Accessing the Alexandria Digital Library from Geographic Information Systems

Dan Ancona, Jim Frew, Greg Janée, Dave Valentine
Alexandria Digital Earth Prototype (ADEPT)
University of California, Santa Barbara
805.893.7684
{da, frew, gjanee, valentine}@alexandria.ucsb.edu

ABSTRACT

We describe two experimental desktop library clients that offer improved access to geospatial data via the Alexandria Digital Library (ADL): **ArcADL**, an extension to ESRI's ArcView GIS, and **vtADL**, an extension to the Virtual Terrain Project's Enviro terrain visualization package. ArcADL provides a simplified user interface to ADL's powerful underlying distributed geospatial search technology. Both clients use the ADL Access Framework to access library data that is available in multiple formats and retrievable by multiple methods. Issues common to both clients and future scenarios are also considered.

1. INTRODUCTION

The geospatial information community has long desired easier and more integrated tools for accessing geospatial data from inside other software environments. The early promise of convenient access to geospatial data over the Internet has blossomed, with many occasionally confusing web-browser-based clients having perhaps a few too many features (ADL itself has produced more than one web-based client). With the recent development of two new clients for the Alexandria Digital Library, we have demonstrated that GIS and visualization applications can access and utilize diverse information stored in digital libraries. The ArcADL client is integrated with ESRI's industry standard ArcView desktop Geographic Information System (GIS), and the vtADL client allows access to the library from the open-source Virtual Terrain Project's Enviro terrain visualization application.

Access to geospatial clearinghouses from GIS applications began with the Z39.50 GEO search profile [1] based on the FGDC Content Standard for Digital Geospatial Metadata [2]. ESRI's Geography Network [3], another example, targets geospatial information cataloged in specific metadata formats. However, not all geospatial or georeferenced information is produced with complete, standardized metadata that GIS applications can exploit. ADL's metadata-agnostic approach allows items cataloged with non-FGDC elements (such at MARC or Dublin Core) to be discovered using searches with geospatial constraints. The ArcADL client can then download and display appropriately-formatted items in ArcView.

vtADL builds on experiments with ADL and "digital Earth" visualization, and techniques borrowed from spatial hypertext systems[4]. Like Shen et al. [5], our work addresses the connection between visualization environments and digital libraries. However, instead of proposing an entirely new protocol, our client-library connection is handled by simple and transparent XML-over-HTTP messaging, taking advantage of the superior scalability of the ADL architecture vis-à-vis harvest-based architectures.

2. ADL QUERY, RESULTS AND ACCESS

Both of the newly developed clients rely on a simple XML-over-HTTP API to communicate with the ADL middleware, which forms the backbone of the distributed Alexandria system. The middleware is responsible for distributing queries over multiple repositories, aggregating the responses, and returning results in a variety of formats.

Complex boolean queries can be formed according to the novel ADL bucket architecture [6]. *Buckets* are an abstraction and aggregation mechanism supporting queries against spatial, temporal, numeric, textual and other types of heterogeneous metadata. After the structured query is passed to the middleware, a result set ID is returned, and the status of the search can be monitored by the client. As results are retrieved from the distributed collections, three standardized metadata views of each result set item can be retrieved: a *bucket* view (the item's metadata mapped to standard buckets), a *browse* view (for obtaining thumbnails and other synoptic representations of the item), and an *access* view (information about all the ways an item can be retrieved and accessed). The item's native metadata (e.g. FGDC) can be retrieved as well. In this work, the access view (the ADL Access Framework) is the critical component, as it provides the linkage between the client, the item's metadata, the item itself, and any services associated with the item. [7]

3. TWO DESKTOP LIBRARY CLIENTS

3.1 ArcADL

The ArcADL client was implemented inside of ESRI's ArcView GIS. ArcADL's objective is to retrieve information from ADL without leaving the GIS environment. ArcView's support of Microsoft's Visual Basic for Applications (VBA), and VBA's straightforward support for XML-over-HTTP transactions, made VBA a natural choice for implementing ArcADL.

Figure 1 shows a screenshot of ArcADL. The small window to the left includes a simple interface to search across three buckets: item locations in space and time, and keywords and other subject-related text. The geographic search is initialized with the coordinates of the currently displayed ArcView map.

