



White Paper

**Collaborative Web-Enabled Data Distribution –
The MetroGIS Experience**

August 2002

By

Alison Slaats

Manager, MetroGIS DataFinder *and*
Web Applications Developer
Metropolitan Council

And

Mark Kotz

GIS Database Manager
Metropolitan Council

Acknowledgements

The authors (Alison Slaats, project technical lead and Mark Kotz, database administrator) recognize the following people from the MetroGIS community for their contributions to the technical design of the DataFinder Café: David Arbeit, MN Land Management Information Center; Chris Cialek, MN Land Management Information Center; Will Craig, University of MN – CURA; Michael Dolbow, Metropolitan Council; Steve Fester, Metropolitan Council; Rick Gelbmann, Metropolitan Council; Joerg Kessler, Metropolitan Council; Randy Knippel, Dakota County; Robert Maki, MN Department of Natural Resources; Curt Peterson, Ramsey County; and Kent Tupper, Dakota County.

The authors would also like to acknowledge the professional work and dedication provided by the following people from Syncline, Inc.: Adam B. Hocherman, Jon Stavis, Josh Lieberman and Saul Farber.

This project would not have been possible without the outstanding leadership and vision of Randall Johnson, MetroGIS Staff Coordinator.

Table of Contents

Acknowledgements.....	2
Table of Contents.....	3
Executive Summary.....	4
1 Introduction.....	5
2 Background on MetroGIS.....	5
3 DataFinder.....	5
3.1 DataFinder Catalog.....	6
3.2 DataFinder Search.....	7
3.3 DataFinder Café.....	7
4 Development of DataFinder Café.....	7
4.1 Background.....	7
4.2 Project Definition.....	7
4.3 Café Application.....	8
4.3.1 Overview.....	8
4.3.2 Café Client Application.....	9
4.3.2.1 Starting the Café.....	9
4.3.2.2 Viewing Data with the Café.....	10
4.3.2.3 Downloading Data from the Café.....	11
4.3.3 Café Server Application.....	12
4.3.3.1 Cascading & Extracting Web Map Service.....	13
4.3.3.2 Remote Nodes.....	13
4.3.3.3 Security Module.....	14
5 DataFinder Café Evaluation.....	14
5.1 Café Status.....	14
5.2 User Feedback.....	14
5.3 Additional Success Measures.....	15
5.4 Next Steps.....	15
5.4.1 Café Enhancements.....	15
5.4.2 Café Extendibility.....	16
6 Project Costs.....	16
7 What have we learned?.....	17

Executive Summary

MetroGIS is a geospatial data sharing collaborative representing over 250 local units of government and other organizations in the seven county Minneapolis-St. Paul Metropolitan Area.

DataFinder (www.datafinder.org), the MetroGIS data sharing web site, provides access to more than 100 metadata records, which can be queried through the online GIS data catalog or the MetroGIS DataFinder node of the National Geospatial Data Clearinghouse. Many data sets are freely downloadable. The site also features several ESRI ArcIMS map services, with web clients to view data directly within the web site.

During 2002, MetroGIS worked on enhancing the data distribution capabilities of DataFinder, with three goals in mind:

1. Provide a more robust method for downloading data. The client for data distribution would have functionality to allow users to subset or clip datasets using an ad-hoc user defined area, or an existing feature, such as a city or school district boundary. Users would also be able to select attributes to download, and choose from a number of different geospatial data formats.
2. Implement Web Mapping Services (WMS) that are Open GIS Consortium (OGC) compliant. This is an emerging method of serving and using GIS data. MetroGIS wants to ensure that the map services available are accessible by clients that adhere to this industry standard.
3. Accomplish these tasks in a truly distributed environment. MetroGIS stakeholders use software from several different vendors. While a goal of MetroGIS is to distribute many organizations' datasets via a centralized web site, we do not want to require 1) centralized data storage or 2) specific vendor software for data storage. The system should allow for either centralized or distributed data storage and serving.

We hope to share through this paper some real world experiences that other agencies and collaboratives can directly apply to their work in sharing data, applying OGC standards, and implementing it all in a distributed and non-vendor specific environment.

1. INTRODUCTION

A primary goal of MetroGIS is the efficient and effective distribution of GIS data via the Internet. This paper discusses the context and process for development of a distribution mechanism, the DataFinder Café, which meets the needs of the MetroGIS community. First, the MetroGIS organization and the existing DataFinder web site are described, followed by a description of the Café application. Next, an evaluation of the Café is provided along with future plans for the system. Finally, the costs and a few of the lessons learned during this project are shared.

