# SemanticSpy: Visualizing Semantic Data in a Media-Rich Environment

Amit Mathew
Department of Computer Science
University of Georgia
mathew@cs.uga.edu

**Abstract**

*In this paper we describe SemanticSpy, a tool for tracking suspects and their activities around the globe. Using an RDF data store, SemanticSpy is able to bring semantic information to life, in the hopes that it will help law enforcement detect trends in a suspect's behavior and give them the power to avert the next major crime or terrorist attack. SemanticSpy utilizes a media-rich environment that can display images, 3D models, documents, and audio to give the user the ability to perceive information in various ways. We describe the architecture used for the development of SemanticSpy and describe its advantages over other visualization systems. Finally, we describe other potential uses for SemanticSpy.*

## 1. INTRODUCTION

In this age of international terrorism and global drug smuggling rings, it is apparent that there is a need for law enforcement agencies around the world to track the movement and activities of suspects around the world. Semantic data is a powerful way of enhancing this information by creating relationships and metadata for the suspects being tracked. The problem becomes one of information overload. The tracking of a single suspect can generate a large amount of semantic data, and important details can be lost in a large volume of text.

SemanticSpy offers a way to harness this information and visualize it in a variety of ways. Phone conversations can be played, photographs of a suspect's meeting can be viewed, and a 3D model of the car a suspect is driving can be shown. Taken separately, these details may be trivial. But with enough semantic-enhanced data, it is possible that the little details can be pieced together to prevent an act of terrorism. For example, if government agencies had a tool that showed images of the similar-sized airplanes that the 9/11 hijackers flew in the months prior to 9/11, that disaster might have been averted.

## 2. FUNCTIONALITY

SemanticSpy offers several different components that aid in the tracking of suspects. These are the suspect view, the activity view, and the global view. The suspect view, shown in the upper right corner of Figure 1, shows a list of suspects that can be selected for tracking. In this figure, the suspects are from the list of the FBI's ten most wanted

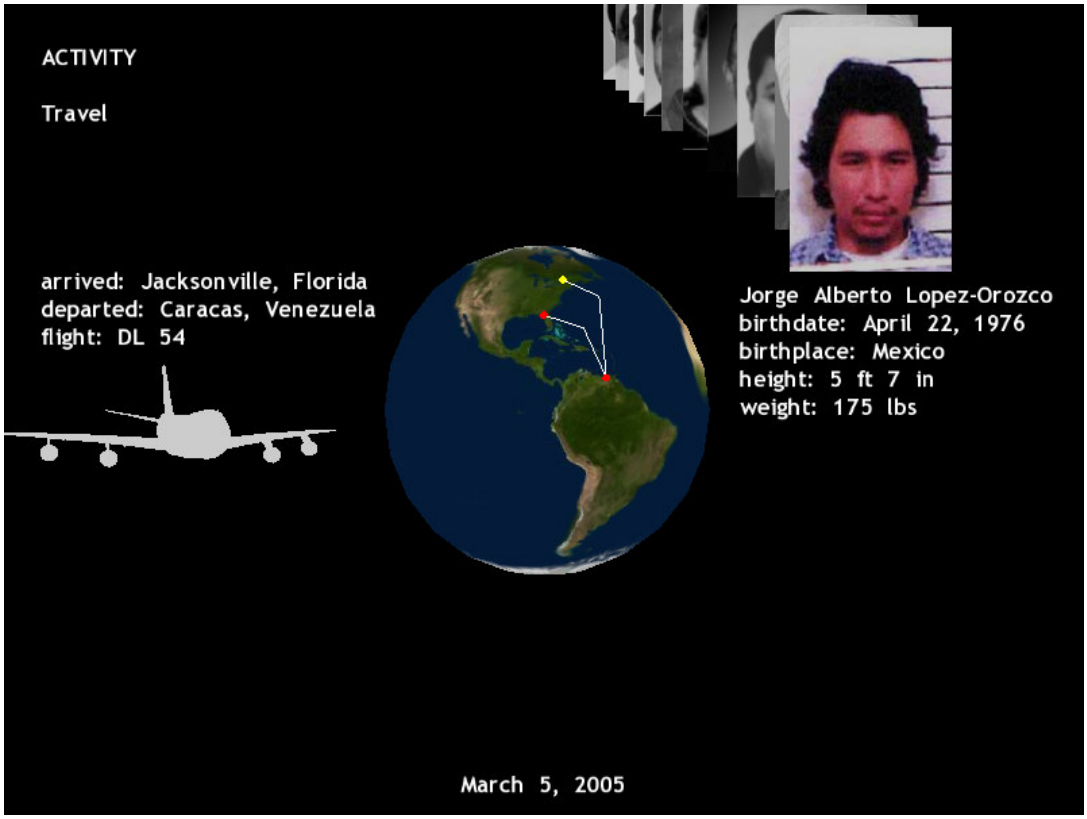

**Fig. 1: A screenshot of SemanticSpy**

criminals [6]. Notice that the suspect's names are accompanied by a photograph of the suspect, which allows security officials to see a suspect and possibly recognize the suspect in another picture. The lists of suspects can be scrolled through, with the currently track suspect shown on the top. Once a suspect is selected, the user can choose a date and see the suspect's activities on that date. In this figure, March 5, 2005, is chosen. When a day on which a suspect has been involved in some activity, information is shown in the activity view on the left side.

