the use of sophisticated, mathematical tools, the project aims to collect all web content generated by international terrorist groups, including content found on web sites, forums, chat rooms, blogs, social networking sites, videos, virtual worlds, and more.

Federally funded through the National Science Foundation [most recent [award](#) for \$500k], the Dark Web's spiders have been crawling through the web for the past five years. As of 2007, they estimated there were about 50,000 sites of extremist / terrorist content when they looked beyond just traditional web pages. This collection is 2 TBs in size making it the largest open-source extremist / terrorist collection in the academic world. Researchers who would like to use this data in their own studies can contact the research center for access.

... [For more information, including summary statistics on content of the Dark Web, see "[Spidering the Dark Web](#)" and the [project website](#).]

- The **IEEE Intelligence and Security Informatics (ISI)** conference, held in June of this year in Taiwan, has many topics of interest. The full program is accessible from this [page](#). Keynote speaker abstracts are available from this [page](#). Keynotes include:

Hsinchun Chen (University of Arizona): "[Homeland Security Data Mining using Social Network Analysis](#)."

Bhavani Thuraisingham (UT-Dallas): "[Data Mining for Security Applications](#)."

Jaideep Srivastava (University of Minnesota): "[Data Mining for Social Network Analysis](#)."

Joydeep Ghosh (UT-Austin) "[Probabilistic Frameworks for Privacy-Aware Data Mining](#)."

Ajith Abraham (Norwegian University of Science and Technology): "[Real Time Intrusion Detection and Prevention Programs](#)."

Patrick Wang (Northeastern University): "[Intelligent Pattern Recognition and Applications](#)."

- [Intelligence director urges IT overhaul](#). The U.S. intelligence community must undergo fundamental cultural and organizational changes... according to the report, [Vision 2015](#), released in July by the **Office of the Director of National Intelligence**. "It is a networked world where what happens in Peshawar affects Peoria – and vice versa." ... Director Mike McConnell said that to succeed in this world, the intelligence community has to: develop capabilities to address cyber space challenges, put customer service first, draw on the capabilities of all the intelligence agencies for specific missions, develop a system that delivers information when and where it is needed, and integrate housekeeping systems such as budget and finance.
-  [Putting the Brain Back into Intelligence](#): As intelligence agencies fail to keep up with massive volumes of data and shrinking reaction times through automation – enter the human analyst. *Peter Buxbaum, a Washington-based independent journalist, wrote an article appearing on SperoNews.com. Some of the highlights from this article, which quotes a user and representatives from **Endeca** and **Centrifuge**:*
 - The **US National Geospatial-Intelligence Agency** expects to gather four petabytes of data annually in coming years. ... "We find a real sense of frustration across the

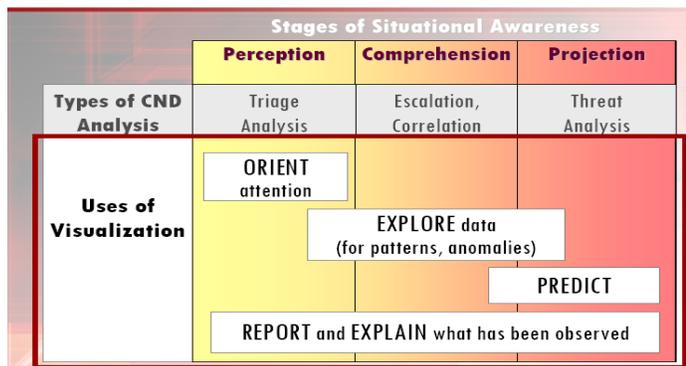
- intelligence community," Guljit Khurana, CEO of **Centrifuge Systems**, a technology company based in McLean, Virginia, told *ISN Security Watch*. "The volume and velocity of data is growing while the time frame in which to act on the data is shrinking."
- One of the applications used by the [US Army's Special Operations Command 95th Civil Affairs] Brigade, Analyst's Notebook by **i2 Technologies, Inc** in Dallas, provides an environment for creating information linkages and timeline analyses. ... Col. Ferdinand Irizarry [Commander of the Brigade] complained his unit has confronted a "dark underbelly of technology." "What we have found is that most of this [analyst] software is not all that user friendly," he said. "It takes a lot of training to teach how to use it. The software is flexible and can serve a lot of different requirements but it is inherently complex. It requires the operator to memorize many commands and be familiar with the full breadth hand depth of the software's functionality."
 - Technology companies are trying to solve these problems by "focusing on human cognition rather than technology. We want to commoditize information searching so that the user doesn't have to think about how it works any more than the average person has to think about how a refrigerator works. All you know is that it is keeping your food cold."
- **The Success of Failure**. This 2007 editorial reviews the cost of NSA's **Trailblazer** (\$1.2B +) and efforts to rescue the program (dubbed **Turbulence**, cost was \$500M), in a story initially revealed to the public by *The Baltimore Sun's* Siobhan Gorman.

B. Cybersecurity

- ★ A 2007 presentation, "Can Visualization Make a Difference in the Outcome of Cyber Conflict?" by **Anita D'Amico** (**Secure Decisions** division of **Applied Visions**) is available in full, online [here](#). She concludes that it can make a difference, but it is not yet viewed as essential. Visualization is more prolific in Computer Network Defense (CND) than any other aspects of cyber conflict, but it is not incorporated into most CND analytical functions. Viz won't get regularly used until it is incorporated into a visualization system aligned with the Concept of Operation. Cool visuals sell; workflow-aligned systems succeed. Visualization systems can provide spatial and temporal context for cyber conflict, beyond CND operations, e.g., surveillance, CNO planning, evaluation of tactics, impact assessment, warrior training.

*We found the entire presentation to be interesting and worthy of reading. A number of the slides from her presentation are shown below. D'Amico references several visualization products, including: **Inxight's** Star Tree, **MITRE** IWViz, **CROMDI/Utah** VisAlert, **PNNL** Starlight, **NeuralStar**, **Secure Decisions'** SecureScope and VIAssist, **ArcSite**, **i2**, and **orgnet.com**.*

Design Visualization Systems Based on How CND Analysts Will Use Visualization



Differences in Visualization System Needs Triage Analysis vs Correlation Analysis

Triage Analysis	Correlation Analysis
<i>Keep visuals simple and constant, to detect changes quickly</i>	<i>Foster ad hoc visual exploration, but keep track of path taken</i>
Visualize primary sensor data (e.g. IDS alerts)	Add secondary data (e.g. whois, CERT advisories, location, intel reports)
Simple 2D graphics	Multi-dimensional displays, often with temporal and spatial context
Use color, blinking and motion in uniform, pre-set conditions	Color, blinking and motion under user control
Distinguish/ fade out old data from new	Retain historic data for pattern discovery
Standard queries to data repository; standard visual filters	Complex, analyst-crafted queries and filter criteria
Automatically update data being visualized at regular intervals	Don't automatically update data in visualization unless user requests it

Uses of CND Visualization Predict and Explain

Forecasting Predict progress of events

- Examples: Trend lines as in SANS Internet Storm, animations of historical activity, animations of predicted events
- Room for growth here: current visualization systems don't support prediction well

Communication Explain findings to others

- Incorporate visualizations into "Big Board" displays, PowerPoint briefings, printed media, web portal access, or shared views of data
 - Examples: SecureScope report builder, VIAssist E-Diary
- Use visualization to explain complex findings or those that are difficult to verbalize