2. BACKGROUND ON METROGIS

MetroGIS is an innovative, regional geographic information systems initiative serving the seven-county Minneapolis-St. Paul metropolitan area. This 3000 square mile area is home to 2.6 million inhabitants, and includes 190 cities and townships, 59 school districts and 39 watershed management organizations. Initiated in 1995, MetroGIS is a voluntary collaboration of local and regional governments, with partners in state and federal government, academic institutions, nonprofit organizations and businesses. MetroGIS's mission is to "provide an ongoing, stakeholder-governed, metro-wide mechanism through which participants easily and equitably share geographically referenced data that are accurate, current, secure, of common benefit and readily usable". The desired outcomes of MetroGIS include:

- Improve participant operations
- Reduce costs
- Support cross-jurisdictional decision making

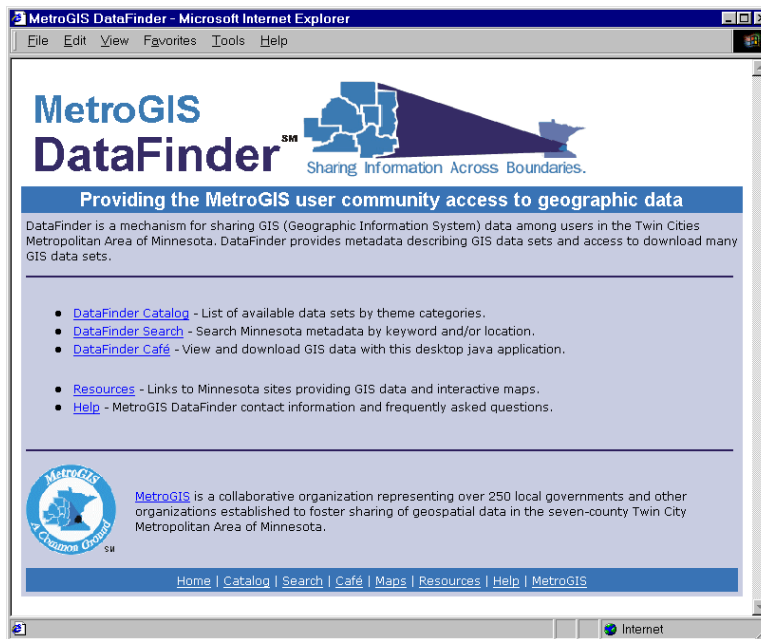
More specifically, the goal is to institutionalize sharing of accurate and reliable geospatial data between the data user and producer communities. MetroGIS has focused on meeting common business information needs for GIS data defined by the MetroGIS community. The result has been the definition and standardization of several regional GIS datasets that meet the needs of multiple organizations. The MetroGIS forum aids both data producers and users by streamlining the dissemination of current data and simplifying the search for it. Additional background on MetroGIS's organizational structure, funding sources and participants can be found on the MetroGIS web site (www.metrogis.org).

3. DATAFINDER

The DataFinder web site implements MetroGIS's data sharing mission. This site (www.datafinder.org), launched in 1998, has changed significantly over time along with organizational growth and technological advances. It now has three functional pieces for data sharing: the DataFinder Catalog, the DataFinder Search and the DataFinder Café. Central to these functions are geospatial metadata records.

DataFinder was primarily created to distribute datasets defined through MetroGIS's common business information needs process, but MetroGIS also encourages the web site to be used for

distribution of other spatial datasets relevant to the Twin Cities area. Any organization which produces geodata for the seven county Minneapolis-St. Paul Metropolitan Area may contribute metadata and/or datasets to be included on the DataFinder web site. Currently 14 different organizations contribute towards a total of over 120 metadata records on the DataFinder site. More than 80 of those datasets are available for download. The numbers of organizations, metadata records, and downloadable datasets are expected to increase over time, especially with the addition of the Café tool.



3.1 DataFinder Catalog

Launched in April 1998, a theme-based GIS data catalog was the first functional piece of DataFinder to be built. It is essentially a list of available metadata records documenting GIS datasets in the Twin Cities metropolitan area. The catalog initially used theme categories developed by MetroGIS, but has since changed to use spatial data categories being promoted by the Federal Geographic Data Committee, which evolved from the ISO metadata standard work (ISO/Technical Committee (TC) 211 Draft International Standard 19115). The catalog also includes some subcategories developed within the Minnesota GIS community.

Users of the catalog can scroll or link down to a theme of interest (e.g. administrative and political boundaries) and view a list of available metadata records (e.g. county boundaries and school districts). Users can then click to view a complete metadata record. Then, depending upon the distribution policy for that particular dataset, the user can either download it directly, or learn how to obtain the data from the source organization.

The DataFinder Catalog is still the most frequently used DataFinder tool, being a particularly attractive option to those who know what they are looking for and where to find it. However, as the number of metadata records on DataFinder has increased, finding a particular metadata record has become slightly more difficult. For this and other reasons, in January 2001 MetroGIS included a metadata search mechanism on the DataFinder site.

3.2 DataFinder Search

As a node of the National Spatial Data Clearinghouse, Datafinder serves metadata records to users who search the Clearinghouse with keywords and geographic extents of datasets. Users can also access this functionality directly on the Datafinder site, with an interface tailored to search for metadata hosted on the MetroGIS node, or other statewide nodes in Minnesota.