For testing purposes, the collections to search across can be manually selected. This feature will be phased out when the middleware supports automatic selection of collections by querying and ranking over collection-level statistics such as spatiotemporal coverage histograms. In the results window to the right, thumbnail images, the primary title for the item and the various formats the item can be accessed in are listed. (These access forms are a rendering of the access report provided by the middleware.) Se-

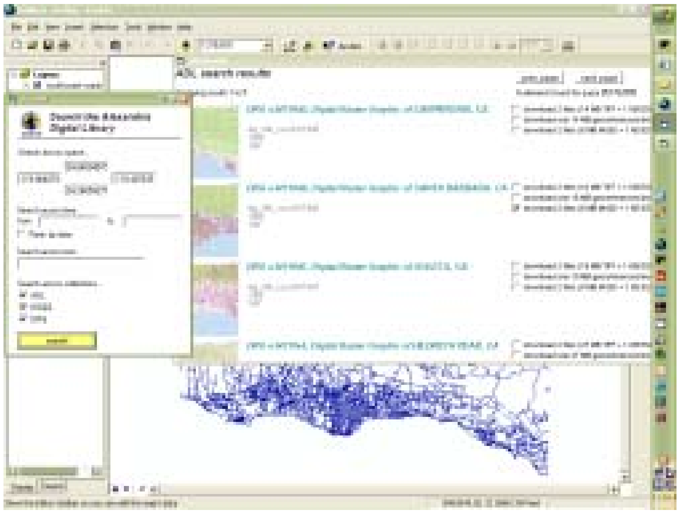

Figure 1 – ArcADL screenshot.

lected objects are downloaded from the appropriate repository and automatically added as layers into the active map.

3.2 vtADL

A similar interface is being developed between ADL and the Virtual Terrain Project's Enviro ROAM algorithm based terrain browser.[8] This cross-platform system doesn't use a scripting language, but the modular architecture of the VTP code has allowed straightforward inclusion of simple ADL features, and Microsoft .NET will allow reuse of the Visual Basic code developed for ArcADL.

The workflow for the current pre-alpha vtADL is simple: after loading a terrain model, the user right-clicks over the terrain to select an area, and then right-clicks again to select "Search ADL" from a pop-up menu. The query is formed behind the scenes in the same manner as described above, and results are fetched and then rendered inside the environment. The optimal method of rendering the results has yet to be determined, but some experiments have already been performed regarding the display of these results in whole-earth and terrain-level views. Labeling features using the ADL gazetteer [9] has also been implemented. Facilitating searches across other buckets is under development, perhaps using 3D widgets (such as a zoomable, interactive 3D timeline widget.)

4. CONCLUSIONS & FUTURE WORK

These two clients demonstrate the ease of writing desktop clients that access the Alexandria Library system. Other possible target environments for client work include scientific modeling environments such as MATLAB and Mathematica. Any software environment that researchers use for analyzing geospatial data is a potential ADL client.

As mentioned above, the ADL Access Framework allows clients to determine whether a library item can be accessed in a manner (format, protocol, etc.) that the client supports. However, clients cannot currently *query* the library by file format or access service; they can only filter access reports post-query. Adding such a query capability is a challenge, as it is difficult to transform the typically rich, hierarchical access descriptions into searchable quantities. Access descriptions may also include conversion services and equivalent formats, which further complicates matters.

The functionality of these two clients may converge with the release of ArcView version 9, which will include a whole-earth visualization module (ArcGlobe). With this release, automatically inserting data into a 3D environment should be straightforward. However, vtADL will still have higher raw visualization performance and will be a viable platform for digital library research (e.g., manipulation of search terms via 3D widgets.) Regardless of what underlying technologies are employed, the development of an integrated, functional, visually rich and user-friendly search environment for geospatial data – the "digital Earth" - continues.

5. ACKNOWLEDGMENTS

This work is supported by NSF award IIS98-17432.

6. REFERENCES

[1] Nebert, D. Z39.50 Application Profile for Geospatial Metadata or "GEO", Version 2.2. 27 May 2000 <http://www.blueangeltech.com/standards/GeoProfile/geo22.htm> (09 Feb 2004)

[2] Federal Geographic Data Committee. FGDC-STD-001-1998. Content standard for digital geospatial metadata (revised June 1998). Federal Geographic Data Committee. Washington, D.C.

[3] Geography Network http://www.geographynetwork.com (09 Feb 2004)

[4] Ancona, Dan, Michael Freeston, Terence R. Smith and Sara Irina Fabrikant. "Visual Explorations for the Alexandria Digital Earth Prototype" *Visual Interfaces to Digital Libraries* 2002, pp. 199-216.

[5] Shen, Rao, Jun Wang & Eduard A. Fox. "A Lightweight Protocol between Digital Libraries and Visualization Systems" *Second International Workshop on Visual Interfaces to Digital Libraries*, July 2002.

[6] Janée, G., and J. Frew. The ADEPT Digital Library Architecture. *Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)* (Portland, Oregon; July 14-18, 2002): 342-350.

[7] Janée, G., James Frew, and David Valentine. Content Access Characterization in Digital Libraries. *Proceedings of the Third ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2003)*, Houston, TX, May 27-31, 2003.

[8] Duchaineau, Mark et al. ROAMing Terrain: Real-time Optimally Adapting Meshes. *IEEE Visualization '97 Proceedings*

[9] Hill, L.L., J. Frew, and Q. Zheng, Geographic Names. The implementation of a gazetteer in a georeferenced digital library. *Digital Library*, 1999, 5(1)