GeoWorlds: A Geographically Based Situation Understanding and Information Management System for Disaster Relief Operations[1]

Murilo Coutinho, Robert Neches, Alejandro Bugacov, Vished Kumar, Ke-Thia Yao, In-Young Ko, Ragy Eleish, Sameer Abhinkar
USC Information Sciences Institute
Marina del Rey, CA
{coutinho, rneches} @ isi.edu

Abstract
GeoWorlds is an extension to digital library technology to support situation understanding and management tasks. Using Humanitarian Assistance and Disaster Relief as a test bed application, GeoWorlds demonstrates feasibility of relating geographical information to a corpus of "other" information in documents. The function of the system is to help a user understand facts and events in relation to space and time. It allows users to take a set of documents, relate them to places and times relevant to their contents, and provide a visual environment for presenting and exploring those relationships.

1. Introduction

A broad class of applications requires tools for situation understanding and management to help users assess complex situations and develop responses. To address such needs, GeoWorlds integrates a rich set of components and services that give end users a wide range of capabilities to explore, cross-correlate and interrelate geographic, spatial and visual information, data and text/document views.

The GeoWorlds project at USC/ISI has focused upon helping humanitarian assistance disaster relief (HA/DR) organizations with rapid-response mission requirements. Such organizations need to be able to quickly stand up rescue teams backed by the information, materiel, and support services available for the region in which the disaster is developing. Within this scenario, GeoWorlds' main goal is to help rescue teams understand facts and events in relation to space and time in a collaborative fashion. The system provides users with the ability to closely link geographical and documentary information about a region to help assess the impact of a disaster, identify assets and partners that can help respond to the problem, and help organize the application of those assets. GeoWorlds capabilities include:

1. *The ability to find and display both geographic and document-based views of an area in which a disaster is playing out*. These include the ability to combine related information into clusters using information analysis clustering services, as well as user-defined clusters and compare the different classification results found;

[1] This work is sponsored by DARPA under Fort Huachuca Contract No. DABT63-95-C-0101.

2. *The ability to find and display both geographic and document-based information in a bi-directional linked fashion*. Users can select a sub-region of a custom geographic display and get shown not just what documents are known about that sub-region, but how they are organized and what characterizes their contents. Conversely, users can act on the document organization (clusters), select documents and images and see them displayed together with related maps, images and documents that help understand the geographic context of the selected set of objects;
3. *The ability to select a document and/or image and see plotted on a map the locations that they reference*. This includes the ability to extract address information from semi-structured documents and retrieve the appropriate maps that reference the extracted locations;
4. *The ability to register services at run-time*. Users can register services and helper tools by either modifying the service's source code to implement a specific service interface, or write a wrapper that implements the desired interface and can be used to exchange information with the service. GeoWorlds provides one interface for each general class of services that are available in the system (see section 2 for more details.) These services help users refine and structure the set of available information in order to support a particular task at hand. These tasks can initiate automated analysis tools which can draw upon both kinds of information (e.g., geographic and document-based information);
5. *A collaborative framework where users can exchange information and collaborate over tasks*. The system supports both asynchronous and synchronous collaboration environments. The collaboration takes place among users logged in a common session, and collaborative tasks within this session are performed by marking, annotating, and editing information in their views;

The above set of capabilities implemented in GeoWorlds is drawn upon several in-house tools, as well as tools developed at other research institutes that are collaborating with the USC/ISI Distributed Collaborative Enterprises group. These include the UC Santa Barbara Alexandria Digital Library project, the UC Berkeley Digital Library Initiative on Multi-Valent Documents, the University of Arizona Artificial Intelligence Laboratory, the University of Illinois at Urbana-Champaign Digital Library program, and the National Center for Supercomputing Applications.

2. Architecture Overview

2.1 Functional-Level Description

At the functional level, all services and capabilities available in the current system can be grouped into six main categories: customized repository, characterization and partitioning, processing-extraction-piping, interactive analysis and view control, collaborative environments, and information visualization. These six categories represent our current working hypothesis about how best to classify and describe the component set of modules that comprise situation understanding and management systems. We believe that a wide range of applications such as HA/DR, can be constructed given interoperable components from these classes. In general, these systems require components to:

- Dynamically store/retrieve sets of information available about the disaster;
- Process and analyze the retrieved information and prune out unnecessary or redundant information;
- View and analyze the information;
- Annotate and share information between rescue teams deployed during the incident;

In GeoWorlds we have defined an API for each category of components. Users can add new components to the system by first implementing the interface corresponding to the category classification the components belong to, and then, register the components using the service registry tool. The current version of the system has at least one service instantiating a component of each of the six main categories. In the following paragraphs we will succinctly describe the six main categories along with the current available GeoWorlds services instances associated with them.

Customized Repository: Users can work with these components to readily create a customized repository of information restricted to those supporting a particular task at hand. Components of this class include UCSB Alexandria Digital Library (gazetteer service), USC/ISI Data Warehouse, UIUC HDF[2] (structured data storage) and USC/ISI DASHER[3] information space analysis tools (managing repositories and gazetteer access.)