A contractor created the search interface for the Minnesota Land Management Information Center (LMIC), a state government entity. LMIC then leveraged their investment by sharing the interface with the Metropolitan Council for use with the MetroGIS DataFinder site. The search tool on DataFinder averages over 50 searches a week resulting in 70 metadata records being examined.

3.3 DataFinder Café

The DataFinder Café is the latest tool to be added to the DataFinder web site, and its development is the focus of the remainder of this paper. It is an application that allows users to browse geographic datasets using a map, and then download those datasets.

4 DEVELOPMENT OF DATAFINDER CAFÉ

4.1 Background

In July 2001, the seven metro counties and the Metropolitan Council began working on an agreement to further a MetroGIS goal of online distribution of the counties' parcel datasets to public sector and academic interests in the metro area. Interested parties are required to complete license agreements with the counties in order to obtain the data free of charge. This agreement was an enormous step for MetroGIS, because the counties' parcel GIS data are a top priority of MetroGIS stakeholders.

However, this also identified some technical and functional problems with the MetroGIS DataFinder web site. For one, DataFinder's existing mechanism for online data delivery could not handle partitioning large datasets into smaller areas of interest. Instead, the general practice had been to make entire datasets downloadable directly from the metadata via an FTP link. Since many parcel data users were only interested in a portion of the entire seven county metro area, this was not very practical. Furthermore, no mechanism existed to restrict downloads to licensed users; only the unrestricted datasets were available for direct downloads.

Based on these problems, and other feedback about DataFinder, a project scope was defined for the creation of a new tool.

4.2 Project Definition

A team of 15 MetroGIS participants from 6 organizations was assembled to ensure that the scope of the project was adequately specified, and that the resulting new tool for DataFinder would

meet the community's data access needs. Members of this ad-hoc team helped during different stages of the project.

MetroGIS hosted a Parcel Data Users Forum in April 2001, and used feedback from the forum and other sources to define a reasonable set of requirements for the project. This scope of work definition stated that the application should satisfy the following objectives:

- Automate processes involved in distributing geospatial data via the Internet to the MetroGIS data user community. Include a security module that provides the ability to distinguish between data requests that require licensing and/or fees, and data requests that require neither.
- Provide the user with the ability to obtain data for self-defined geographic areas of interest (i.e., clipping out a subset of the dataset).
 - Allow users to select among available geographic datasets and attributes (or fields).
 - Provide users with multiple geospatial data format choices.
- Identify the possibility of integrating Open GIS Consortium's proposed Web Map Services standard into the solution.
- Provide a method for data providers to share their data via an application, but keep the data on their servers (i.e., create a distributed environment for data sharing).

In July 2001, MetroGIS published a Request for Proposals (RFP) for the development of an Internet-enabled data distribution mechanism that encompassed and expanded upon the requirements listed above. The full "scope" section of the RFP can be viewed at http://www.metrogis.org/data/datafinder/data_distribution_rfp_scope.pdf. Seven proposals were received and two companies were interviewed. Syncline, Inc. of Boston, Massachusetts, was selected for the project. Syncline is a company that specializes in products and solutions that allow for the visualization, analysis, management and sharing of spatial data via the Internet.

Over several months, Syncline and MetroGIS further refined the project requirements, including two meetings between Syncline and the MetroGIS ad-hoc group. This was extremely important because it identified as many potential problems as possible in advance of actual application development. Two documents resulted: 1) a Functional Requirements Document (FRD) and 2) a System Design Document (SDD). The FRD can be viewed at: http://www.metrogis.org/data/datafinder/ieddm_func_req.pdf

4.3 Café Application

4.3.1 Overview

The FRD and the SDD provided the blueprint for implementing the DataFinder Café. The Café can be logically separated into two parts – the server-side piece and the client piece. The client piece is the user application that allows for data browsing and data download. The server-side piece consists of modules that process requests from the client. These processes include providing graphics of GIS layers to the client, compiling data requested for download and changing the format of the data as necessary. Both the server and client pieces of the Café are described below.

4.3.2 Café Client Application

To achieve the functionality requested by MetroGIS, Syncline recommended using a more sophisticated client than would be possible in a web browser. The decision was made to use Java Web Start for client application development. Java Web Start provides the ability to develop an application that is as flexible and robust as a desktop application, but yet also has the capabilities of web-based applications.

For example, Java Web Start allows for the user to drag and drop data layers into the map, to draw ad-hoc polygons, and to utilize “right-click” mouse functions. While such functionality could be provided through a web-based client with Java applets, it would take considerably longer to load each time – a condition unacceptable to the MetroGIS project team.

For the user, this means that to run the DataFinder Café, they are first required to perform a one-time install of Java Web Start, and then install the Café application. While the initial setup of the Café may take the user a few minutes the first time, the ensuing load time and performance is better than a web-based client, because the actual application resides on the user’s computer.

