

# nSpace and GeoTime - VAST 2006 Case Study

Pascale Proulx, Lynn Chien, Robert Harper, David Schroh, Thomas Kapler, David Jonker and William Wright

Oculus Info Inc.

## ABSTRACT

GeoTime and nSpace are new innovative visual analytic tools that enable investigation and understanding of complex, ambiguous situations. nSpace is a system of systems supporting the whole analytic workflow. GeoTime provides insight into events and behaviors in time and space. The analytic capabilities supported by these tools were used by junior analysts to investigate the 2006 VAST contest dataset and to produce an intelligence assessment report. Their work won first place in the corporate category. This paper is a detailed case study describing how the nSpace/GeoTime tools created an analytical environment that enabled two novice analysts to examine the scenario, discover patterns, trace connections, assess evidence, visually represent meaningful hypotheses with associated evidence, track progress, collaborate with others and then produce a final report.

**CR Categories:** H.5.2 [Information Interfaces & Presentations]: User Interfaces – Graphical User Interfaces (GUI); I.3.6 [Methodology and Techniques]: Interaction Techniques.

**Keywords:** visual analytics, human information interaction, sense making, geo-spatial information systems, temporal analysis.

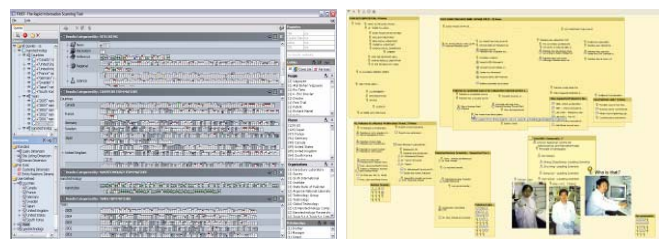
## 1 OVERVIEW OF THE ANALYTIC TOOLS

GeoTime and nSpace are two novel visual analytic applications that have been developed in collaboration with analysts to support the investigation of large and complex multi-dimensional datasets. GeoTime supports the visualization and analysis of entities and events over time and geography [6] and is currently in transition to deployment for analysts to use on a day-to-day basis. nSpace is used for triaging massive data and for analytical sense-making [5][12], and is currently undergoing pilot deployments.

As shown in Figure 1, *nSpace* combines the multi-dimensional linked views found in TRIST (The Rapid Information Scanning Tool) with the visible, cognitive mechanisms of the Sandbox. TRIST provides query planning, rapid scanning over thousands of search results in one display, and includes multiple linked dimensions for

result characterization and correlation. Change difference visualization allows the efficient comparison of different sources and the monitoring of issues from day-to-day. Analysts work with TRIST to triage their massive data and to extract information into the Sandbox evidence marshalling environment. The Sandbox is a flexible and expressive thinking environment that supports both ad-hoc and formal analytical sense making. Key capabilities for the Sandbox include “put-this-there” cognition, animated automatic layout of evidence with analytical templates, gestures for the fluid expression of thought, assertions with supporting/refuting evidence gates, analytic methodologies such as analysis of competing hypotheses (ACH) and scalability mechanisms to support larger analysis tasks.

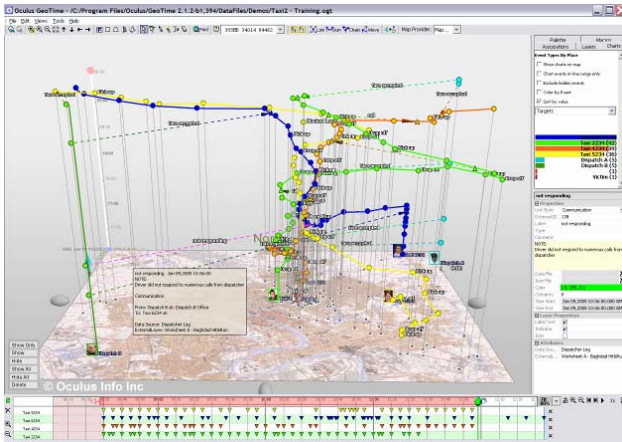
nSpace is an integration platform that also marshals tools and data sources for analysts. Using open web services interfaces and protocols [11], nSpace integrates computational resources such as reasoning services [2], agent-based modeling and advanced computational linguistic functions [1] [4] including entity extraction, supervised and unsupervised clustering, and automatic ontology construction [7]. nSpace spans multiple data sources using a variety of Service Oriented Architecture (SOA) integration mechanisms such as OpenSearch (RSS / Atom), REST and SOAP Web Services and native JDBC connectivity. Configuration settings allow multiple computational resources, in different organizational contexts, to be combined in this unified analysis environment.