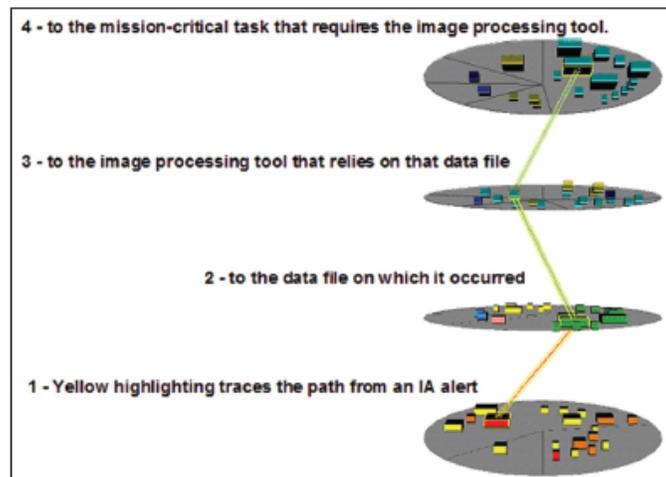
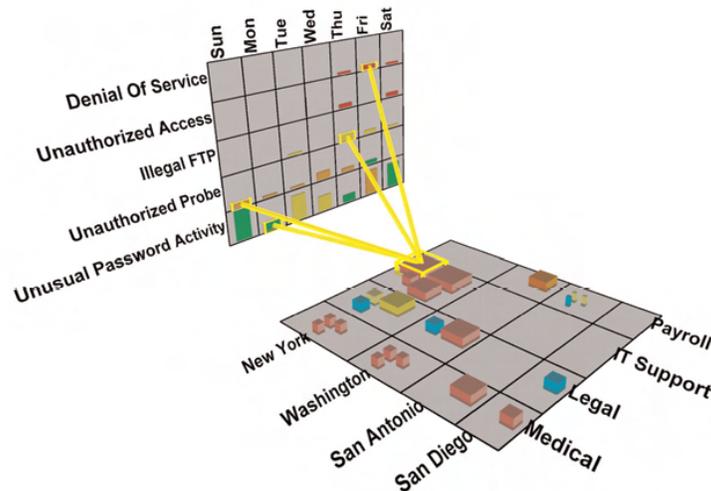
Spatial Context MITRE IWViz Relates Location to IDS Alerts



- More on Secure Decisions (the group that produced the preceding slides):** In May of 2007, the Secure Decisions division of Applied Visions, Inc. was awarded the first phase of an Air Force SBIR research contract (made by the AFRL Information Directorate under their solicitation entitled "Situation Awareness and Impact Assessment for Cyber Network Defense.") ... Secure Decisions will research and develop the **Cyber Asset-to-Mission Mapping Model (CAM3)**, a system to automatically relate cyber assets to higher-level network capabilities and missions dependent upon those assets. Secure Decisions will develop technology that automates the collection and structuring of cyber asset-to-mission data to facilitate Impact Assessment. CAM3 represents the complex relationships between network nodes, network services, cyber capabilities, mission-critical tasks, and organizational missions, enabling an analyst to trace and predict operational impact starting from a breached node and ending in an organizational mission. (More detail [here](#).) Also, a 2007 DARPA success story was published about Secure Decisions' visualization of cyber attacks on DOD computer systems, called SecureScope. More detail is available [here](#). The visualizations in the brochure are reproduced below.

Cyber Asset-to-Mission Mapping Model (CAM3) for Impact Assessment

- Relates cyber asset to *network services*
- Relates network services to *cyber capability*
- Relates cyber capabilities to *mission-critical tasks*
- Relates mission-critical tasks to the *overall organizational mission*



The SecureScope™ layered wheel scene traces the path from a security alert to its business or mission impact

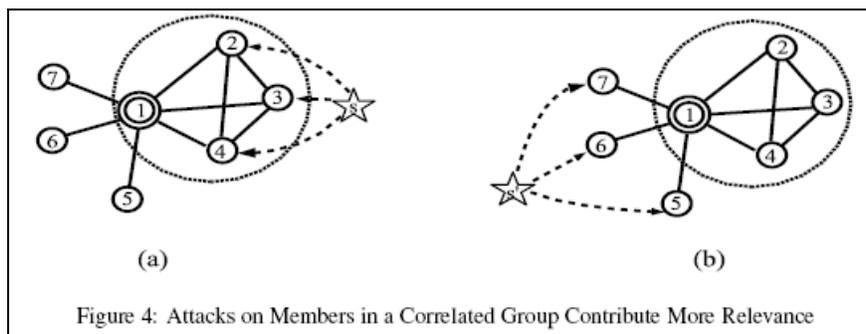
- ★ NSF's Computer & Information Science & Engineering (CISE), Information & Intelligent Systems (IIS) has funded a project called "[Collaborative Research: Mining Biomedical and Network Data Using Tensors](#)" for about \$615k total to **Chris Faloutsos** at **Carnegie Mellon** and **Vasileios Megalooikonomou** at **Temple University**. The project duration is September 2007 to August 2010. The abstract for this project is shown below, and the project also has a [website](#).

Given a large collection of functional Magnetic Resonance (fMR) images over time, how can one find patterns and correlations? Similarly, given a never-ending stream of network traffic information, how can one monitor for anomalies, intrusions, and potential failures? The main idea behind this proposal is to treat both problems using the theory of tensors. Despite the seemingly wide differences in the two settings, they both boil down to finding patterns in multidimensional arrays, sparse or dense. Tensors are exactly generalizations of matrices, and correspond roughly to "DataCubes" of data mining. Matrix analysis and decompositions are part of the standard toolbox for data mining, providing methods for dimensionality reduction, pattern discovery and "hidden variable" discovery. Extending these tools to higher dimensionalities is valuable and tensors provide the tools to do this generalization. However, these tools have not yet been put to use in large volume data mining. This is the main contribution of this proposal.

The investigators propose (a) to design tensor decomposition algorithms that scale for large datasets, with special attention to sparse datasets, and to never-ending streams of data and (b) to apply them on two driving applications, fMRI data analysis and network data analysis. The investigators propose to analyze large volumes of fMRI data performing the following sub-tasks: cluster voxels with similar behavior over time for a given subject and/or task or across subjects and/or tasks, classify patterns of brain activity, and detect lag correlations and spatio-temporal patterns among fMRI time sequences. ... For the network traffic monitoring setting, fast detection of anomalies is important, to spot malware, port-scanning attempts, and just plain non-malicious failures.

- CAULDRON software:** Researchers in **George Mason University's [Center for Secure Information Systems](#)** have developed new software that attempts to reduce the impact of cyber attacks by identifying the possible vulnerability paths through an organization's networks. Says **Sushil Jajodia**, university professor and director of the Center for Secure Information Systems. "This new software is an automated tool that can analyze and visualize vulnerabilities and attack paths, encouraging 'what-if analysis'." CAULDRON ... To manage attack graph complexity, CAULDRON includes hierarchical graph visualizations with high-level overviews and detail drilldown, allowing users to navigate into a selected part of the big picture to get more information. Their research was funded by grants from the **defense, homeland security and intelligence communities and the Federal Aviation Administration (FAA)**. The FAA has already installed CAULDRON in its Cyber Security Incident Response Center to prioritize security problems. [See "[How to detect network vulnerabilities](#)," which includes some links to papers presented on the software.]
- NIST / GMU and "Attack Graphs":** Computer scientists at **NIST** and **George Mason University** are applying security metrics (patent pending technique) to computer network pathways to assign a probable risk of attack to guide IT managers in securing their networks. "We analyze all of the paths that system attackers could penetrate through a network," says computer scientist **Anoop Singhal**, "and assign a risk (to each component of the system. Decision makers can use our assigned probabilities to make wise decisions and investments to safeguard their network." Risk is determined by using these attack graphs and [NIST's National Vulnerability Database](#) (NVD). [For more information, see this [article](#).]

- **Attackers' behavior builds better blacklists.** Computer scientists from **SRI International** and the **SANS Institute** presented a paper at the USENIX Security Symposium in San Jose on a technique that correlates an attacker's preference for victims' networks as a way to prioritize additions to a blacklist. ... By matching up the preferred victims of a known attacker, the researchers have been able to develop per-network blacklists that perform better than either massive global lists or more focused local lists, according to the paper ([full text](#) of "Highly Predictive Blacklisting").