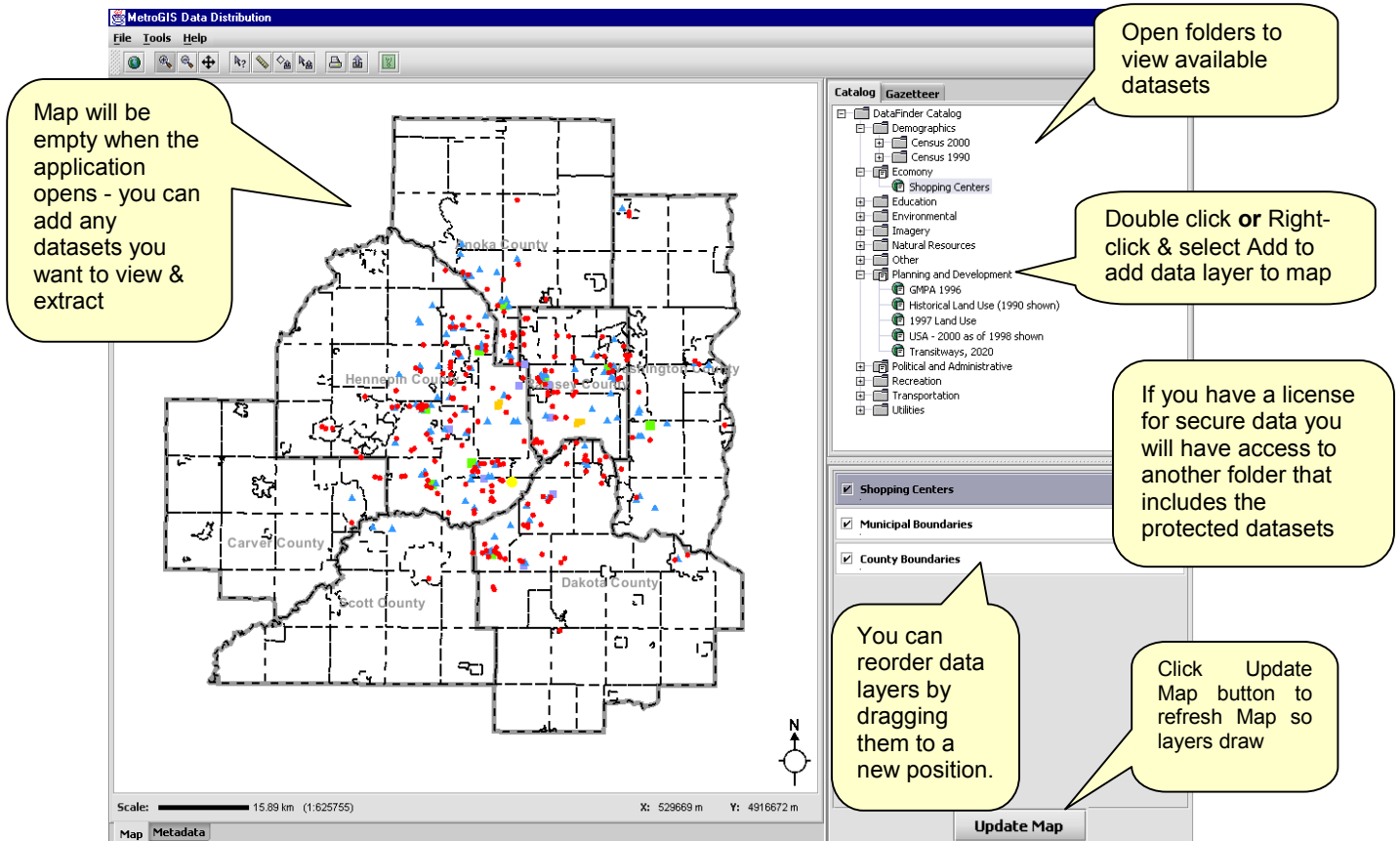
The next time the user starts the Café application, Java Web Start checks to see if a newer version of the Café is available at the server. If so, the newer components are automatically installed. This ensures that the user always has the most up-to-date functionality in their client, and also allows for MetroGIS to seamlessly deliver periodic updates to the Café. A description of the Café client functionality follows.

4.3.2.1 Starting the Café

When starting the Café application, users are given the option of entering a user name and password, or logging on as an anonymous user. Anonymous users have access to all unlicensed data. Licensed users of the parcel data enter their user name and password to be able to view and download parcel data and any other datasets they may be licensed for in the future.