**Figure 1: nSpace is a two screen system with TRIST for information triage on the left and the Sandbox for evidence marshalling and sense-making on the right.**

*GeoTime* improves the perception of entity movements, events, relationships, and interactions over time within a geospatial (or any conceptual diagrammatic) context. As shown in Figure 2, events are represented within an X,Y,T

coordinate space, in which the X,Y plane shows geographic or diagrammatic space and the vertical T axis represents time. Events animate in time vertically through the 3-D space as the time slider bar is moved. For analysts, GeoTime's single view representation of a combined temporal-spatial three-dimensional space amplifies concurrent cognition of entity relationships and behaviors in both space and time [6]. Analysts can see the who and what in the where and when.



**Figure 2: GeoTime system for analysis of events and transactions in time and space. Time is in the vertical dimension.**

nSpace and GeoTime are both examples of external mental aids which enhance human comprehension and reasoning abilities. With the correct visualization, interaction and animation approaches, human cognition is amplified by increasing human mental resources, reducing search times, improving recognition of patterns, increasing inference making and increasing monitoring scope [3] [10]. Visibility, interaction and animation of large amounts of complex information increases human cognitive productivity. Analysts can see and work with more information, more quickly with more comprehension.

## 2 THE VAST CONTEST

### 2.1 Objectives

The purpose of the Visual Analytics Science and Technology (VAST) Contest ([conferences.computer.org/vast/vast2006](http://conferences.computer.org/vast/vast2006)) is to promote the development of benchmarks for visual analytics and establish a forum to advance evaluation methods. A fictitious scenario is used and the contest dataset in 2006 consisted of 1,200 news stories from a city newspaper, photographs, maps, spreadsheet of voter registry information and a phone call log. The goal was to determine if wrongful activities are taking place in the city of Alderwood in the state of Washington and report on the situation and the assessment of the situation. Questions to be answered included: Who are the players engaging in deceptive activities in the plot(s)? What is the time frame in which the plot(s)

unfolded? What events occurred during this timeframe that are most relevant to the plot(s)? What locations are relevant to the plot(s)?

### 2.2 Evaluation Metrics and Results

The submissions were evaluated by the correctness of the answers, the quality of the explanations of the process used, and the description of how the tools facilitated the analysis. The judges were experienced in both analysis and visual analytics, and the methodology was informed by recent work in evaluation metrics [8] [9]. Two full day judging sessions took place in separate locations and a total of ten judges reviewed the materials submitted by all the teams. A two hour, live analysis session using the tools with real analysts was also conducted at the VAST Conference. The Oculus submission using nSpace and GeoTime won first place in the corporate category and was acclaimed for the capabilities of the tools in finding patterns through filtering and visualizing massive and unstructured data.

In particular, the judges noted that our submission hit the major threat plots, went in depth into the politics of the city council and that our assessment was "excellent". According to the judges, we had "many correct answers" and some incorrect answers. The judges also noted we "found more suspicious stuff in the city council than was perhaps going on" and that "much analytical effort was spent going down the wrong path". With respect to this last remark, in our experience, this is the nature of day-to-day analysis, and so serves to reinforce that the contest task is a representative analytical task.

## 3 THE ANALYSIS PROCESS WITH THE VAST DATASET

The analytic capabilities supported by the nSpace and GeoTime tools were used by two junior analysts to investigate the 2006 VAST contest dataset and to produce an intelligence assessment report. The following describes how these tools created an analytical environment that enabled the novice analysts to examine the scenario, discover patterns, trace connections, assess evidence, visually represent meaningful hypotheses with associated evidence, and then report on their final assessment.

### 3.1 Brainstorming without Boundaries

The Sandbox in nSpace assists in divergent thinking. Ideas are unrestricted and thoughts can flow freely and be recorded by pointing and typing anywhere in the Sandbox. The annotations can then be quickly organized to create more structure when necessary. The VAST contest provided the contestants with a task sheet that introduced the scenario to be investigated and listed specific key questions. The junior analysts began by jotting down these key questions and generating possible hypotheses using prior and tacit knowledge, along with strategies or plans to answer the questions. Sentences that appeared to be the locus of the task were noted while key words taken from task descriptions, such as "city hall" and "laboratory", were organized into a keyword group, as shown in Figure 3.



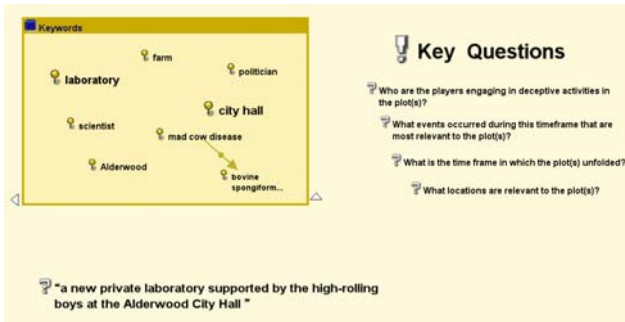


Figure 3: Brainstorming in the Sandbox

exploring it using the multi-dimensional linked views supported by the tool. An overview of the TRIST workspace, at one point in the analysis, is shown in Figure 4. TRIST is organized into panes for queries, dimension setup, dimension viewing and a separate entity dimension area. The center results view provides multi-dimensional characterization of information objects in result sets that allows for fast, user controlled scanning. Results can be organized by dimensions such as source, country, technology, date as well as automatic ontology and unsupervised clustering discovery dimensions. User constructed tailored dimensions and their associated bins can be selected or determined by the analyst in the lower left.

