PROJECT TITLE: Automated Data Integration in Support of *The National Map*

PRINCIPAL INVESTIGATORS:

Principal Investigator:	**E. Lynn Usery**
Title:	**Research Geographer**
Affiliation:	**USGS, MCMC**
Address:	**1400 Independence Road**
Telephone/fax number:	**573-308-3837**
Email:	**usery@usgs.gov**
Co-Principal Investigator:	**Michael P. Finn**
Co-Principal Investigator:	**Wm. Steve Helterbrand**

DURATION OF PROJECT: 3 years

FY 2003 FUNDING REQUEST: Fully burdened dollars $ 307,521

STUDY AREA/ REGIONS: St. Louis, Missouri; Atlanta, Georgia

RESEARCH THEMES: Data integration, raster/vector, layer/feature approaches

KEY WORDS: data integration, geometric matching, feature, National Map

ABSTRACT:

Data integration is a significant problem for *The National Map*. We propose to examine data integration from a layer-based approach, developing a conceptual framework based on resolution, geometric accuracy, and topological consistency, and apply it to five of *The National Map* data layers, digital orthoimages, elevations, land cover, hydrography, and transportation. From the experience with the layered approach and the data developed, we will examine a feature approach to integration based on a model previously developed and implemented as a feature library. We anticipate significant results leading to an automated approach based on the conceptual framework, the empirical results, and the use of those in metadata to drive an automated process.

BACKGROUND:

Integrating the eight data themes for *The National Map* (USGS, 2001) is a significant problem consisting of several components including differences in datums, projections, coordinate systems, data models, spatial and temporal resolution, precision, and accuracy. Design of an automated approach to insure integration is a long-term research objective and a staged approach is proposed. Each of the integration issues requires specific consideration and only by resolving each issue can one develop an overarching solution to *The National Map* integration problem. For example, problems of datums, projections, and coordinate systems are resolvable through exact or approximate mathematical transformations provided the source and target systems are known. Welch and Homsey (1997) document a USGS datum integration problem between existing topographic maps, Digital Line Graphs (DLG), and Digital Orthophotographic Quadrangles (DOQ's). Integrating data across different vertical and horizontal datums and meeting National Map Accuracy Standards requires all data be transformed to a single datum (Welch, 1995). If the source datum, projection, and coordinate system are unknown, then an approximate solution may be the best result achievable.

Significant research has examined the data integration problems between remotely sensed images and GIS (*e.g.*, Ehlers *et al*, 1989; 1991; Cowen *et al*., 1995; Star *et al*., 1997). For data models, the difference in the raster and vector geometric models mean inherent integration problems and loss of accuracy between them (Abel and Wilson, 1990; Flowerdew, 1991). However, the problem is greater since within the vector model, topology may be explicitly stored in cellular or algebraic (simplicial) modes or not stored at all and the accuracy of vector lines varies significantly (Egenhofer and Herring, 1991). Within the raster model, square cells (pixels) can be assumed, but the size (resolution) of those cells varies from extremely coarse (30 m for elevation and land cover) to very high (0.3 m for ortho images). The type of values, categorical (land cover) and numeric or continuous (elevation and ortho images) of the individual pixels also affects methods and accuracy of data integration (Congalton, 1991). Given the plethora of the types of problems and permutations of those in integrating 8 data layers, the proposed work is to address one component of the integration problem as a step toward a comprehensive solution.

The University Consortium for Geographic Information Science (UCGIS) has identified data integration as both a long term (5-15 years) research challenge entitled Spatial Data Acquisition and Integration (http://www.ucgis.org/research98.html) and particular aspects of the problem, Geospatial Data Fusion, as a short-term (3-5 years) research priority (http://www.cobblestoneconcepts.com/ucgis2summer2002/researchagendafinal.htm). Data integration strategies and methodologies have not kept pace with recent advances in resolution of satellite sensors, radar and LIDAR technologies, *in situ* sensors using built-in Global Positioning System (GPS) capabilities, and other advances in geographic data collection and processing in geographic information systems (GIS). "It remains difficult to analyze even two spatial data sets acquired at different times, for different purposes, using different datums, positional accuracy (x,y,z), classification schemes, and levels of in situ sampling or enumeration precision." (Jensen *et al*., 1998, p. 1).