4.3.2.2 Viewing Data with the Café

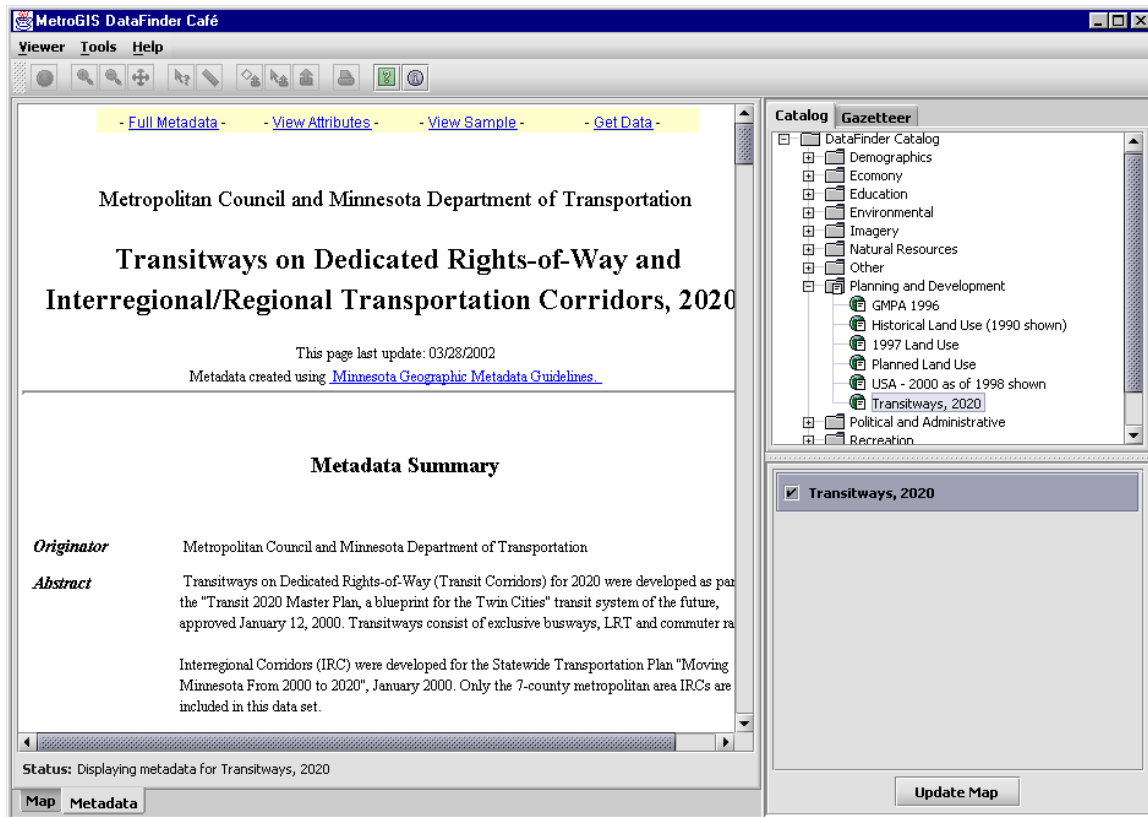
When entering the Café, the user is first presented with a blank map frame on the left side of the window. On the right side, the user is shown a list of geospatial theme categories, the same ISO derived categories used in the DataFinder Catalog. The user may then click any theme category to see the available datasets within it, and add individual datasets to the map for viewing. With over 60 datasets already in the Café, the user can browse many combinations of datasets that suit their individual needs. Each theme has data-specific, predefined symbology assigned.



It is important to note that the focus of the Café is data distribution and not data viewing. Thus, from the start of the project, it was agreed that the viewing capabilities would be limited to those necessary for data acquisition purposes. Therefore, no changes to symbology are permitted by the user. This decision has allowed MetroGIS to put some clear boundaries on the scope of the Café project, preventing potentially significant “scope creep” from the myriad of functionality associated with map symbology and rendering.

The DataFinder Café includes a Gazetteer function, allowing users to zoom to predefined areas of interest such as counties, municipalities, school districts or watershed management areas. Standard zoom and pan map navigation functions are also available, along with a feature identification tool and a measurement tool.

A MetroGIS policy requires all datasets available through DataFinder to be accompanied by complete metadata. Users can access metadata records by selecting a dataset, and clicking the metadata tab. The MetroGIS metadata standard follows the Minnesota Geographic Metadata Guidelines, which are essentially a subset of the FGDC standard. Once the ISO standard is formally approved, the state of Minnesota will revise their metadata guidelines (with MetroGIS involvement) and MetroGIS will adopt the new standard.



4.3.2.3 Downloading Data from the Café

Three methods exist for defining the geographic extent when downloading data from the Café. Users can draw their own ad hoc polygon, they can extract based upon the extent of a feature in an existing dataset (e.g., a city boundary), or they can extract based upon the current viewing extent in the application. Additionally, when users choose the option to extract based on an existing feature, they will also be able to buffer this area by typing in a distance and selecting units, or by selecting from a pre-defined list of buffer distances. The other two extract options do not include this function.

When defining the geographic extract functionality, the design team decided that these extraction methods should not "clip" the features to the exact boundary specified. Instead, every feature that intersects the boundary should be included in its entirety. This method may deliver features that extend beyond the extraction boundary, but it also preserves attribute integrity. For example, an attribute that shows total population for a polygon feature is still accurate when that

entire polygon is delivered. This would not be the case if that polygon were to be cut in two by a clipping operation.