The activity view shows the type of activity that occurred and its associated metadata. Currently, SemanticSpy supports three types of activities: travel, meeting, and phone call. The different types of activities can have different types of metadata. For example, a travel activity can detail the type of travel used, whether it be by air, sea, or land. If the suspect traveled by air, an airline, flight number, and an image or model of the airplane they traveled in can be shown. If the suspect traveled by a rental car, information about the rental car company can be shown. Each travel activity is associated with two geographic points, the point of departure and point of arrival. These points are shown in the global view, shown in the center of Figure 1.

The global view is made of a three-dimensional sphere, textured with an image of the earth displayed with terrain. The global view is helpful in several ways. It allows users to absorb information "at a glance." For example, in Figure 1, a user can find the exact

details of each of the suspect's trips, but from the image it is immediately clear that a suspect traveled first from somewhere in North America to somewhere in South America, and later back to North America. This easy access to general information prevents the user from being overloaded by information, and allows them to more likely notice the information that can stop an international criminal. The currently active points are highlighted in red, while points of previous travel are highlighted in yellow. Plotting points is not restricted to travel. It can be used to for communicating any relationship between two points, such as a phone call between parties in two different countries. Another helpful feature of the map is that it shows the topographical information of the Earth. This is helpful because a user may notice a suspect traveling to suspicious terrain, such as a desert.

The other types of activities supported by SemanticSpy are meetings and phone calls. Meetings are displayed with a single geographic point. Meetings can be accompanied by a photograph of the meeting if one is available, as well as relevant metadata, like who was present at a particular meeting. It may be beneficial to the user to investigate other people involved in a meeting. The other type of supported activity is a phone conversation. Phone calls can be accompanied by an audio recording of the conversation and information about the phone call, like who was called and from what number. All these elements can be used to understand the motivations of a suspect's actions.

Although only three activities are currently used, it is simple to add more activities to SemanticSpy. SemanticSpy can be configured to accept new types of activities and to show additional information that may be pertinent to a particular type of activity, like an image or an audio clip.

## 3. ARCHITECTURE

SemanticSpy's architecture is separated into two major components: the front-end visualization system and the back-end RDF data store. The visualization system is powered by MAGE [1], a high-performance game and virtual reality engine, with support for simulation and visualization tasks. MAGE supports the viewing of 3D models, 2D images, and audio, while still maintaining interactive framerates. MAGE also has a scripted material system, making it easy to update suspect's photographs or 3D models on-the-fly.

SemanticSpy is written in C++ to provide adequate performance for real-time interaction. It is currently only runs on Windows, but that will change as soon as MAGE is ported to other platforms. SemanticSpy utilizes MAGE's functionality to provide a large number of features with a minimal amount of code.

BRAHMS [3] is used as SemanticSpy's RDF data store. BRAHMS is an in-memory system that is very fast, making it an appropriate choice for a real-time visualization system. Currently BRAHMS is queried on application startup exclusively, but future work can allow dynamic queries using BRAHMS to support user requests for SemanticSpy.

Semantic information is loaded from the input RDF file by BRAHMS at runtime. The RDF file must contain certain attributes to work properly, but SemanticSpy allows the

input of arbitrary metadata for suspects and their activities. Once the RDF file is loaded, BRAHMS creates a snapshot file that is optimal format for BRAHMS to load RDF content. The snapshot data is then loaded into SemanticSpy for programmatic access.

The points on the globe are determined by GPS coordinates in the RDF input file. The GPS coordinates are converted to Cartesian coordinates and plotted on the globe. The lines connecting the points are the midpoint of the two points, with the height determined by the distance between the two points.

Media is loaded dynamically from the RDF file based on certain predetermined tags. For example, an <audio> tag triggers loading an audio file from disk. The RDF doesn't provide the path to the audio file directly, but provides a named instance of the audio source in the MAGE scripting system. Similarly, a <model> triggers the loading of a 3D model.

## 4. ADVANTAGES

The two major features of SemanticSpy that differentiate it from existing tools are its use of semantic data and its rich multimedia environment. It is important to understand the value of these two features and how they benefit the user.