### 3.2 Setting the Stage for Scanning

The analysts began their investigation by importing the entire corpus of 1,200 news stories into TRIST and

Figure 4: TRIST loaded with the VAST dataset. Some of the dimensions used in the VAST contest are shown in the center. The document legend is shown in the top center window bar. The other major components of TRIST are labeled. All the documents in the “Philip Boynton” bin in the Key Players dimension are selected. The same documents are highlighted in the other dimensions.



**Figure 5: Adding a found individual to the “Key Players” dimension by dragging and dropping from the extracted entities dimension.**

The analysts used keywords taken from the brainstorming stage in dimensions to cluster the data and extract information. For example, “city hall” and “laboratory” were used during the preliminary filtering process. All information objects (e.g. documents) associated with those keywords were clustered together into dimensions with bins of categories, and then scanned quickly using tooltip-like mouseovers. In an analysis, keywords are organized under dimensions such as *Key Players* and *Key Issues*. These tailored dimensions (in this case focused on key issues, players, and locations) were created and expanded as the analysts acquired further intelligence by reading through related articles associated with the initial keywords.

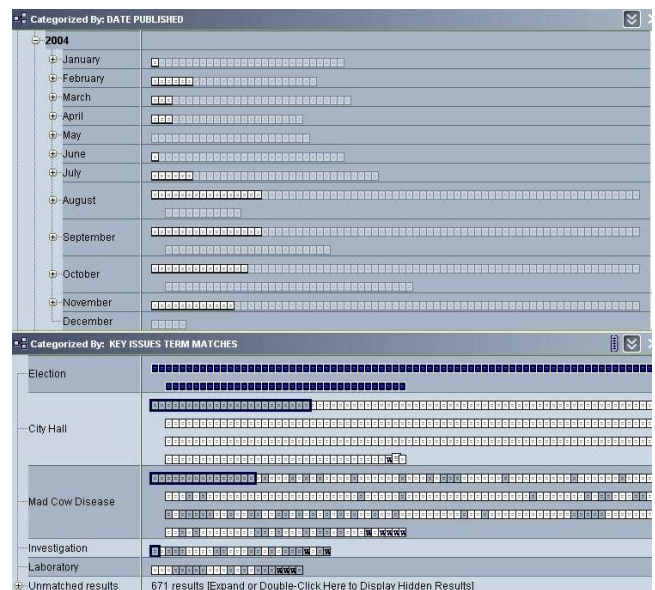
### 3.3 Scanning with Multi-Dimensional Linked Views

TRIST supports the exploration of data oriented to any particular line of thought. For example, to get a sense of trends in the issues over time, the analyst selected each issue by selecting the information objects in that issue’s bin. Those same information objects are then highlighted in each of the other dimensions. Figure 6 shows an example.

The analyst can use TRIST to see who is connected to which issue. For example, starting in the Key Issues dimension, the analyst selected all results in the laboratory bin which automatically highlighted all people, places and organization entities mentioned in the corpus related to the laboratory using the automatic entity extraction dimensions. With the relevant entities seen in context, selected entities were added to user-defined dimensions such as *Key Players* using rapid drag and drop gestures, which can be seen in Figure 5. In general, this process was repeated for any interesting topic found.

After gathering all suspected key issues, players, and locations from the preliminary data extraction, the analyst further acquired awareness by looking at links between two dimensions. Because the analyst wanted to start answering the first question “Who are the Players engaged in deceptive activities?”, they began in the *Key Players* dimension and selected all results in the bin for *Rockford*, the City Manager of Alderwood. Selecting his bin automatically highlighted all people, places and

organizations mentioned in the corpus related to Rockford. The analyst then switched her attention to the *Key Issues* dimension and scanned for activities linking *Rockford* with the issues. The analyst also used the *Date Published* dimension, which revealed the time range of the articles mentioning Rockford, and was able to scan the *Rockford* results chronologically.