When the system was evaluated using 720 million log entries ... it outperformed global and local blacklists in more than 80 percent of the cases. ... And, unlike lists made from local firewall logs, the researcher's blacklists can proactively block attacks.

- **★ China's Cyber-Militia** – Chinese hackers pose a clear and present danger to U.S. government and private-sector computer networks and may be responsible for two major U.S. power blackouts. The **CIA's** chief cyber-security officer, **Tom Donahue**, is quoted in this extensive article on the subject, appearing in the *National Journal*, May 31, 2008. **Joel Brenner**, the government's senior counterintelligence official and former inspector general of the **National Security Agency**, is quoted extensively, as is **Tim Bennett**, the former president of the **Cyber Security Industry Alliance**.
- **China and the Internet.** According to the **China Internet Network Information Center**, over 253 million Chinese are using the Internet, the highest number ever for the country. ... Says the China Internet Network, "This is the first time the number has drastically surpassed the United States, becoming the world's No. 1."
- **Before the Gunfire, Cyberattacks.** Weeks before physical bombs fell on Georgia, a security researcher was watching an attack against the country in cyberspace. (*The New York Times*) Internet experts in the United States said the attacks against Georgia's Internet infrastructure began as early as July 20, with coordinated barrages of millions of requests – known as distributed denial of service attacks – that overloaded and effectively shut down Georgian servers. Researchers at **Shadowserver**, a volunteer group that tracks malicious network activity, reported that the Web site of the Georgian president, Mikheil Saakashvili, had been rendered inoperable for 24 hours. ... "It costs about 4 cents per machine," said Bill Woodcock, the research director of the Packet Clearing House, a nonprofit organization that tracks internet traffic.. "You could fund an entire cyberwarfare campaign for the cost of replacing a tank tread, so you would be foolish not to."
- **Cybersecurity will take a big bite of the budget.** President Bush's single largest request for funds and "most important initiative" in the fiscal 2009 intelligence budget is for the Comprehensive National Cybersecurity Initiative ... CNCI – or "Cyber Initiative" – is designed to develop a plan to secure government computer systems against foreign and domestic intruders and proactively prepare for future threats.

C. The Big Data Problem

In his 2001 book *Simplicity*, business management expert Bill Jensen noted that conservative estimates showed business information worldwide doubling every three years. More recent estimates say this **data doubles every 11 months today and is expected to double every few days by 2010.**

-- Statement by Jim Davis, SAS Senior Vice President and Chief Marketing Officer [[Source](#)]

- ★ [The MMDS 2008 Workshop on Algorithms for Modern Massive Data Sets](#) was held at Stanford University in June. The [program](#) is available here, and **full text presentations** from many of these talks are available [here](#). The conference addressed algorithmic, mathematical, and statistical challenges in modern large-scale data analysis. The goals of MMDS 2008 were to explore novel techniques for modeling and analyzing massive, high-dimensional, and nonlinearly-structured scientific and internet data sets.

The **networked data and algorithmic tools** section included these talks:

Reid Andersen (Microsoft): "[An algorithm for improving graph partitions.](#)"

Michael W. Mahoney (Yahoo!): "[Community structure in large social and information networks.](#)"

Nikhil Srivastava (Yale): "[Graph sparsification by effective resistances.](#)"

Amin Saberi (Stanford): "[Sequential algorithms for generating random graphs.](#)"

Yuan Yao (Stanford): "[Topological methods for exploring pathway analysis in complex biomolecular folding.](#)"

Piotr Indyk (MIT): "[Sparse recovery using sparse random matrices.](#)"
- [Sifting the Data: \\$3 Million Award Will Build a Foundation for New Ways to Analyze Massive Data Sets Using Visual Analytics](#). The **Georgia Institute of Technology** has received a five-year grant to lead and coordinate a new initiative that will develop foundational research in massive data analysis and visual analytics. A research team headed by **Haesun Park** will investigate ways to improve the visual analytics of massive data sets through machine learning, numerical algorithms and optimization, computational statistics, and information visualization. The \$3 million joint **National Science Foundation** and **Department of Homeland Security** grant establishes Georgia Tech as the lead academic research institution for all national **Foundations of Data and Visual Analytics (FODAVA)** research efforts. [Seven other FODAVA Partnership Awards will be announced later this year](#), all working in conjunction with eleven Georgia Tech investigators to advance the field. "FODAVA seeks to put an improved science base under one portion of the problem – how can we transform large, complex data sets into reduced computational models or mathematical formalisms that retain the information content while better supporting the human in extracting critical information from the data," said **Lawrence Rosenblum**, program director for graphics and visualization at the National Science Foundation.

- *Wired Magazine* carried a [series of articles](#) on the Big Data Problem, “The End of Science ... Welcome to the Petabyte Age.” Articles include:
 - **“Pricing Terrorism: Insurers Gauge Risks, Costs”** ([here](#)).
 - **“Visualizing Big Data: Bar Charts for Words”** ([here](#)).
 - **“Winning the Lawsuit: Data Miners Dig for Dirt”** ([here](#)).
 - **“Tracking the News: A Smarter Way to Predict Riots and Wars”** ([here](#)).
 - **“Sorting the World: Google Invents New Way to Manage Data”** ([here](#)).
 - **“Predicting the Vote: Pollsters Identify Tiny Voting Blocs”** ([here](#)).
- **[Computation Institute to bulk up data analysis capability with \\$1.5 million grant](#)**. The Computation Institute, a joint effort of the **University of Chicago** and the DOE’s **Argonne National Laboratory**, has received a \$1.5 NSF grant for a computer system, called the Petascale Active Data Store (PADS)... The system has been optimized for rapid data transactions, both on campus and around the globe. ... “The Tesla nodes will allow us to experiment with algorithms that combine traditional CPUs and special-purpose GPUs to extract results from data faster than in the past,” said **Ian Foster**, Director of the Institute.
- **[German research focuses on tomorrow's opportunities](#)**. Germany’s **Fraunhofer Gesellschaft's Applied Research Group**, comprising various institutes, has identified 12 strategic research areas in which it is seeking answers to these challenges of the future, including “A Clear Overview in the Data Jungle:”

The internet, company networks, sensors, simulations and electronic processes deliver a wealth of information. ... ‘Visual analytics’ provides a solution, making the best possible use of the different abilities of humans and computers.

D. Funding Opportunities

This section compiles open solicitations for R&D efforts that include network analysis and visualization. The material came from searches of the FedBizOpps database on concept terms (e.g., “network analysis,” “networks + (visualization OR analytics),” “graph analysis”). The solicitations may or may not be open to DOE Laboratories. (BAAs that are clearly not appropriate for a DOE Laboratory are marked as “FYI only” below. We have also used this designation for interesting solicitations that are closing soon or that have already closed.)

We believe the compiled set of funding opportunities should be of use to NGC in that it presents a picture of how government agencies are framing or prioritizing NGC-relevant research.

NOTE: The Air Force announced last year that they were creating a [Cyber Command](#). The target date for Phase I of the stand up is October 1. This apparently has generated substantial funding for work that is related to data visualization of networked information. A number of older solicitations related to modeling / software have also been re-posted (those that are open for numerous years). Most of the solicitations shown below are coming from the AFRL / Rome site, and most are soliciting white papers (not full proposals – generally full proposals are “by invitation only” based on a review of the white papers).

In the last week it has become clear that **plans for the Cyber Command have been “suspended”** and could even be cancelled (see this [source](#), for example). News reports paint a picture of organizational turf wars rather than a re-evaluation of the importance of cyber as a mission space for DOD.