The word *conflation* is often used as a synonym for integration of multiple sets of spatial data from different sources (Saalfeld, 1988). Conflation may be used in the following ways: to transfer attributes from old versions of feature geometry to new, more accurate ones or to different geometries; to detect changes by comparing images of an area from different dates; or to automatically register one data set to another through the recognition of common features. Data integration methods are usually ad hoc, designed for specific projects involving a specific pair of data sets and of no generic value. A general theoretical and conceptual framework is needed to be able to accommodate at least five distinct forms of data integration (Jensen *et al.*, 1998, p 2):

1) in situ measurement-to-in situ measurement (calibration, adjustment, variance, etc.)
2) in situ measurement-to-foundation map (point-to-map; registration, verification)
3) vector-to-foundation map (map-to-map; vector segmentation scheme integration; different scales, different geographic coverage, etc.)
4) image-to-foundation map (image-to-map; for elevation mapping, map revision, etc.)
5) image-to-foundation image (image-to-image; involving different spatial, spectral, temporal, and radiometric resolutions)

The National Map requirements for data integration include all of the types listed above. However, in order to move forward we must restrict our scope to a subset of the data integration problems that are tractable and can be adequately investigated in the course of a 3-year project with the resources available. We will thus limit our approach to examining aspects of bullets 3 and 4 as they apply to specific layers of *The National Map*. Specifically, we will examine vector-to-foundation map with hydrography, transportation, and land cover layers (Cobb *et al*., 1998) and image-to-foundation map with ortho images, Digital Elevation Models (DEMs), hydrography, transportation, and land cover (Arnberg, 1981).

As in the UCGIS Research Agenda, we distinguish between data integration and data fusion. Data integration is the process of assuring consistency among various data elements in terms of geometric, topologic, and attribute accuracy and precision. Data fusion requires that the data elements be integrated but then creates a single composite dataset from the integrated elements. Commonly data fusion is used to combine different resolutions of remotely sensed data (Chavez, 1986; Welch and Ehlers, 1987). For example, an integrated dataset of hydrography and shaded relief insures that the streams of the hydrography dataset flow in the valleys of the elevation data. The two datasets may be in separate files, can be displayed together in softcopy or hardcopy formats, but maintain separable existences. A fusion of the shaded relief and hydrography datasets requires that the two sets be combined in some manner resulting in a single composite dataset and the two inputs are no longer separable. One method of fusing the hydrography and shaded relief datasets is to “burn” the vector hydrography locations into appropriate raster cell values of the shaded relief image. Once this occurs the two datasets are not separable since the shading cell values have been lost and replaced by the hydrography values.

The ideal data integration theoretical framework would be based on a feature-oriented data model (Usery, 1996a; 1996b) and would insure perfect feature-matching on a one-to-one basis. However, geographical data are rarely perfect and a framework is needed which allows integration of diverse data types in which some items remain unmatched and others are matched only with limited but specified confidence. The integration must handle horizontal (adjacency)(Chrisman, 1990), vertical (overlay), and temporal integration. Often feature-matching is the basis of the integration process requiring features to be identified in the two data sets to be integrated. While this is the ultimate goal, we propose a staged approach to achieve this framework with our initial work focusing on integrating data layers based on geometry and topology. This stage will be followed by a feature-matching integration approach based on a feature-oriented data model.

HYPOTHESIS/QUESTION:

1) Layer-based integration of geographic databases can be accomplished automatically using geometric and topologic constraints that are a function of accuracy for vector datasets and a combination of data resolution and accuracy for raster datasets.
2) The closer the two datasets are in resolution and accuracy, the easier the integration becomes.
3) Feature integration is dependent on the geometric and topologic constraints of layer approaches and attribute and relationship associations and integration into single objects.
4) Temporal integration can be accomplished through a feature approach with multiple instantiation of features from different time periods.