**Figure 6: An example of linked views for trend analysis. Selecting the “Election” issue bin highlights those same information objects in the time dimension. The analyst can then scan by jumping from highlighted document to highlighted document in the date dimension.**

One example of a surprising discovery was several discrepancies in the dataset timelines found while using linked views for each key player. Documents from every year between 2002 and 2004 announced the old City Manager David Fonfara leaving for a new job. Similarly, Rockford was announced to become the new City Manager in two different years. It was also observed that Councilor Alex de la Cruz moved from Alderwood once a year. Further, highlighting the documents in each Alderwood politician’s bin revealed their names shown only in the months that they were in the job position. For example, the previous City Manager Fonfara’s bin only showed articles



between December to February, the interim manager's bin only revealed articles from January to April, and the new manager Rockford's bin only revealed articles between January to July. This was consistent with the monthly timeline of their position but the years were not reliable. It was learned later from the contest judges that these discrepancies occurred because the dataset producers decided to change the year of the news stories during the last stages of the production but had overlooked changing the years for all the documents. This discrepancy was not discovered until the analysis submitted by Oculus was reviewed.

### 3.4 Saving and Organizing in the Sandbox.

Throughout the triage process, a variety of relevant information was discovered, including key events, people, relationships, sources, etc. The analyst saved these relevant fragments of information by dragging them into the Sandbox. A short, to-the-point fragment is enough since nSpace inherently supports linking of fragments to full sources. In the Sandbox, the analyst quickly organized relevant content using put-this-there interaction. The

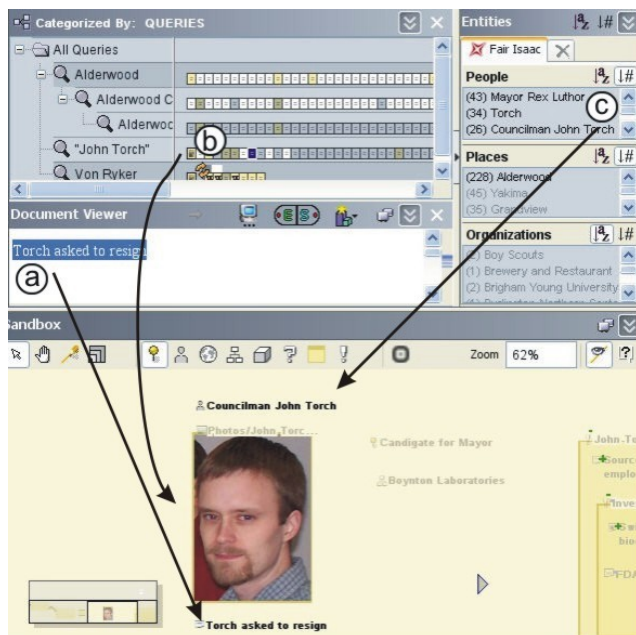


Figure 7: Dragging and dropping snippet, entity and image information objects into the Sandbox

analyst grouped events about the laboratory in chronological order, while another group was created to classify the voting behavior of council members. From these classifications, observations regarding alliances, voting frequency, and decision making were derived. The analyst continued their investigation by exploring relationships between the key players. Each key player entity was expanded into a group object in order to embed further detailed profile information based on data discovered by scanning with the *Key Players* dimension. Fragments that gave insight to individual key players were

dragged and dropped into the associated entity groups, as shown in Figure 7.

### 3.5 Creating a Social Network and Detailed Profiles

During the process of gathering information on key players, the analyst developed a profile structure that included sub-group objects for professional and personal background, positive and negative relationships, and the person's personal and professional character dispositions. The analyst then decided to save this structure as a template, named "Character Profile", which could then be reused and shared. This template was then applied to all the other key players and all the associated snippets of found data. The new Character Profile template automatically organized the associated fragments according to that same template structure using animation and supervised clustering. Some examples are shown in Figure 8.

Since applying the Character Profile template revealed the backgrounds of and relationships between each key player, the analyst was able to distinguish the networks between the key players. By drawing links between them, each with a label and associated evidence for the relationship, the analyst was able to keep track of their relationships from a more visual point-of-view, as shown in Figure 9. Two important connections that helped answer some of the questions posed by the contest were found from the developed profiles and social network: the alliance between certain members of the City Council, and the discreet connection between the laboratory and the City Hall, who turned out to be a woman by the name of Laurel Sulfate. However, an unexpected relationship between what appeared to be two completely unrelated entities was also discovered while developing the social network. The employees of the laboratory in Alderwood were former directors of a disreputable laboratory in Switzerland.



Figure 8: Automatic profiles in the Sandbox.

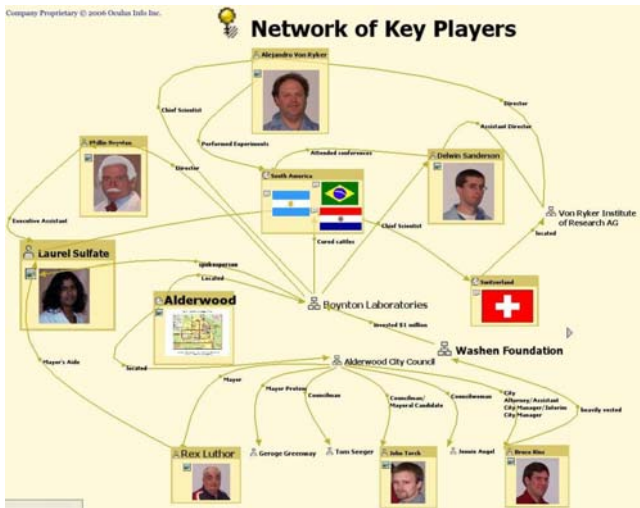


Figure 9: Key Players Network in the Sandbox.