AFRL / Rome:

- [Cyber Command and Control \(C2\) Technologies](#) (BAA0809-RIKA) AFRL – Rome Research Site. The Air Force Research Laboratory, in partnership with the **Global Cyber Integration Center** (GCIC) is soliciting white papers for various scientific studies. AFRL will focus on technologies to enable C2 of cyberspace assets and integrating their effects with kinetic operations. Specific technologies should deliver capabilities that include, but are not limited to **Cyber C2 collaboration**, interoperability, **visualization** and integration of friendly orders of battle, **course of action development**, ... We acknowledge that C2 of non-kinetic assets and the effects they produce and integrating them with kinetic assets/effects is at a formative stage. ... **Total funding for this BAA is approximately \$20M.** ... **All potential applicants are eligible.** ... FY 09 [whitepapers] should be submitted by February 1, 2009.
-  [Proactive Intelligence \(PAINT\) - PART 1 of 2](#) (BAA-07-01-IFKA-PART-1). This is a solicitation for white papers only, recently reissued by AFRL-Rome. “AFRL, in conjunction with the **Intelligence Community's Disruptive Technology Office** (DTO), is soliciting white papers for innovative, creative, and high-risk research to advance the state-of-the-art in technologies and methods for Proactive Intelligence (PAINT). ... This BAA proposes an overall challenge problem: How should the Government locate and discover terrorist networks; describe them in the context of the political, religious, social and economic networks that intersect with, influence, and are influenced by, the terrorist network; explain the structure and topology of the terrorist network; predict the formation, evolution, vulnerabilities, and dissolution of the network; and identify strategies to shape or influence the network through selective action?” **Nuclear nonproliferation** was recently added to the challenge problem. The solicitation is open through FY11 and posts deadlines for submission of white papers in each fiscal year. (Part 2 of this solicitation is available [here](#) and contains additional topics of interest. Total funding is ~ \$48.9M for FY07-FY11.)
- [Special Capabilities in Information and Surveillance \(SCIS\)](#) (BAA-08-07-RIKA), AFRL – Rome Research Site. This solicitation for white papers includes a number of topics of interest, including: Novel approaches to monitor, visualize and recommend alternative courses-of-action in the management and defense of massive, heterogeneous wide-area networks (**5 million nodes**); Innovative methods to visualize complex, self-organizing systems; **Novel three-dimensional data visualization** and projection methods; and Exploitation of human visual physics and physiology of visual perception for efficiently conveying or hiding information. The solicitation is open through September 30, 2012.
- [Applying Information Technology to Command, Control, Communications, Computer and Intelligence \(C4I\) Systems](#) (BAA-05-06-IFKA). The Information Directorate, Command and Control Engineering Branch ... Rome Research Site, is soliciting white papers... Areas of interest include data retrieval / mining, data storage, software / system monitors, data visualization, and many others. Total funding for this BAA is approximately \$49.9M beginning in FY06. ... **FYI only:** one amendment to this BAA was a request for white papers (the deadline has passed) regarding the application of Information Technology to improve Joint Space Operations Center (JSpOC) Intelligence, Surveillance, Reconnaissance Division (ISRD) intelligence (INTEL) functions by application of advanced algorithms and software tools for: automated aggregation of multiple foundational INTEL sources (OSINT, HUMINT, SIGINT, IMINT) and data fusion of INTEL data for the space domain (more detail in the 4/30/2008 [modification](#)).

- [Military Networking Technology for Global Information Exchange \(GIE\)](#) (BAA-03-15-IFKA). AFRL Rome Research Site is soliciting white papers ... in the context of a Global Information Grid (GIG). *One part of this BAA solicits work on Network Modeling and Simulation:* “ ... This solicitation also seeks research on appropriate online network measurement methodologies, which in conjunction with modeling and simulation tools will provide the basis for online network analysis and control.” The BAA is open through September 30, 2008.
- [Airborne Networking Technologies](#) (BAA-06-05-IFKA). ... to support Air Force communications as it applies to Modeling and Simulation, Enterprise Management, and Networking and Communication Links for the future Airborne Network (AN). ... One topic includes “Information Assurance for Airborne Networking” – Innovative ideas are being sought to counter cyber-threats against both network control and data cyber-security attributes suitable for low data rate, low bandwidth, limited connectivity AN Environment. ... Another key area is End-to-End Information Assurance performance metrics and integrated network visualization concepts.

Other Agencies:

- [University and Non-Profit Scientific and Engineering Research](#) (Solicitation Number: N0017808Q3004) Navy – NSWC Dahlgren Division. This solicitation, open through September 30, 2008, includes a topic on data mining: ... Require expertise in the areas of network analysis and link analysis. In addition expertise in the area of text data mining is needed. Text data mining needs include feature extraction, clustering, and classification. Finally expertise in visualization and projection of high dimensional data is solicited.
- **FYI only:** [Research and Development Support](#) (N00173-08-R-JS09). The Naval Research Laboratory (NRL) has a requirement to provide research and development support for a wide range of NRL research activities that include, but not limited to: Systems Engineering and Research Support in Modeling and Simulation; Intelligent Information Systems; Virtual and Augmented Reality Systems and Applications; Visual Information Interfaces for Data Analysts; (Note that [NRL’s general BAA](#) also includes a number of topics that could be considered to include network analysis and visualization. These include [advanced information technology](#), in particular, and also [artificial intelligence technologies](#). This solicitation closes December 31, 2008)

II. Other Applications and Items of Interest

A. Medical / Life Sciences / Pharma

- **Technische Universität Dresden (Germany):** [“Unraveling Protein Networks with Power Graph Analysis”](#) (Royer L, Reimann M, Andreopoulos B, Schroeder M), *PLoS Computational Biology* 4(7), July 11, 2008.

Abstract: Networks play a crucial role in computational biology, yet their analysis and representation is still an open problem. Power Graph Analysis is a lossless transformation of biological networks into a compact, less redundant representation, exploiting the abundance of cliques and bicliques as elementary topological motifs. ... Power Graphs compress up to 85% of the edges in protein interaction networks and are applicable to all types of networks such as protein interactions, regulatory networks, or homology networks.
- [Scientist Unravels Deadly Pathogens](#). A scientist in the Department of Biomedical Engineering at the **University of Virginia Health System** has deciphered the metabolic properties of two dangerous pathogens – discovering how they thrive and which genes, when knocked out, cause them to weaken. **Jason Papin**, Ph.D., principle investigator in the Computational Systems Laboratory at UVA, [used network analysis methods to probe *Leishmania major* and *Pseudomonas aeruginosa*](#). Out of thousands of possibilities, the computer finds the most probable combinations necessary to sustain the life of the pathogen. [More on Jason Papin’s work [here](#).]
- [Analytics and Illumination in the Pharma Industry](#). The pharmaceutical industry, by its nature, has not been able to place customers under the microscope to examine behaviors, preferences and patterns that might influence the bottom line. However, sophisticated new analytics methodologies are finally giving pharma [company]s valuable insights – and measurements – that can help them operate more effectively.

... Patrick Homer, life science sales and marketing practice principal at SAS: ... pharma is now looking to analytics technology and applications to do the same for its industry that they did for the financial services and retail industries five to seven years ago, he said ...

Colin Shearer, senior vice president for market strategy at SPSS: In one project he cited as an example, a customer used its data-mining technology – SPSS Clementine – to predict the variables that drive consumer behavior for a certain treatment type.

Mainstream projects, though, tend to focus on the doctors – the only legal channel through which patients can receive pharmaceuticals. Another SPSS client used its text-mining application to collect and analyze both structured and unstructured data – notes, essentially, that doctors kept in their diaries. Analyses turned up patterns that accompany prescribing behaviors, Shearer said.
- [Smoking Cessation and Social Networks](#). In an article published recently in the *New England Journal of Medicine*, researchers Nicholas Christakis and James Fowler examined a social network of 12,067 people, who were part of the Framingham Heart Study from 1971 to 2003. They concluded that:

 - Smoking behavior spreads through close and distant social ties, groups of interconnected people stop smoking in concert, and smokers are increasingly marginalized socially.
 - Relationships really do help you quit smoking: If your spouse quits, the chances that you will continue smoking decrease by 67%; If a friend quits, the odds decrease by 36%; If you’re employed by a small company and [a] co-worker quits, the odds that you will continue drop by 34%; If your sibling quits, you are 25% less likely to continue.
 - The more educated a contact, the more influential. And, if you yourself are highly educated, you might be more easily influenced.