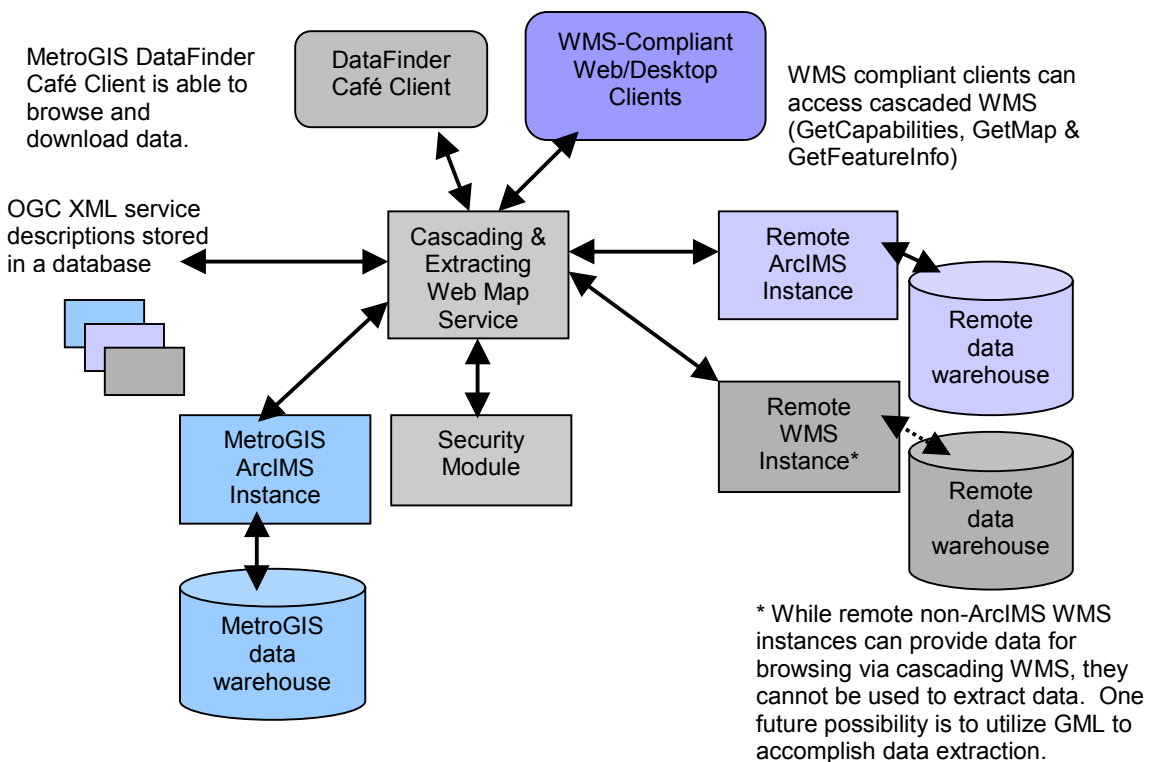
Once a user has defined their geographic extent of interest, the Café download wizard allows users to select the attributes (data columns) they wish to download. This is particularly useful for datasets with very large attribute tables, and for users that need only a few of those attributes.

Next, the user is able to select from a variety of spatial data downloading formats. The Café employs FME's Spatial Direct product to perform the data conversion functionality. Currently four spatial data output formats are available from the Café: ESRI Shape format (.shp), MapInfo format (.tab), Bentley Microstation format (.dgn), and FME transfer format (.ffs).

All datasets are downloaded in a compressed (.zip) file format. The compressed file contains the actual datasets, and an HTML metadata file for each dataset downloaded. It may also contain a disclaimer, a use/maintenance reminder for the dataset, or another type of document pertinent to the dataset.

4.3.3 Café Server Application

The server side of the Café application is what drives much of the Café client's functionality. The server side system is shown in the diagram below.



4.3.3.1 Cascading & Extracting Web Map Service

The main part of the system on the server side is the Cascading & Extracting Web Map Service. This service can access map services from a number of different sources and combine them together to send them to the Café client.

For example, if a user wants to draw shopping centers and rivers in their Café map, the client sends a request to the server. The cascading map service then finds those datasets within one or more map services, and requests a graphic of those map layers. Next, the service combines the requested graphics into *one* graphic, demonstrating the power of a “cascading” service. That single graphic is then sent back to the requesting client – showing the user their requested datasets in the map. This series of transactions occurs each time the user refreshes the map by panning, zooming or adding new data layers. The cascading map server ascertains which map layer comes from which map service by accessing the OGC-compliant XML stored in a database that exposes this information.

A similar series of events occurs when the user requests a specific extent of data to be downloaded. First, the Café map client sends the request to the extracting map service module. That module then uses the XML stored in a database to find which map service contains the data, and then requests the data to be extracted by the user extent. In response, the map services that host the data return zip files, containing the requested data, to the extracting service. If more than one zip file is returned, the extracting service compiles the data into *one* zip file to return to the user for download. If the user has requested a data format other than ESRI shapefile, the format of the data is changed (before zipping) by running a side request to the SpatialDirect software. The communication to the SpatialDirect software is carried out in a separate service, communicating with the server using Java Message Service (JMS).

The ability to extract data is not included in the OGC WMS standard, prompting Syncline to amend the XML files with custom vendor tags specific to this function. Additional custom tags also indicate the location of the “notice” file that is included in the zip file along with the metadata HTML file.