### 3.6 Capturing Questions, Thoughts and Hypotheses

The Sandbox supported the analyst in the development and assessment of meaningful hypotheses, which were captured as assertions. These assertions make explicit the points the analyst was trying to prove or disprove. Evidence was associated with an assertion using simple drag and drop gestures through the refuting or supporting gates of the assertion. In the Sandbox, an assertion has “Support and Refute Gates” along the sides. The act of binding evidence

to an assertion is done fluidly by dragging the evidence to the assertion group through the appropriate gate. The supporting gate is along the right edge and the refuting gate is on the left of the group box. The gates light up red or green when evidence is brought into the assertion [6].

The analyst generated several competing hypotheses, to assess and determine patterns of behavior within the scenario. These were generated using both a bottom up (interpretations of the evidence) and top down (assumptions about what is implied by the top level assertion) fashion. For example, after discovering the relationship between the laboratories in Alderwood and Switzerland, the question of whether the laboratory in Alderwood could possibly be a part of a bigger scandal materialized and hypotheses were made in order to marshal the various evidence. The diagnosticity of supporting evidence was supported by linked views and the highlighting of identical evidence in multiple competing hypotheses.

As plot elements became more tangible, the visual representation and layout of the Sandbox began to match the mental model of the analyst. Figure 10 shows an example of assertions that were brainstormed and filled as further evidence developed from analyses made through the collaboration of nSpace and GeoTime. The thumbnail in the lower left is an overview of the entire Sandbox analysis-in-progress.

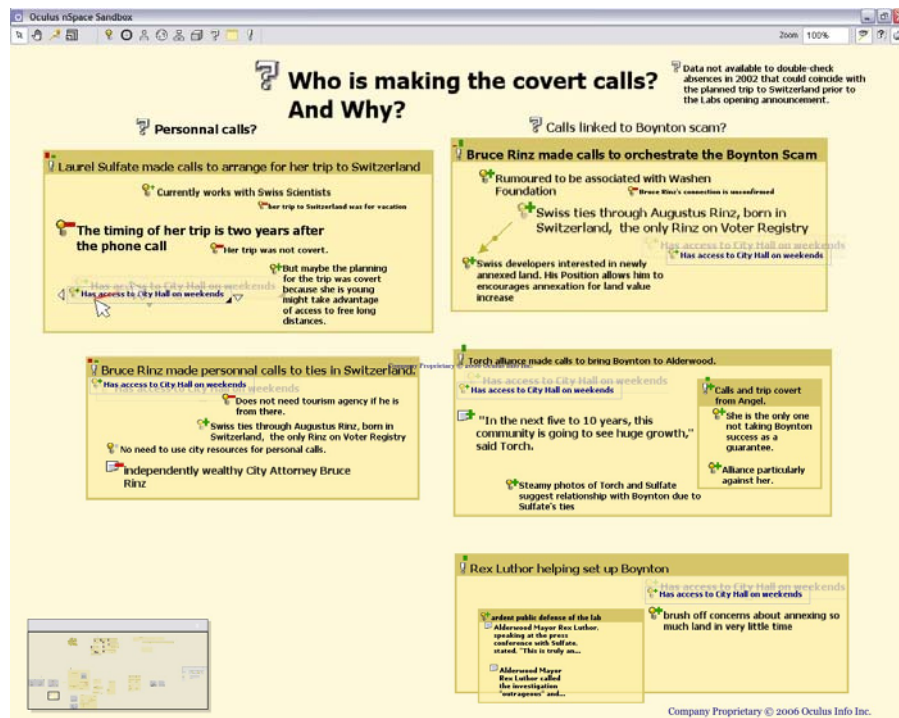


Figure 10: Assertions used for some hypotheses. The degree of support or refute weight for each piece of evidence is shown with subtle green and red symbol indicators on each piece and then aggregated in the upper left of each assertion group. The diagnosticity of evidence can be seen by selecting an item and seeing it highlighted in the other hypotheses. In this example, “Access to City Hall” has low diagnosticity because it is used in all hypotheses.

### 3.7 Importing Data into GeoTime

Entities and events identified as potentially relevant in nSpace were imported into GeoTime using a combination of automated import using an Excel-GeoTime plug-in, as well as manually entering data using the appropriate built in wizards. Each event was linked to the related data source, such as a news clipping, for access to the full context from within GeoTime.