- [Cellphone Data Track Our Migration Patterns.](#)**
Precise data on the sweep of human movements could aid emergency relief efforts in natural disasters, as well as improve urban planning, public transportation and traffic control. It could also help public-health experts better understand the crowd dynamics of epidemic diseases. "What has changed is that we are being tracked on a daily basis no matter what we do," said network physicist **Albert-Laszlo Barabasi** at **Northeastern University**. In a recent paper in *Nature*, Dr. Barabasi and his colleagues report on the physics of commuting, using six months of cellphone-company data to track the seemingly random movements of 100,000 anonymous European mobile-phone users. By analyzing more than 16 million records of call date, time and position, the researchers distilled these diverse travel routines into a mathematical pattern, like those previously observed among some predators. "It is not an overstatement to say these tools are fostering a whole new type of social science," the editors of *Nature* wrote in an accompanying editorial. [Research published as "[Understanding individual human mobility patterns](#)" (Marta C. Gonzalez, Cesar A. Hidalgo, Albert-Laszlo Barabasi), *Nature* 453, 779-782, June 5, 2008. ([Abstract](#))]
- Projects on Mobility**

 - [The SENSEable City Laboratory](#) at MIT uses cell-phone data to study the complex dynamics of urban life and globalization.
 - At **Northeastern University** in Boston, [The Center for Complex Networks Research](#) uses mobile phone data to analyze the patterns of human mobility.
 - In lieu of direct measures of travel, [The Currency Tracking Project](#) uses banknotes passing from person to person to track human mobility at 'Where's George?'
- [Health care, a strong pulse: Contractors boost capabilities to address expected health care reforms under the next administration.](#)** With a new presidential administration around the corner, a sense of possibility and a renewed focus are energizing the major contractors in the field.initiatives such as the **Health and Human Services Department's [Nationwide Health Information Network](#)** and the **Centers for Disease Control and Prevention's [BioSense](#)**. Along with growing interest in data mining to better use health information to improve patient care, those activities point the way to more spending for health IT, industry executives say.
 - The CDC has funded a project at the **University of Arizona's AI Lab** called **BioPortal**, which has been exploring the use of social network analysis for SARS with a dataset from Taiwan. While this portal is firewalled (www.biportal.org), a description of the work from 2006 appears in this [presentation](#).

B. Business Analytics

- TechNewsWorld* provides a three-part series on "[Breakthroughs in Analytics](#)." Part 1 of this series explores the technologies behind analyzing customer behaviors. Part 2 looks into the tools of business intelligence and Web analytics. Part 3 focuses on analytics tools for niche markets. The series included a collection of market size data, noting that the business analytics ... industry topped US\$20 billion in 2007, according to IDC estimates. ... JupiterResearch put the size of the Web analytics market at \$463 million in 2006.... Oracle remains a leader of the business analytics market... closely followed by **SAS Institute, SAP, IBM, and Microsoft**, the report stated.

* BioSense is a national program intended to improve the nation's capabilities for conducting real-time biosurveillance, and enabling health situational awareness through access to existing data from healthcare organizations across the country. One goal is to develop technologies for automated detection of suspected cases of infection with bioterrorism relevant agents from clinical data streams.

- [Innovation Still At BI Forefront after Round of Consolidation](#). A series of **high-profile acquisitions** last year altered the business intelligence (BI) market. First, **Oracle** bought **Hyperion Solutions** for \$3.3 billion. Then **SAP** acquired **Business Objects** for \$6.8 billion. Finally, **IBM** snapped up **Cognos** for \$4.9 billion. But the BI sector remains vibrant, as hundreds of smaller firms are upending the status quo....
- “**Litigation Readiness**” Application: [MetaLINCS E-Discovery Software Suite: Leading E-Discovery Analysis Capability...](#) The [MetaLINCS E-Discovery Software Suite](#) is a software platform for electronic data analysis for litigation and regulatory readiness. ... “Adding the MetaLINCS E-Discovery software suite to our solutions will enable us to provide very advanced capabilities such as content and social network analysis for our customers,” commented Thadd Hale founder and president of Advanced Discovery Services. *[MetaLINCS is owned by Seagate Technology and the software product here is primarily geared to discovery and document management for legal cases.]*
- [Secrets of Successful IT Teams: Socially Connected Employees](#). Social Network Analysis applied to IT staff.

C. Other Items of Interest

- [Clarkson University Professors Receive Top Marks From Critics](#). Three researchers from Clarkson have published a paper focused on modeling complex networks such as the Internet and the World Wide Web, flight connections, and social contacts, which have recently been under scrutiny for their omnipresence in everyday life. The group's **sequence nets** serve as suitable models of these everyday-life complex nets, but are much simpler to work with. The journal referee concluded, "This paper presents an elegant and useful formalism for classifying and analyzing networks. . . [it] has great future potential and influence." Research is published as:
[Sequence nets](#). (Jie Sun, Takashi Nishikawa, and Daniel ben-Avraham, Clarkson University, *Phys. Rev. E* 78, 026104 (2008)). Abstract: ...Just like threshold nets, sequence nets in general possess a modular structure reminiscent of everyday-life nets and are easy to handle analytically (i.e., calculate degree distribution, shortest paths, betweenness centrality, etc.). ... sequence nets retain many of the desirable analytical properties of threshold nets while yielding richer possibilities for the modeling of everyday-life complex networks more faithfully.
- [Google tool identifies linchpin species. Search system predicts what prey are needed to keep an ecosystem working](#). (*Nature*) Google's search algorithm can be used to determine which prey are most important for an ecosystem to thrive. That's the claim of a researcher who studies food webs, the complex networks that describe who eats whom in an ecosystem.
- [Public Companies Facing Increasing Audit Pressure To Monitor User Behavior Inside the Network According to New Survey](#): Survey Points to Interest in Managed Security Services for Identity-Based Monitoring; Reliance on Managed Services In Retail, Pharmaceutical, Energy, Financial Services and Chemical Industries. “Security and Network operations teams have worked hard to secure the perimeters of their networks, and prove their security posture to auditors. But now, the audit screws are tightening down inside the network,” said Richard Greene, EVP, Worldwide Operations, Securify.