APPROACH:

Our approach will consist of two stages, one empirical and the other theoretical. The empirical work will provide a significant base from which we can draw more general theoretical concepts and develop an overall solution. We will use two specific test sites in St. Louis, Missouri, and Atlanta, Georgia, and develop an integrated set of the data layers for roads, streams, elevation, and land cover upon an image base. Throughout this integration process, which will be a combination of manual, semi-automated, and automated methods, we will be developing experience for designing the automated methods. In the second phase, we will use the developed knowledge to design a completely automated system for integrating these five data layers in a feature-based model. Since we will have the empirical base as an accurate gauge, we can change parameters for the automated process and determine effects.

Integrating Geographic Data using a Layer Approach – The Empirical Stage

To develop an appropriate base of empirical results to establish the theoretical limits needed for various combinations of resolution and accuracies of data, we will use existing datasets for test sites in St. Louis, Missouri, and Atlanta, Georgia. The data we will use for each site depends on availability, but, as a starting point, we will use all datasets in Table 1 in the testing process.

Table 1
Datasets for Empirical Integration

Data	Source	Type	Resolution	Accuracy	NM Layer
Elevation	NED	Raster	30 m	2-10 m	Elevation
Hydrography	NHD	Vector	--	13 m	Hydrography
Images	133 Urban Areas	Raster	1 ft	1 ft	Orthoimagery
Land Cover	NLCD	Raster	30 m	60 m	Land Cover
Transportation	133 Urban Areas	Vector	--	Various	Transportation

We anticipate that for the specific test areas, higher resolution datasets for elevation and land cover will be available from *The National Map* partnering program. Regardless, the objective will be to use available data to establish empirically the accuracies of data necessary to support integration for specific resolutions of data. The specific approach to integration is detailed below.

The possible combinations of data type, resolution, and accuracy yield 12 possible integrations shown in Table 2. The first step in integrating data from two different layers is matching specific features. This often is the most difficult part of the process. Since a layer-based dataset contains no specific geographic features with complete, self-contained attributes and relationships, as in a feature-based model, the location of common objects, *i.e.*, points, lines, areas, or pixels, in the layer becomes the basis of the feature-matching process. The actual matching process can be accomplished based on geometry, topology, and/or attribution.

Table 2
Integration Possibilities

	Data types	Resolution	Accuracy	Feature Matching	Tolerance Needed
1	Vector/vector	Same	Same	Point to point	No
2	Vector/vector	Same	Different	Point to points	Yes
3	Vector/vector	Different	Same	Point to points	Yes
4	Vector/vector	Different	Different	Points to points	Yes
5	Raster/raster	Same	Same	Pixel to pixel	No
6	Raster/raster	Same	Different	Pixel to pixel	Yes
7	Raster/raster	Different	Same	Pixels to pixel	Yes
8	Raster/raster	Different	Different	Pixels to Pixel	Yes
9	Raster/vector	Same	Same	Pixel to point	No
10	Raster/vector	Same	Different	Pixel to point(s)	Yes
11	Raster/vector	Different	Same	Pixel to point	Yes
12	Raster/vector	Different	Different	Pixels to points	Yes

Geometry Feature-Matching

The geometric approach to feature matching essentially requires matching the positions of features or geometric objects in the two layers. For each of the 12 possibilities for layer integration shown in Table 2, column 5 provides a possible feature-matching basis. For example, in row 10, a raster/vector layer integration is to be performed. Because the resolutions match, but the accuracies are different, the feature matching may require multiple points from the vector layer to find a single pixel match in the raster layer. To achieve a match where accuracies and resolutions are not the same between the two layers, tolerance values may be needed (Column 6 of Table 2). As an example, suppose the accuracy of the vector data in a layer to be integrated is ± 3.0 m in the horizontal. We could not then expect an exact point match to another vector layer (row 2 of Table 2) or another raster layer (row 10 of Table 2). Therefore, we would be forced to use a tolerance value of at least ± 3.0 m to insure the match. This is a simple case for matching.

If we examine raster data with different resolutions and different accuracies (row 8 in Table 2), the need for tolerance values to force matching becomes more obvious (Quattrochi and Lam, 1991).