The Café system is compliant with the OGC web mapping service (WMS) standards. When the cascading web service combines graphics together, it also sends the graphics and any other requests, such as tabular information, back to the client in WMS-compliant XML. The benefit of using this standard is that it can be understood by not only by the DataFinder Café, but also by any WMS-compliant client. As this standard becomes more widely used, DataFinder map services will be available via many more software packages. For example, the DataFinder map services may be viewed via the WMS viewer at www.wmsviewer.com.

4.3.3.2 Remote Nodes

As the diagram shows, the system can have inputs from multiple map services. The datasets that are shown in the DataFinder Café client may come from many different organizations and many different servers, but this is not evident to the user. This satisfies a MetroGIS requirement for a

truly distributed system: centralized data distribution through a one-stop shop, without the need to centralize data storage. It ultimately reduces data redundancy and ensures up-to-date datasets.

Currently, the remote nodes in the system can serve data to the Café client through either ArcIMS map services or OGC-compliant WMS map services. However, for data to be downloadable, the map service *must* be an ArcIMS map service. Other remote WMS instances can provide data through the cascading WMS for browsing only. One possible remedy for this problem is to utilize GML for data extraction. However, at the time of development, MetroGIS felt that the GML standard and the related Web Feature Service (WFS) were not stable enough to integrate into the solution. This is one of the enhancements that MetroGIS hopes to implement in the future.

4.3.3.3 Security Module

The security module of the system regulates which users have access to which datasets. The users and their access permissions are stored in an Oracle database, which the security module checks each time a user logs into the Café client. The user's access rights determine which datasets are shown in the Café catalog. A web-based administrative console allows for easy updates and edits to this security system. The system is flexible enough to define accessibility rules for entire organizations while still allowing each user from that organization to have individual user names.

5 DATAFINDER CAFÉ EVALUATION

5.1 Café Status

As of August 2002, the DataFinder Café has 60+ datasets available for download via the Café client. Those datasets can be browsed through a WMS using the Café client or other WMS-compliant clients. Data producers can provide their datasets for hosting on the MetroGIS server or they can use a remote node to the Café if they are running ArcIMS. When this functionality was put in place, in April 2002, MetroGIS conducted a feedback survey to ask users for their input on the DataFinder Café.

5.2 User Feedback

In order to obtain feedback on the DataFinder Café, over 80 individuals were encouraged to test the application, and each tester was encouraged to fill out an online survey to comment on and rate several aspects of the application. A summary report was created of the test process, objectives, participants, comments and rating results, and actions taken to address issues identified by the beta testers. This report can be viewed online (http://www.metrogis.org/data/datafinder/beta_summary_final.pdf). The comments submitted by users both during and after the testing period indicate that the new application is generally being well received. A selection of these positive comments is as follows:

- "Very easy to download & install."

- "Very intuitive design; pays off if you are an ArcView user, but could figure it out without difficulty."
- "I like the fact that when you download the parcel data you receive metadata as well as use and care information. Gives people a second reminder."
- "This site is incredible, what a fantastic application, great job. We will be setting the link on our site for our customers to use it to get data."
- "Overall, this is what I envisioned MetroGIS would become when first involved back in 1995-6 – GREAT WORK!"

In addition to these comments, several good ideas were provided for ways to enhance the Café. These suggestions form the basis for the next steps MetroGIS will be taking to improve and enhance the DataFinder Café.

5.3 Additional Success Measures

In addition to the input received in the feedback survey, MetroGIS data producers and the State of Minnesota have demonstrated significant interest in this project. MetroGIS acknowledges this as a strong endorsement that the Café is successfully meeting data distribution needs.

Many MetroGIS data producers have shown their interest in serving their data on DataFinder and the MetroGIS server, or serving as remote nodes to the Café. This could help data producers eliminate many of their day-to-day data bundling and distribution tasks. MetroGIS feels that the Café offers a flexible solution for participants, fostering increased participation from MetroGIS organizations.

In addition to interest from data producers, some of the most encouraging support for DataFinder Café has come from the Land Management Information Center (LMIC), a division of Minnesota Planning which is the state's planning agency. LMIC has recently received a grant from the state Technology Enterprise Board to implement a statewide extension of Café functionality. The new project, called GeoIntegrator, seeks to expand the Café concept, allowing state agencies and local governments outside the Twin Cities Metro area to share data from a single portal, speeding up data access and decreasing data redundancy.

5.4 Next steps

5.4.1 Café enhancements

While the current version of the DataFinder Café is well received and fills an important role, some important future steps are already planned. The enhancements detailed below stem from feedback provided during the beta test conducted in April 2002.

One of the problems with the current Café is created by two conflicting functions of the application. One function allows users to download only datasets that are drawn on the screen. The other function sets scale dependencies for large datasets, so that a user would not be able to draw an entire county of parcel data, for example. The purpose of this latter feature is to prevent users from getting trapped in an enormously long draw time. Under this scenario, users have no

way to download an entire county parcel dataset, because they can't draw it. Since downloading a full county is expected to be a common task, one of the most important near term enhancements pursued in 2002 will be implementing separate viewable and extractable extents for datasets.