### 3.8 Collaborating to Divide and Conquer

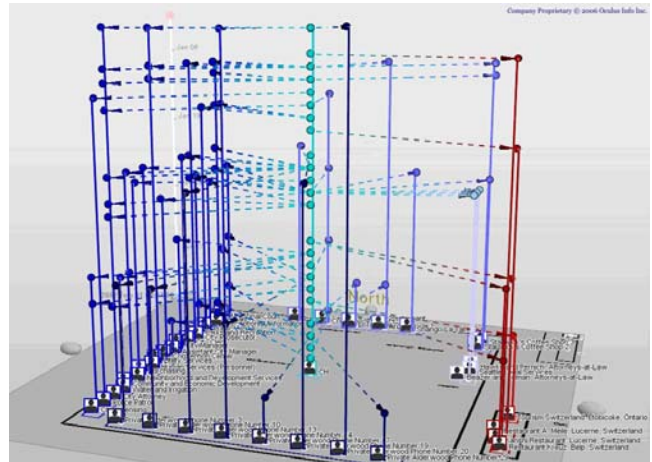
In order to further investigate the relationship between the Alderwood and Switzerland locations, events at each location as well as between locations were explored. The use of both nSpace and GeoTime was able to initiate a synergetic exchange of ideas, where hypotheses generated in either software tool could be cross-referenced, refuted, and supported by the other tool. While one analyst continued to extract, query, and assemble associated data between Alderwood and Switzerland through nSpace, another analyst used the data from nSpace and plugged it in to GeoTime to attain a geographic and temporal visual materialization of events. For example, a phone log showed international calls being made from City Hall and a quick search in TRIST revealed that the numbers belonged to Swiss restaurants and a Swiss tourist agency in Canada. The analysts therefore decided not only to map important events in GeoTime, but also to map the call log in order to cross-reference the phone calls to the events in order to derive further patterns.

### 3.9 Adding Visual Structure with Diagrammatic Layouts

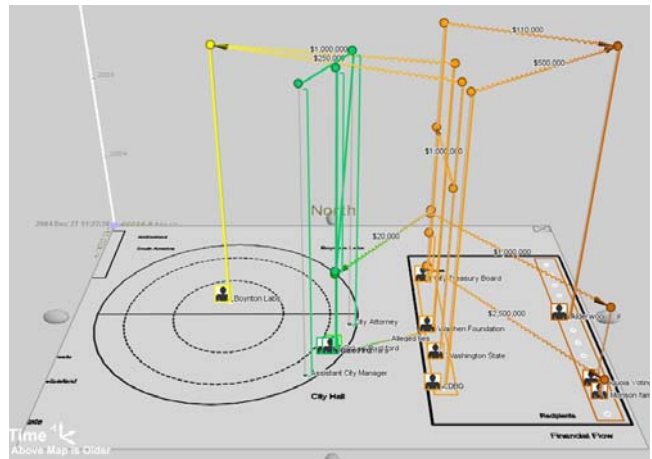
Once in GeoTime, the analyst added structure and context by classifying related entities into groups. “Places”, which are normally geo-spatial locations, in this case were defined as more abstract locations in a diagram. Diagrammatic places are not constrained by specific lat/long coordinates, and were simply moved around with drag and drop gestures to suit the analyst’s notions of organization and relationship. Various symbolic shapes were created using GeoTime’s *Ink* tool to help visualize the structure of these diagrammatic frameworks, and text was added to supplement the frameworks. Color was also used to convey intended relationships and functions. An example screenshot is shown in Figure 11.

An analysis of all financial transactions was also done using this diagrammatic GeoTime technique. Any transaction found using TRIST was noted in the Sandbox by the analyst and then transferred to GeoTime. The resulting GeoTime configurable space is shown in Figure 12. This supported a bottom-up type of analysis to discover any trends, patterns or anomalies. However, no patterns or anything unusual was found. It turned out that

all the financial information was noise inserted by the contest scenario authors.



**Figure 11: GeoTime was used to analyse the significance of phone calls. Abstract “places” are on the ground plane and events in time are shown vertically. Transactions, in this case phone calls, are shown as dotted lines connecting two or more events.**



**Figure 12: An examination of financial transactions using diagrammatic spaces in GeoTime.**

### 3.10 Isolating Patterns and Capturing Hypotheses

GeoTime visualization makes complex patterns of events in time and space easier to recognize. Various built-in tools in GeoTime also facilitate the detection and recognition of interesting events. For example, in an attempt to root out the suspicious phone log findings, the GeoTime *find* tool was used by the analyst to execute searches for “Switzerland”. The results revealed a host of events, entities and places all with links to Switzerland. These were dragged to a Switzerland *layer*. Layers are containers used to categorize and organize interesting observations and evidence for quick future referrals.



At any point, the insights provided by the GeoTime visualization can be captured in annotated graphical reports. These reports can then be integrated with the other arguments in the Sandbox and subsequently prompt the analyst to go back to TRIST to look for additional supporting or refuting evidence for these hypotheses. GeoTime graphical reports can also be used in a final report.