Characterization and Partitioning: Users can use these components to browse sets of documents and, in the process, develop characterizations of them. In doing so, they can simultaneously specify a taxonomy of topics of interest to them and use that taxonomy to describe and organize the documents. Components of this class include University of Arizona neural-network clustering tools, and USC/ISI DASHER document partitioning services [7].

Processing-Extraction-Piping: Users can use these components to perform specific, commonly used services such as extraction (or addition as annotations) of information about places and times referenced in documents, and link them to geo-locations. Components of this class include all wrappers, scripts and routers available in the GeoWorlds system and used to pass information back and forth between system components with different data format requirements.

Interactive Analysis and View Control: Users can use these components to interact with documents and the extended information about them. This interaction may consist of exploring relationships between the documents and lay out presentations that will help rescue teams understand possible connections between the events and facts taken place during the disaster. Components of this class include UC Berkeley Multi-Valent Documents tools for marking / annotating.

Collaborative Environments: Components of this category are responsible for supporting synchronous and asynchronous collaboration between teams working on different aspects of the HA/DR tasks at hand. GeoWorlds currently uses the NCSA Habanero

[2] Hierarchical Data Format
[3] Defense Acquisition Services for High Performance Electronic Commerce.

collaboration framework to support synchronous collaboration among rescue teams, and DASHER clip book servers to support asynchronous collaboration among users.

Information Visualization: Components of this category are responsible for supporting user interaction with documents and geographic information using highly graphical techniques to represent relationships between them. GeoWorlds currently uses ESRI's Arcview as its default map viewer and DASHER as its default document viewer.

2.2. Implementation-Level Description

The GeoWorlds system essentially implements a client-server architecture. The server side is a standalone Java application while the client side is embedded in the Habanero Collaborative framework and therefore supports synchronous collaboration between registered applications.

The server side supports services and tools that are either not currently collaborative or not intended to become so. The current version of GeoWorlds has ESRI's Arcview registered as the GIS component. This CoTs desktop GIS software is a Windows application and is impossible to port it as is into Habanero's collaborative framework. GeoWorlds publishes a "GISViewer" interface that all registered GIS components should adhere to. Any application implementing this interface can then be plugged into GeoWorlds as a registered GIS component and be dynamically loaded when deemed appropriate. We have written a Java wrapper that implements this interface to enable communication between Arcview and the GeoWorlds system. The Java wrapper forwards all requests to Arcview via Windows DDE (Dynamic Data Exchange) calls, invoking scripts written in Avenue, Arcview's scripting language. The Avenue scripts are bundled in an Arcview project that is loaded in when the GeoWorlds server launches Arcview for the first time. The GeoWorlds server is also a DDE Server and thus is able to accept requests from Arcview, which then acts as a DDE client.

ALOHA [2], distributed and supported by the National Safety Council, is currently registered as GeoWorlds' plume analysis component. Based on weather conditions, the type of chemical, and the physical dimensions of the source, ALOHA computes the maximum spread of the plume and also provides concentration levels within that extent as a function of time. As it provides very little support for communication with external applications, we first export the analysis results to an external file and then import that file into GeoWorlds to visualize the impact of the incident. The plume is imported into Arcview as a shape file and intersected with other layers to extract a list of places affected by the plume. We are currently looking at adding other plume analysis tools like Slab [5], as registered analyzer components.

GeoWorlds exposes a "DataWarehouse" interface as part of its API. Any data source that has to be integrated into GeoWorlds register wrappers that implement this interface. These data sources can be present both locally or distributed on a network. All communication with remote databases is achieved through Java's Remote Method Invocation. In the current version we have created wrappers for UC Santa Barbara's Alexandria Digital Library and USC/ISI's own Data Warehouse. Thus the users of

GeoWorlds can transparently query multiple data sources and get back results independent of their structure and format.

GeoWorlds uses USC/ISI's DASHER [7] Information space Analysis tools as its document management component. DASHER helps users quickly gather knowledge or expertise on or about a given region of space or topic by managing collections of documents obtained via web searches and other data sources (such as registered data repositories and experts created customized information spaces.) In contrast to numerous services available for information search and retrieval, DASHER helps users make sense of their data, by providing tools for characterizing, partitioning and visualizing documents and document collections. It combines multiple methods like natural language text extraction and ontology based categorization to help users understand their data sets. GeoWorlds communicates with DASHER using Java's Remote Method Invocation.

The GeoWorlds Client program connects to the Server via RMI and has access to all information related to the currently active incident. It is used to start collaborative sessions in order to share and exchange information between applications. USC/ISI's DASHER Information Space Analysis tool and UCB's Multi-Valent Document application have been ported to the Collaborative Environment. In addition to these the framework supports applications like whiteboard, chat, and audio chat.