III. Company Activities

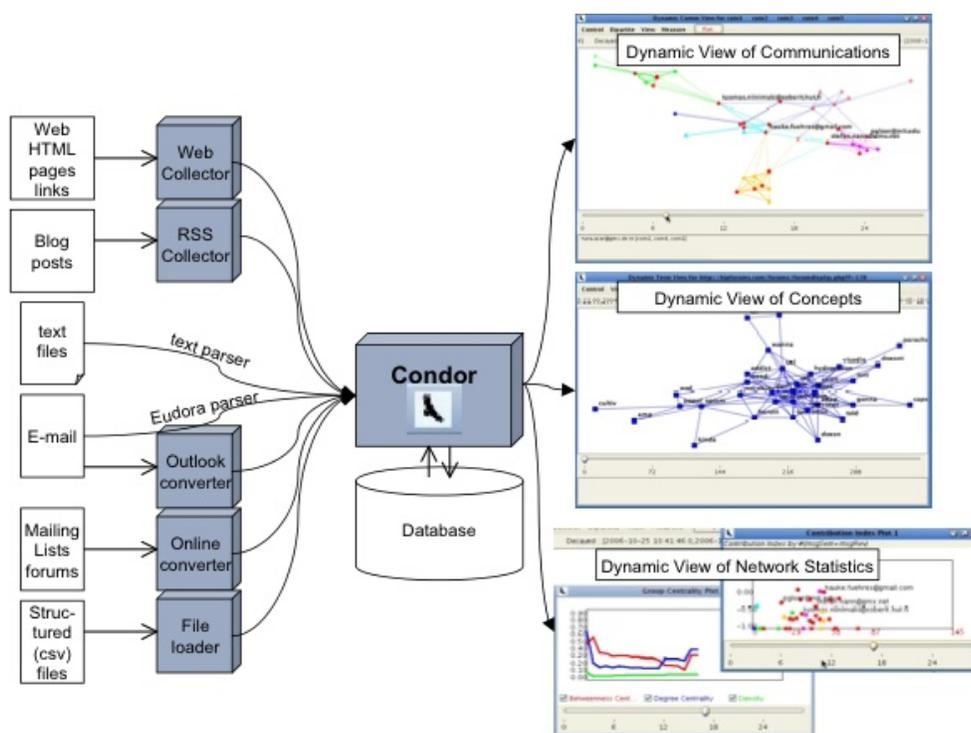
A. In-Depth

- ★ **GalaxyAdvisors** ([website](#)). **Peter Gloor**, an MIT professor and chief creative officer of GalaxyAdvisors, is researching swarm intelligence and crowdsourcing, using some social network analysis software called **CONDOR**. It analyzes the relationships between people and the topics they are discussing in online communications channels, building a map of relationships and connections, and enhancing predictive value of the content. ... software measures online word of mouth to help companies understand and track consumer "buzz," including opinions, preferences, issues and coming trends.

Peter Gloor has been working on visualizing knowledge and analyzing social networks for the last 15 years, leading to the development of Cybermap, a navigation tool for the Web, and to TeCFlow, a tool combining Web navigation with social network analysis. He founded galaxyadvisor's predecessor TeKFlo, Inc. in 2004. He is a Research Scientist at the Center for Collective Intelligence at MIT's Sloan School of Management where he leads a project exploring Collaborative Innovation Networks. ([Additional source](#), including a link to a recent presentation by Gloor; Gloor's bio is [here](#).) Some additional information about Condor is shown below.

Condor allows you to create visual maps, movies and many graph metrics of relationships. Relationships can come from social networks, Web site link structures, and concept maps of unstructured documents, online forums, phone archives, e-mail networks, and many more.

Condor takes as input outlook mailboxes, Eudora mailboxes, Web mailing lists and online forums, Web links, and flat files. It parses those documents and incrementally stores them in a database. It then offers the option to visualize and analyze this data in manifold interactive and visual ways



- ★ [Sprylogics International Corp.](#) (Ontario, Canada) develops advanced search, analysis, and information display tools and services. *[The company advertises its Evidens product as “the analytical power formerly reserved for government agencies.” The CEO is a former Lieutenant Colonel with the Israeli intelligence.]* These solutions enable users to methodically search, collect, analyze, discover, visualize and report large amounts of raw data and transform that data into analyzed and actionable intelligence. Content solutions enable users to identify non-obvious financial risks based on the Social Network Analysis of companies, individuals and other entities. The core technologies driving Sprylogics' solutions are embedded in the **Analyst** and **Evidens analytical workflow** and **Cluuz Search Engine platform** which enables both consumers and corporate users to methodically search the Internet and internal corporate resources and find the information they are looking for. Cluuz search results are visually displayed through patent pending semantic graphs and result in improved decision making capabilities.

In November of 2007, the company held a two-day conference on "**Money Laundering & Terrorism - Exploring Technological Solutions**" which was co-hosted by Sprylogics and the St. Vincent Financial Intelligence Unit. At the conclusion of the conference, a number of regional financial institutions have signed up to commence a trial period of Sprylogics' recently launched beta version IntegraDB product. **IntegraDB Due Diligence Search Engine** contains information from Sprylogics' proprietary Market Integrity Database ("MID") of disciplined individuals and entities. The information is drawn from publicly available sources including regulatory filings. This content is analyzed and uploaded into the Market Integrity Database by the Company's team of intelligence analysts based in Tel Aviv, Israel. A name or an entity analyzed in the Web-based IntegraDB search engine yields a Social Network Analysis of the name or entity searched. The end result is structured data or intelligence which can be used in the due diligence process and acted upon in real time. ([Source](#).) *[For a free trial, go to www.integradb.com.]*

- ★ [DeepMile Networks](#) (Arlington, VA), founded in 2006, provides niche business intelligence, and its DeepSource software. DeepSource was created for government and commercial clients grappling with how to absorb and make sense of huge amounts of data, helping them gain insight from masses of information by identifying the relationships among people, products, companies, or topics being sought. Combining elements of data harvesting, visualization, and social network analysis, DeepSource is being used in the pharmaceutical industry to monitor and evaluate international markets for drug development; to conduct due diligence in the legal and financial services sectors, and to map human networks for counter-terrorism in national security. DeepSource can bring clarity to complex scenarios or to monitor shifting networks of actors by building a reliable picture of those involved. Based on the context or meaning of the underlying source data, information is analyzed and visually mapped, showing the linkages and connections that tie together the entities or topics that need to be understood. ([Source](#))
- [ThinkAnalytics](#) develops data mining and knowledge discovery solutions designed to automate the analysis and interpretation of large data sets. With a focus on developing ... sophisticated, yet easy-to-use, data mining software to automate business analysis. Originally founded as K.wiz Solutions in Glasgow, Scotland in 1995... The company provides leading predictive analytics technology for Telecommunications / Cellular, Media, Banking & Retail industries... [Products](#) include:
 - The [ThinkAnalytics Recommendations Engine](#) provides industry leading personalization. Utilizing state of the art real-time modeling and preference capture techniques, Think Recommendations Engine aims to broaden customers tastes suggesting new and interesting experiences.
 - [Think Enterprise Data Miner](#) offers the most comprehensive and open data mining platform to allow software and Application Solution Providers (ASP) to add sophisticated analytical and predictive intelligence to product offerings, delivering scalability, high performance, flexibility and real-time analysis

B. Company News in Brief

This section contains annotated headlines of news from companies profiled in the benchmark report and some newly identified companies of interest.

- **21st Century Systems:**

[HiRSA Technology Selected for Jordan Border Security Program](#). Its High Resolution Situational Awareness (HiRSA™) software system will provide the surveillance and situational awareness information technology backbone for the end-to-end Jordan Border Security Program...

[21CSI to Demonstrate Critical Technology for New Integrated Swimmer Defense System](#) – advanced technology showcased at **Naval Undersea Warfare Center**. The company providing SentinelNet™ as the integrating software for ISD, linking sonar systems, radars, infrared cameras, and other surveillance equipment to provide round-the-clock, real time data for immediate detection and decision-making.

- **Aptima:**

[Releases Visual Scenario Generator 4.0 to Accompany DDD 4.0 Adaptable Simulation Environment](#): Tool provides rapid scenario development for research, planning and design, and training; applications include “terrorist prediction algorithms.” (Product info [here](#).)

- **Autonomy:**

“[Nuclear Waste Cleans up its Search Problem](#).” EnergySolutions, which had formed from the merger of nine companies in two years, and each company had different document management systems and file types needed an enterprise search engine to identify intellectual property and other important documents. The company looked at a number of options, including Endeca, but selected Autonomy. (According to a report by Forrester Research in May, Autonomy is a top choice in its category for large enterprises, largely based on the fact it can read old documents and file types.)

- **Business Objects:**

[Business Objects Dives Into Predictive Analytics](#) – The SAP subsidiary's first predictive analytics module takes direct aim at the market dominated by **SAS Institute**...The module, based on technology from **SPSS Inc.**, is the latest example of a traditional BI company looking to move beyond its expertise in reporting and analysis tools.... **IBM's Cognos** struck a similar deal with SPSS in March, and also plans to integrate IBM-developed predictive analytics within Cognos. Both efforts take a direct competitive shot at **SAS Institute**, by far the market leader in predictive analytics.