In one GeoTime analysis, shown in Figure 11, the event data as well as the phone call transactions showed a wide range of communications, but interesting patterns were identified and isolated. A pattern that emerged from the data was the timing of the international phone calls, which were made right before city council members left town. Further, these calls were only made on weekends, as opposed to other phone calls to Alderwood and Washington that were made on weekdays. This activity supported the hypothesis that there might have been someone at City Hall planning a covert trip to Switzerland, perhaps through Canada (there was a 416 area code in the call log), and the evidence was added to support an assertion in Figure 10. Other evidence to support or refute the theories included arguments and conclusions derived from other parts of the analysis in both applications.

The question of who was the person responsible for the calls became one key question and the analysis was directed to respond to the question. As shown in Figure 10, most evidence supported the hypotheses of certain political figures such as Bruce Rinz, John Torch, and Rex Luthor being responsible for the phone calls while much refuting evidence existed for the other hypotheses such as Laurel Sulfate making calls to arrange for a trip to Switzerland.

### 3.11 Identifying Missing Data and Tracking Task Completion

As with many instances of intelligence analysis, at any stage, there were always gaps in the information to be filled in order to make a more sound assessment. nSpace is not only capable in analyzing and assessing evidence, it is also useful in helping indicate what issues were outstanding, what data was missing and what links needed to be completed, before conclusions could be drawn. In the case of the VAST scenario, after answering the key questions and assessing the situation, the junior analysts reviewed the Sandbox and looked for the questions that were recorded throughout the process of analysis to check if all the bases had indeed been covered. Questions such as what exactly was Laurel Sulfate's role between City Hall and the Alderwood laboratory, and whether the disreputable laboratory in Switzerland was actually a fraudulent entity still remained to be answered. One of the recommendations at the conclusion of the analysis was to acquire further data for these issues as a next step.

### 3.12 Reporting

An answer sheet was completed according to the VAST contest requirements. Tables were filled in with answers

and references. When relevant text snippets were dragged and dropped from the Sandbox into the MS Word report document, associated references were automatically included. The contest also required a debriefing summary of the story and what was found. Objects from the Sandbox (e.g. groups with notes, snippets and assertions) were dragged and dropped into MS Word with their structure, formatting and references preserved.

## 4 STRENGTHS

Using nSpace, the analysts were able to freely brainstorm, identify important entities, find evidence, generate and refine hypotheses, marshal evidence, make connections and report findings. Combined with its intuitive interaction, such as "put-this-there", as well as automatic computational linguistic functions [4], nSpace supported both ad-hoc and more formal analytical sense making techniques.

GeoTime provided a geo-temporal environment that supported the analyst in understanding entities and events, including communications, money flows, positions in organizations, relationships and the physical movement of key people. GeoTime's fused one-data-view system eliminated the typical analyst's dependency on multiple windows of spreadsheets, tables, maps and other simultaneous cross-referenced separate views that can increase cognitive load and cause confusion and errors.

During the course of the analysis, a number of people reviewed the work and provided input. This same-time, same-place collaboration was possible, without everyone having to read and explore everything in the detailed scenario, due to the ability of nSpace and GeoTime to support *visual collaboration*. Playing key events through time in GeoTime quickly brought a colleague up to speed with the overall story. Since the Sandbox provides an explicit representation of much of the analytic thinking, this facilitated presenting the analysis and eliciting specific feedback on particular hypotheses.

## 5 LIMITATIONS

Currently, nSpace does not scale well past approximately 5,000 extracted entities. This means that some pre-processing of the data was necessary before using nSpace to explore the voter registry dataset. nSpace is currently undergoing an optimization review to increase scalability. Additionally, nSpace does not currently support a specialized timeline view, requiring the analyst to manually lay out events chronologically, or to use an automatic time template.

While GeoTime inherently supports temporal analysis, it is currently overly laborious to import events into GeoTime from nSpace. Data transfer chores inhibited the fluid thinking enabled by diverse interactive views. One of our current research efforts aims at supporting the fluid creation of GeoTime scenarios within GeoTime, eliminating the need for external plug-ins or laborious wizards. Another promising effort under investigation is to automatically



extract geo-temporal metadata from Sandbox content using focused computational linguistics extraction functions and so more easily be able to use evidence in time and space. Mixed initiative interaction techniques would help deal with dirty data including missing values and incorrect or ambiguous extraction results.