Semantic data has many advantages over data stored in a traditional database, but the two key aspects are its emphasis on relationships and its use of metadata. Relationships are important in many aspects. One is the idea of human networks and its relationship to organized crime, terrorism, and other illicit activities. Suspects participating in illegal activities often are part of social networks that need to be analyzed to determine leadership and level of involvement. Furthermore, relationships are important when trying to anticipate a suspect's actions. For example, a suspect may have a relationship to a chemical plant which may be legit (if the suspect was a chemical engineer) or not (the suspect is investigating using hazardous materials for malevolent purposes).

The second important aspect of the Semantic Web is metadata. Metadata is often treated as a side effect or unnecessary detail in traditional settings, but for intelligence agencies, metadata can make the difference in catching a criminal. Metadata allows a user to detect patterns that would otherwise go unnoticed. For example, the fact that two seemingly unrelated people are traveling in a car may be innocuous, the fact they are traveling in the two cars with exact same make and model may be important.

## 5. RELATED WORK

Much of the work done in Semantic Web visualization today are in the area of graph visualizations, especially of large ontologies. However, there exists some work dealing with Semantic information in a spatial context. Themescape [5] provides topographical maps that show the relations between different types of documents. Ghinea et al [2] use spatial data in the context of the human body to study back pain. GIS-type applications have also been developed before the invention of the Semantic Web. Eick [4] developed a method for visualizing communication networks using a three-dimensional globe.

## 6. FUTURE WORK

The development of SemanticSpy has opened up the possibility for productive future research. SemanticSpy can be extended with more multimedia, like video, and more types of activities. Other activities that can be tracked are job interviews, emails and other documents, and social gatherings. Also the global visualization system can be extended to support zooming and satellite views.

A GUI system is a future feature that would allow navigating dates and a large number of suspects and metadata much easier. This will be added after MAGE provides support for GUI systems. In the future, SemanticSpy may also add the ability to do dynamic queries during operation. This is currently not possible because BRAHMS does not support queries, but this can easily be changed as soon as querying is added to BRAHMS.

The idea of SemanticSpy can be extended to other domains that require geospatial tracking of entities described by semantic data:

*Employee travel*
Large corporations often have many of their employees traveling to different parts of the world. It is beneficial for a company to keep track of where exactly their employees are at a given time. This can be important in the case of emergencies.

*Transportation companies*
A transportation company may need to know where all their assets are at a given time. Also, the companies would be interested in the metadata for each of their assets. For example a shipping company may wish to know where all their shipping containers are and the occupancy of those containers. If the containers with smallest load were going to the furthest destinations, the company may wish to revise their shipping system.

*Military*
Military organizations can use SemanticSpy to track their assets on a battlefield. Modern conflict can span continents and it is difficult to visualize troop and equipment concentration. Furthermore, the military can use the visualization system to view enemy troop movements. Visual patterns of movement may allow the military to predict future attacks.

*Computer networks*
SemanticSpy can be used to provide real-time information of network traffic around the globe. Network specialists can monitor traffic patterns and can track suspicious traffic that may indicate the presence of a virus or denial-of-service attack. Malicious activity is especially suited for geographic visualization because attacks tend to commonly occur in certain areas of the world.

## 7. CONCLUSION

SemanticSpy allows law enforcement officials to view information about suspects in new and powerful ways. SemanticSpy uses a multimedia-rich environment to capitalize on

the human's powerful senses and pattern recognition abilities. Furthermore, SemanticSpy is an experiment in visualizing semantic data in a new way, by creating domain-specific visualization tools that empower the user to make connections.

## 8. REFERENCES


1. Amit Mathew and Leonidas Deligiannidis, "MAGE-VR: A Software Framework for Virtual Reality Application Development". *Proc. of The 2005 International Conference on Modeling, Simulation and Visualization Methods* (MSV'05), pp. 191-197, June 2005.
2. Gheorghita Ghinea and David Gill. "Using Scalable Vector Graphics and Geographical Information Systems in a Back Pain Study." Visualizing the Semantic Web. Springer 2002. pp. 180 – 197.
3. Maciej Janik, Krys Kochut. "BRAHMS: A WorkBench RDF Store And High Performance Memory System for Semantic Association Discovery", Fourth International Semantic Web Conference ISWC 2005, Galway, Ireland, 6-10 November 2005.
4. S.G. Eick, Aspects of Network Visualization. *Computer Graphics and Applications* 16(2), 69-72. 1996.
5. ThemeScape Product Suite. Cartia Inc. http://www.cartia.com
6. FBI's Ten Most Wanted: http://www.fbi.gov/mostwant/topten/fugitives/fugitives.htm