[Business Objects Ranked Number One Business Intelligence Vendor by Revenue by IDC](#). IDC reports that the company has a 14.2 percent market share in business intelligence tools, and when combined with SAP, the company leads the market with a 19.2 percent share.

[SWOT Analysis: Business Objects' Predictive Workbench](#) – the strengths, weaknesses, opportunities, and threats of Business Objects' Predictive Workbench. ... governmental agencies including **Homeland Security**, Medicare, and the Internal Revenue Service are certainly likely to increase their usage of predictive analytics in their efforts to **identify suspicious behavior or prevent fraud**.

- **Centrifuge Systems** (formerly **Tildenwoods**):

[“Launches Ground-Breaking Information Visualization Product”](#) – dubbed the Centrifuge Server, see more product detail [here](#).

- **Endeca**:

[Autonomy, Endeca rate among top enterprise search vendors](#) – Forrester Research says Autonomy, Endeca, Fast and Vivisimo offer the best products. The best enterprise search products on the market come from Autonomy, Endeca, the Microsoft subsidiary Fast and Vivisimo, but Google’s Search Appliance continues to dominate the market in terms of brand awareness and sheer number of customers, Forrester Research says in a new report.

- **Detica**:

★ [BAE to Buy Detica to Expand in Security](#) for **\$1.06B** (£531 million) (see also [here](#)). BAE, Europe’s largest defense contractor by sales, said the proposed acquisition of the consulting company would boost efforts to target the homeland-security markets in the U.K. and the U.S.

[Threat intelligence needs significant re-think to deal with today’s information-intensive world, says Tom Black](#) (report on keynote speech by CEO, who discusses “shifting our focus from digital footprints to threat blueprints – or ‘threatprints’ – which we anticipate ideally before the terrorist or criminal has been radicalized or recruited”).

- **FMS Advanced Solutions**:

“Applying Next Generation Visualization and Analytics in 2008 and Beyond” seminar presentations are available from this [page](#), including one on [“N-Dex Analysis and Visualization”](#) and [“Challenges and Solutions in Data Analysis and Visualization”](#) (co-presented with **Ed Waigand** of the FBI)

- **i2 Inc.**:

[Introduces iCentral Enterprise Investigation and Intelligence Management Software](#) (iCentral streamlines processes, delivers 360 degree knowledge management – product detail page [here](#))

★ [Silver Lake Sumeru Completes Acquisition of i2](#) ChoicePoint announced in late April that it has entered into an agreement to sell its i2 Inc. and i2 Ltd. to Silver Lake Sumeru, a leader in private investments in technology, technology-enabled and related growth industries, in a cash purchase of \$185 million. The acquisition was completed in June 2008.

- **InferX**:

[InferX Corporation Receives Approval for Its Patent Application on Knowledge Inferencing and Data Visualization Method and System](#). The patent covers key components of InferX’s distributed Predictive Analytics software platform, which examines both structured and unstructured distributed data sources on a near real-time basis.

* While Perspectives could not find a published application by the name shown above, another one listing Jerzy Bala as inventor, was published in May of this year. Application 20080104007 is for a [“Distributed Clustering Method”](#), which is about a processing method for distributed data clustering, which involves splitting data points into groups, creating subsets containing data from only one class, and outputting complete rules whereby data points are all located in subsets (full text [here](#)).

- **IBM:**

[Sapiens and IBM Announce new Software \[INSIGHT™\] for Smaller Insurance Carriers](#). ... users can find an assortment of analytical tools, such as Key Performance Indicators, predictive analytics, alerts, visualizations and assorted reports through easy-to-use dashboard controls.

["IBM looks to the future"](#) (includes discussion of the **LanguageWare** product, which enables the analysis and visualization of unstructured text, for issues such as social network discovery and timeline analysis. See this [webpage](#) for more detail on this product.)

- **ISS (Intelligent Software Solutions):** *(Most news items below could not be directly hyperlinked; access them from the "news" tab at the top of their homepage. For more information on any of the products mentioned below, use the "solutions" tab accessible from the homepage.)*

ISS to Deploy PANACIA Information Fusion Solution Around the Globe. The award is to support the deployment of their multi-intelligence fusion system known as "Predictive Awareness and Network-centric Analysis for Collaborative Intelligence Assessment" or PANACIA, to Air Force Distributed Ground Station sites around the world.

A brief [profile](#) appeared in the *Colorado Springs Business Journal*, which noted that the company is adding employees, and has had some \$54.5M in contracts in 2007-2008.

Air Force Awards ISS a \$4.5M Option to its Secure Information Access Analysis and Dissemination (SIADD) Contract.

- **Jute Networks** ([website](#)), a North Carolina based startup, offers a network relationship management software application and some public relationship maps. Designed for professionals, the software brings together web based collaboration, social network analysis and data visualization and is offered as a web based on-demand solution. In July, it went live with a [map](#) of political donors in Asheville, North Carolina.
- **KCC / COPLINK:** [CompStat Analyzer](#) is now available – this is a crime trend analysis and performance enhancement solution that integrates seamlessly with the company's COPLINK® solution suite.

CompStat Analyzer provides both GIS-based mapping functions and a range of graphing and charting tools for analytical functions, including temporal/spatial analysis. The program uses OLAP-cube like technology to allow the user to "slice and dice" all data in the COPLINK database. The mapping portion of the program employs ESRI...

- **Mazu Networks:** [Mazu Networks Announces Response-Time Analytics for Mazu Profiler](#) – latest release marks the first and only solution on the market to offer behavior-based analytics for performance, availability, network utilization and security – unlike a manual threshold approach, Mazu Profiler offers analytics to automatically and dynamically monitor the performance and availability of business services across the IT infrastructure.
- **Orgnet.com and Political Trends in Book Purchasing:** **Valdis E. Krebs**, a founder of orgnet.com, a company that performs social-network analysis, has spent time categorizing books sold on Amazon.com to identify political trends in purchasing patterns. His discoveries include the fact that an increasing number of books are drawing interest among buyers on both the left and the right. "The buying data shows that the old conservatives have more overlap with the progressives than they do with the neocons." ... ([Source](#))

- **SAS:**

[SAS organizing federal government subsidiary to more effectively deliver powerful data analytics to agencies.](#)

[New Capabilities From SAS and Visual Numerics Integrated With Teradata Warehouse Miner Help Streamline the Data Mining Process.](#) Teradata introduced Warehouse Miner 5.2, which embeds 50 new statistical data mining functions from its partners, to accelerate the data mining process.

[SAS Grows Fastest of Top Five BI Vendors; Leads in Advanced Analytics Market Share](#) – SAS' BI Tools Revenue Grew 16.6% in 2007; BI Tools Enter Mass Market. SAS' sales of advanced analytics software, totaling US\$439.8 million in 2007 – with its 31.9% market share being more than twice that of the nearest competitor – and showed 15.2% growth over the previous year.
- **SPSS:**

[SPSS Predictive Analytics Provides Great Returns for Tax Collection Agencies.](#) In the area of tax collection, government agencies use SPSS Predictive Analytics software to quickly risk-assess accounts and accurately identify those with a high probability of delinquency.
- **SRI International:**

[Social Network Analytics Technology Supercharges Popular Online Military Communities.](#) The iLink technology is a machine learning based system that models users and content in a social network. It uses these models to deliver highly relevant information to members of the network across a broad range of scenarios. For users, it can point them to valuable content, discussions and, perhaps most importantly, other users with shared interests and goals. For discussions, it can continuously provide useful relevant content, as well as recommend other users who ought to be part of a particular discussion.
- **Tableau Software**, a privately held company based in Seattle, grew out of a **Stanford University** research project designed to help people better see and understand the information in databases. Tableau markets its data visualization tool to tech-savvy end users as a "light and fast" analytics application in a market dominated by complicated and expensive business intelligence tools. ([Source.](#))
- **Visual Analytics, Inc.:**

[Selected for NY Division of Criminal Justice Services Operation Impact.](#) The agreement is to implement the VAI Data Clarity® Enterprise Software Suite to establish and maintain **Crime Analysis Centers (CAC)** in designated locations across the state – the 17 counties outside of NYC that account for 80% of the upstate and Long Island crime.
- **Gartner Magic Quadrant Reports:** Several of these reports on software companies were released recently: (methodology used is explained [here.](#))
 - [Magic Quadrant for Customer Data-Mining Applications.](#) **SAS** and **SPSS** remain the leading vendors in the customer data-mining application market, supporting the requirements of many CRM initiatives. However, **ThinkAnalytics** emerges as a visionary, while **Angoss Software** and **Portrait Software** assume the role of challengers.
 - [Magic Quadrant for Business Intelligence Platforms, 2008.](#) The market for business intelligence platforms is moving away from a position of being dominated by pure-play vendors. This is being driven by a trend for consolidation, with several large application and software infrastructure vendors initiating major BI acquisitions in 2007.