Another important enhancement will be a projection/coordinate system conversion utility, tentatively proposed for 2003. Currently, the seven counties in the metro area, as well as most local government units, use 5 different state sanctioned county coordinate systems. State and regional users typically use the UTM zone 15 coordinate system. While the Café uses UTM coordinates, the ability to re-project remote map services and to re-project data for download will increase its usefulness to the MetroGIS community.

An additional technical step involves providing additional output data formats. ArcInfo export (e00) file format, AutoCAD DWG, and DXF are proposed for 2002 and 2003 enhancement projects.

Another enhancement that will extend data download from remote nodes would be to integrate the use of the Open GIS Consortium's Geographic Markup Language (GML) data format into the Café. This would allow the DataFinder client to extract data from non-ArcIMS map servers such as Minnesota MapServer or Intergraph's GeoMedia Web. This would significantly help MetroGIS reach its goal of being vendor neutral. Supplemental Funds for this enhancement may be available from LMIC and via a NSDI Web mapping Services Grant. The scope of the project is still evolving but is expected to be undertaken in 2003.

5.4.2 Café Extendibility

On an administrative and policy level, MetroGIS is working closely with the State of Minnesota to assess the benefit of expanding the Café model statewide. MetroGIS is very interested in leveraging its investment in this technology by sharing it with partner organizations. The contract with Syncline permits MetroGIS to share the complete application with any government unit in Minnesota, provided that sharing it furthers MetroGIS goals and objectives.

6 PROJECT COSTS

The project to develop the DataFinder Café has several direct and indirect costs. The direct costs include SpatialDirect software (\$22,050) and Syncline application development (\$76,145). The cost for the enhancements proposed for 2003 (GML, additional format support, and projection conversion) is approximated at \$60,000. A NSDI Web Mapping Service grant has been received and the state of Minnesota and the City of St. Paul have expressed interest in assisting with the funding needed to accomplishing these enhancements.

The Café could not have accomplished without investments that had been made to provide the infrastructure upon Café rests; investments made by GIS department at the Metropolitan Council, where the Café is hosted. These investments include the initial development of DataFinder (about \$100,000+/-), a web-server class computer, ESRI ArcIMS software, Oracle database software and skilled technical staff support. If this system were to be replicated

elsewhere, these investments would need to be factored into the overall cost of the project. More information about the initial investment in DataFinder can be found at <http://www.metrogis.org/data/datafinder/index.shtml>.

Future, anticipated system support costs include maintenance fees on ArcIMS (\$2,500, annually), Spatial Direct (\$5,500/annually), and troubleshooting support from Syncline as needed for the Café application. In addition, continuous staff time will be necessary to maintain the Café and its supporting software and hardware (.5 FTE for a skilled GIS/Web professional has been allocated).

7 WHAT HAVE WE LEARNED?

MetroGIS considers the DataFinder Café project to be a major success. For a reasonable investment, it is providing desired functionality that does not exist elsewhere, and it is attracting partners for subsequent enhancements. Meanwhile, we learned several things that may be of interest to others undertaking similar projects.

- A fundamental component to the project's success was the concerted effort for nearly a year prior to the start of the project to develop broad consensus on the need for the project and a clear vision of the functionality desired. The use of the ad hoc team of GIS professionals from several MetroGIS organizations was integral to defining a clear definition of the desired functionality.
- Another fundamental component to the project's success was a strong emphasis and accompanying effort on quality documentation to insure expectations were clear all involved. The creation of a functional requirements document and a system design document was the key to maintaining an understanding of the project objectives and scope. This documentation was written through a mutual effort between Syncline and MetroGIS staff, and it defined the expectations for both Syncline and MetroGIS. This proved its worth when beta testers requested the ability to change symbology as an enhancement. We had already addressed this issue in the documentation phase and knew it would be out the project's scope.
- Similarly, it was beneficial to treat this project as simply "Phase 1" of a long-term vision of data distribution. Symbology and other enhancements could be implemented in a later phase. Thus, we were able to implement our initial goals without sacrificing future possibilities.
- The use of standards is very important for creating a scalable and long-lasting product. For example, the Café uses Web Map Services (WMS), and we are confident that as more WMS clients become available, more organizations will be able to leverage the Café's services. In addition, the Café requirement of standardized metadata for all datasets means that users are assured of complete, consistent documentation for all data. They can assess if a dataset meets their needs quickly, because information is easy to find and easily comparable between records.
- While standards are important to follow, the needs of data users and providers must also be considered if the project is to be a success. For example, while adhering to the WMS standard will prove to be valuable in the long run, most of the data producers in MetroGIS

who have map service capabilities use ArcIMS. Thus, having the system use ArcIMS services directly made it easier for these producers to be involved in the project and to include their remote data. More organizations could contribute their data from a remote node with less up-front work on their part – a definite incentive to participate.