3. Demonstration scenario

To illustrate how the system works, imagine that people witness an explosion that releases Phosgene, a heavier-than-air toxic gas, in a populous area. They inform an "Emergency Operations Center" (EOC), giving them a rough idea of the incident's location and nature. The EOC register the incident with GeoWorlds. GeoWorlds queries registered data-warehouses and retrieves thumbnails and meta-data. Based on that, the user selects the appropriate information (street maps, aerial imagery, etc.) for download and viewing in one of the registered GIS components. A registered plume analyzer is then launched to compute the spread of the toxic material taking factors like the wind direction, speed, and the weather conditions as input. This result is then overlayed on the downloaded map data to determine the geographic extent of the areas affected by the plume.

The GIS component then extracts the list of places (e.g., schools, hospitals, airports, tourist attractions, etc.) affected by the incident, and sends them to the Information Space Analysis component of the GeoWorlds system. Meanwhile, various teams like the "Emergency Medical Response" (EMR), "Transportation", etc. have already been deployed to the scene of the incident. Each of these teams or their corresponding "Anchor desks" run a copy of the GeoWorlds Client, through which they can share and exchange data collaboratively. Each client has full access to all incident-related data gathered at the EOC, and subsequently attempts to gather as much information on their specialized topic as possible, focusing in the incident region. In the scenario, when asked by the EMR for help in "moving perishable medical supplies", the "transportation" team invites the EMR team to join their session to execute the search collaboratively. They query the GTE yellow pages and extract the addresses of companies dealing with "refrigerated trucks" in the region of interest. The EMR team then transfers this data to their copy of the information space and saves it for future reference thus building on their existing knowledge base. This information is also mapped onto the GIS Viewer to help evaluate its geographic relationship to the affected area.

4. Conclusion and Future directions

The GeoWorlds system, as it stands today, is best understood as an initial experimental effort. Although not ready for general use, we believe it demonstrates the desirability of integrating a range of functionality. In particular, it shows the utility of having the ability to coordinate geographic information with document collections, and to collaboratively view and discuss the related information. The current version of GeoWorlds follows a component-based approach to enable continuous increase of functionality and portability, but is too closely tied to specific components. Our goal is to have a system where new components that implement any of our category class interfaces, can be added to the system in a plug-and-play fashion and where components within the same category can be swapped at run-time (facilitating comparison between equivalent components.) In order to achieve this goal, we are working on extending the current system to support a dynamic distributed service architecture based on JINI, as well as implementing a security model for components/services registration and use.

5. Acknowledgements

We would like to thank many people. In particular, we would like to thank Mike Sullivan (BBN/GTE), Hsinchun Chen, Dorbin Ng (UAz), Robert Wilensky, ByungHoon Kang (UCB), Larry Jackson, Terry McLaren, Al Woss (NCSA), Terence Smith, Jim Frew, Greg Janee (UCSB), Bill Pottenger, Nancy Yeager (UIUC), John Wilson and Ed Young (USC). We would also like to thank Ron Larsen for his support and encouragement.

6. References

[1] M. Coutinho, R. Neches, A. Bugacov, K.-T. Yao, V. Kumar, I.-Y. Ko, R, Eleish, *GeoWorlds: A Geographically-based Information System for Situation Understanding and Management*, TeleGeo '99, Lyon, France, May 6-7, 1999 (http://www.insa-lyon.fr/Labos/LISI/telegeo).

[2] Aloha (Areal Locations of Hazardous Atmospheres), Chemical Air Dispersion Model, National Safety Council (www.nsc.org/ehc/cam/alohapg.htm).

[3] Chabert, A., Grossman, E., Jackson, L., S., Pietrowiz, S., R.., and Seguin, C., *Java object-sharing in Habanero*, Commun. ACM 41, 6 (Jun. 1998), Pages 69 - 76.

[4] Chen, H., Schuffels, C., and Orwig, R., *Internet Categorization and Search: A Machine Learning Approach*, Journal of Visual Communication and Image Representation, Volume 7, Number 1, Pages 88-102, 1996.

[5] Ermak, D., L., *User's Manual for Slab: An Atmospheric dispersion model for denser-than-air releases*, June 1990, Lawrence Livermore National Laboratory and Lakes Enviromental (www.lakes-enviromental.com).

[6] Frew, J., Freeston, M., Kemp, R., Simpson, J., & Smith, T. (1996). *The Alexandria Digital Library testbed*. D-Lib Magazine (July/August 1996).

[7] Neches, R., Abhinkar, S., Hu, F., Eleish, R., Ko, I.-Y., Yao, K.-T., Zhu, Q., Will, P., *Collaborative Information Space Analysis Tools*, D-Lib Magazine, October 1998.

[8] Phelps, T., A., and Wilensky, R., *Toward Active, Extensible, Networked Documents: Multivalent Architecture and Applications*, Proceedings of Digital Libraries '96.

[9] http://heineken.gsfc.nasa.gov/eosgis/eghome.html