IV. Resources / Overviews

A. Major Publications

- ★ The **National Academies** has released "[Behavioral Modeling and Simulation: From Individuals to Societies](#)" (Available to read online from this [page](#).) Social network and link analysis are discussed at length in the chapter on [Meso-Level Formal Models](#), including a discussion of limitations and future directions. **Kathleen Carley of CMU** is one of the authors of this publication. The book is described well in an press release from Aptima:

[Chief Scientist Jean MacMillan Co-Editor for National Academies Publication:](#)

State-of-the-art in modeling individuals, organizations, and societies; recommendations to address critical military needs. The National Academies Press has released a major publication on modeling individual and group behavior, co-edited by Aptima Chief Scientist Jean MacMillan. The book, entitled "[Behavioral Modeling and Simulation: From Individuals to Societies](#)" (more on this publication, [below](#)), represents three years of work ... by a National Research Council study panel.

The book explores how IOS models can support military missions, particularly through the application of computational models of human behaviors in social units. A thorough review of the state-of-the-art in organizational modeling, ranging from conceptual and cultural models, to systems dynamics, decision theory, social networks, and games, leads to a discussion on how to address unmet modeling needs, pitfalls, and lessons learned. The R&D roadmap suggested by the Committee focuses on use-driven research to address the major challenges of IOS modeling.

- Evaluation Theory, Models, and Applications*** by Daniel L. Stufflebeam and Anthony J. Shinkfield (Jossey-Bass, 2007) ([book review](#) in the June 11 issue of *JAMA*): ... "A lively discourse is afoot in health evaluation describing how applied researchers are using emerging methods such as geo-mapping and spatial analysis, concept mapping, photography, and [social network analysis](#) to improve studies of germ-social-human-community relationships."
- FYI: Risk analysis for complex systems** ... typically involves analyzing sequences of failures, figuring out every part of the system that might fail and what effects that might have, and putting all those pieces together - essentially, a bottom-up way of looking at things. MIT professor, **Nancy Leveson**, after nearly three decades of working on such problems, has ... **[developed]** a new top-down way of analyzing the risks of complex systems, which leads to a more integrated approach to managing the risks. Leveson calls her new approach **STAMP**, for **System-Theoretic Accident Model and Processes**. ... chapters of the book are available online on Leveson's web site (<http://sunnyday.mit.edu/book2.pdf>). "Nancy Leveson has developed a control-based modeling approach to systems safety which can be applied to complex networks of hardware and humans," says professor Jeffrey Hoffman, a colleague of Leveson's in MIT's Aeronautics and Astronautics Department. "Her work has elicited considerable interest inside **NASA**, where safety analysis has traditionally concentrated on the reliability of individual pieces of complex systems."

B. Reviews and Overview Articles

- [Growing Field of Search Options](#). While the field is dominated by Microsoft and Google, a number of start-ups are looking to establish themselves by offering niche capabilities that may prove vital as search becomes more endemic throughout the enterprise. A firm called **Endeca**, for example, concentrates on database and data warehouse search, primarily for XML, RSS and documents. **Vivisimo** is gaining a name for itself by [tying search to social networking](#). (See also this [article](#) on top enterprise search vendors.)
- [Semantics Gives The Web Meaning – For Machines](#). The European Union's investment in semantic web research has already far surpassed other regions and countries, including the USA. The EU's Sixth Framework Programme (FP6, 2002-2006) for research has funded 17 semantic web projects and about **€50m annually** is allocated to continued research in this area under FP7, which runs until 2013.

Two new projects look set to make a global impact on the semantic web. First, the [REVERSE](#) network has worked on a set of interoperable reasoning languages for advanced web systems and applications. Second, the [LarKC](#) project is looking at semantic reasoning to solve a fundamental problem of the semantic web: its size. LarKC will be a platform for massive, distributed, incomplete reasoning. It will achieve scalability both through its lack of completeness (it decides when it has queried 'enough' data) and its parallel processing (on clusters of high-performance computers or through a distributed network of 'home' computers).

C. Resources on the Web

- **Blogs:**
 - [TNT \(The Network Thinker\)](#) is a blog run by **Valdis Krebs** of orgnet.com. This blog is focused on "exploding" old concepts and thinking about economies, organizations, communities and groups. "We will focus on new forms of connectivity and emergence in organizational, community, and social networks and how these new structures lead to resilience, adaptability, agility, and innovation."
 - A related blog, [Network Weaving](#), is a "social network blog about the creation of robust and vibrant economic and community networks... using network mapping, weaving and leadership development." It is also run by Valdis Krebs, along with Jack Ricchiuto and June Holley.
 - [Large-Scale Social Network Analysis](#): "This blog is dedicated to discussing the tools, techniques, and challenges relevant for working with very large social network data sets. In particular, the focus is on the theoretical, algorithmic, and statistical challenges facing social network researchers."
- **CASOS Network Analysis [Dataset List](#)**: This site has links to datasets available from Carnegie Mellon University as well as an extensive list of those from other sources.
- [Software for Social Network Analysis: Pajek and Friends](#). This blog posting contains some good links on SNA software.
- New Idea Engineering Inc., a consulting company that helps businesses implement enterprise search ... provides a directory of "open source, no cost, low cost and commercial software tools components and products" at [Search Components Online](#).

- [DataShaping.com](#): This site maintains a good data mining meta directory: “general resources with emphasis on fraud detection, CRM, advertising technology, web mining, probabilistic trading, scorecards, risk management, market research, business intelligence, artificial intelligence, statistical technology, information retrieval, computational marketing.

D. Conferences

- **Intelligence and Security Informatics**. (Available online [here](#).) This book, part of the *Lecture Notes in Computer Science* series, constitutes the refereed proceedings of the three international workshops in Taipei, Taiwan, in June 2008. The Pacific Asia Workshop on **Intelligence and Security Informatics** (PAISI 2008) cover topics such as information retrieval and event detection, internet security and cybercrime, currency and data protection, cryptography, image and video analysis, privacy issues, social networks, modeling and visualization, and network intrusion detection... The Pacific Asia Workshop on **Cybercrime and Computer Forensics** (PACCF 2008) furnishes 10 papers about forensic information management, forensic technologies, and forensic principles and tools. The **Workshop on Social Computing** (SOCO 2008) ... topical sections on social web and social information management, social networks and agent-based modeling, as well as social opinions, e-commerce, security and privacy considerations.
- The 4th [Annual Text Analytics Summit](#) was held in June. The conference has a nice mix of solutions vendors and users describing applying text analytics to their particular issues.
- The 4th [UK Social Networks Conference](#) was held in London in July. The conference covers a number of topics on theories, methods, or applications of social network analysis. The program covered business, online, policy, social capital, politics and other subjects. Detail is available [here](#). Keynote speakers included Martin Everett of the University of East London, Tom Snijders of Oxford, and Ron Burt of the University of Chicago.