## 6 LESSONS LEARNED

Initially, the two junior analysts analyzing the scenario integrated nSpace and GeoTime with a top down approach due to the abundance of textual data available in the VAST dataset. GeoTime was thus used as a tool for generating evidence for a pre-postulated assertion rather than a tool that could make inferences and provide hypotheses. Thus, data would not be input into GeoTime unless a specific question that may have involved spatial and temporal relationship visualization needed to be answered. For example, in order to determine whether or not the scandal in Alderwood was an affair between a married city council member and Laurel Sulfate, the analysts mapped the meeting locations and times in GeoTime to see if supporting evidence could be found. However, it was not until the pattern between the out of town city councillors and the phone calls was discovered using GeoTime that the analysts realized it could be a useful tool for bottom up analysis as well. GeoTime is as capable of generating hypotheses through analysis, as it is capable of providing evidence for pre-conceived hypotheses. It was then realized GeoTime could have been utilized more effectively to investigate the VAST dataset than it actually was.

When importing evidence or conclusions from GeoTime into the Sandbox, screenshots from GeoTime acted as image objects. It would be an advantage and more functional if clicking on the screenshot from GeoTime in the Sandbox would open the GeoTime scenario in order to observe the evidence in more detail. Other specialized analytical tools could be integrated with the Sandbox in a similar way.

One obstacle that was encountered during the analysis was the question of whether to account for the possibility of red herrings. Because the tools extracted and assembled the massive data quickly and easily, the possibility of red herrings arose and further analysis was therefore conducted in order to safeguard against that possibility. Perhaps, too much analysis was done. The tools were able to find patterns that were not placed in the data set intentionally by the contest scenario writers. These patterns were used as evidence and impeded what could have been a more fully sound analysis. An important lesson taken from this aspect of the competition was the need for proficiency in weighing the relevance of found patterns and observations. It is perhaps possible to enhance something like the diagnosticity of evidence technique for assertions with a relevancy assessment to minimize the chance of following the wrong thread.

## 7 CONCLUSION

The VAST analysis contest demonstrated that nSpace and GeoTime can effectively enable novice analysts to take on real-world analysis problems and work them to a successful conclusion using their full reasoning and imaginative capabilities, and not be bogged down in rigid workflows or difficult to use tools. Advanced computational linguistics, massive data and visualization techniques were harnessed to advantage for the analyst.

The VAST contest was treated by the nSpace and GeoTime research teams as a valuable opportunity to test the products by having fresh junior interns, who had not previously used the tools, attempt to solve a representative real world problem. A significant number of both minor and major issues, ideas and improvements were identified and many resolved and implemented. These improvements have resulted in much more capable nSpace and GeoTime products which are now being used by the intelligence community. Our participation in the contest also generated numerous new ideas for the future.

## REFERENCES

- [1] Aarseth, P., M. Deligonul, J. Lehmann, L. Nezda and A. Hickl., "TASER: A Temporal and Spatial Expression Recognition and Normalization System", Proceedings of the 2005 Automatic Content Extraction Conference (ACE 2005), Gaithersburg, MD.
- [2] Bringsjord, S., Arkoudas, K., Clark, M., Shilliday, A., Taylor, J., Schimanski, B. & Yang, Y., "Reporting on Some Logic-Based Machine Reading Research". in Etzioni, O., ed., Machine Reading: Papers from the AAAI Spring Symposium (Technical Report SS-07-06), pp. 23-28, and also [www.cogsci.rpi.edu/research/rair/slate](http://www.cogsci.rpi.edu/research/rair/slate).
- [3] Card, S., MacKinlay J., and Shneiderman B., Readings in Information Visualization, Morgan Kaufman, 1999.
- [4] Freitag, D., M. Blume, J. Byrnes, R. Calmbach and R. Rohwer, "A Workbench for Rapid Development of Extraction Capabilities", In Proc. International Conference on Intelligence Analysis, 2005.
- [5] Jonker, D., W. Wright, D. Schroh, P. Proulx and B. Cort, "Information Triage with TRIST", In Proc. International Conference on Intelligence Analysis, 2005.
- [6] Kapler, T., and W. Wright, "GeoTime Information Visualization", IEEE Symposium on Information Visualization, 2004.
- [7] Moldovan, D., S. Harabagiu, R. Girju, P. Morarescu, F. Lacatusu, A. Novischi, A. Badulescu and O. Bolohan, "LCC Tools for Question Answering", Proceedings of the TREC-2002 Conference, NIST, Gaithersburg, 2002.
- [8] Morse, E., M.P. Steves, J. Scholtz, "Metrics and Methodologies for Evaluating Technologies for Intelligence Analysts", In Proc. International Conference on Intelligence Analysis, 2005.
- [9] Scholtz, J., "Beyond Usability: Evaluation Aspects of Visual Analytic Environments", IEEE VAST Symposium, 2006.
- [10] Ware, C., "Information Visualization – Perception for Design", Academic Press, 2000.
- [11] Whittenton, J., Editor, "Defense Intelligence Integrators Guide (DIIG)", version 1.2, February, 2007.
- [12] Wright, W., D. Schroh, P. Proulx, A. Skaburskis and B. Cort, "The Sandbox for Analysis – Concepts and Methods", ACM CHI 2006.