While Table 2 provides a basis for feature matching using the positions of points and pixels, we can also examine feature matching in a layer-based integration approach using positions of lines and areas. These basic vector elements can be matched by lines of pixels and contiguous areas of pixels in a raster layer (Usery, 1994a). Again, we require tolerance values for geometric measurement to force the matches.

Topology Feature Matching

Similar to the geometric approach, the topologic approach forces a match between objects in the two layers, in this case topologic elements, *i.e*., nodes, lines, and areas, and the matching is established on topologic characteristics including adjacency and containment. For example, for two vector layers, for two specific lines to match, they must share the same bounding nodes and separate the same areas. This restriction is more stringent than the geometric matching and becomes difficult when integrating layers of different types. Often, the topologic match will force the generation of new geometry and topology as in the case of a typical overlay to integrate transportation and land cover layers.

Our objective in the empirical approach will be to determine appropriate resolution and accuracy requirements to insure exact feature matching of topological elements. The actual implementation will use a combination of automated procedures to position to the approximate location in the image and interactive procedures to finalize the feature matching. Our final goal of an automated procedure will use the empirical work to establish a framework and rules for the feature matching process. A variety of approaches has been used including plane-graph node matching (Lynch and Saalfeld, 1985), artificial intelligence methods (McKeown, 1987), and semantic rule bases (Cobb *et al*, 1998). Inclusion of the tolerance values suggested above is based on previous work involving uncertainty effects on feature matching (Foley *et al*., 1997).

Data Fusion -- Merger of Layers and Situating Features

Once the match is accomplished, the second step in the integration is merging and situating the features (objects). This can be done through adjustment of the individual layers to force the matching features to correspond or through actual merging or fusing the individual layers. In the first case, approaches include storing transformation parameters with the layer that are necessary to force the match to the second layer. The layers remain intact as separate layers but can be merged “on-the-fly” as needed by application of the transformation parameters. This approach is particularly useful when the data only need to be merged for display purposes. If the layers are to be integrated permanently and merged or fused, a single composite of the two inputs is the result.

The fused product is again dependent on the input layer types and characteristics. Perhaps the most common fused product is the merger of two different spatial and spectral resolution image products. An example of such a product is the Ikonos pan-sharpened image (Figure 1). In this particular image, the 0.8 m resolution panchromatic band of the Ikonos data is used as the intensity component of the image and the 4.0 m resolution red and infrared bands of the multispectral data are used as the hue and saturation components. The process to create such fusions is provided in conventional software packages and can use a variety of algorithms such as intensity-hue-saturation transformations, principal components analysis, and others.

Figure 1. Pan-sharpened Ikonos image created by fusing the 0.8 m panchromatic and 4 m multispectral bands of data. The feature shown is an airport over Camp Lejeune, North Carolina.

While image fusions create useful products and may form parts of the data integration requirements of *The National Map,* we see a greater need for fusions of vector datasets such as transportation and hydrography into raster datasets including digital orthographic images, DEM's and shaded relief representations of terrain. For these types of products, we will again explore the requirements for accuracy and resolution to create effective products that will be a part to *The National Map*. The results of the feature matching geometric work should apply largely to the fusion of these data to create composite products.

The exact methods for fusing vector data into raster images often include "burning" the lines into the images. Effectively, at a particular pixel location through which a vector passes, the pixel value in the raster image is replaced with a constant value representing the vector line. The actual raster value for that pixel is lost and the vector becomes a permanent part of the raster image. This approach works well for simple vector lines, but *The National Map* requires consideration of appropriate symbology for cartographic representation.

We will examine symbology issues in the integration process and the production of merged and fused products. Specifically, which symbols work in a fused product and which do not? Our base will be the standard USGS topographic symbol set applied to transportation and hydrography layers. We anticipate that symbol sizes and colors may need to be altered to produce an effective fused image product.

Integrating Geographic Data using a Feature Approach

The integration of data based on layers is necessary to handle our current generation products and the initial implementations of *The National Map*. Geographic data representation is rapidly moving to a feature-based approach. Examples include the U.S. Census TIGER files, the USGS National Hydrography Dataset, and current generation software systems such as the ArcGIS Geodatabase (ESRI, 2001). We propose to extend the work on layer integration and use it as a base to explore integration in a feature-based system. In a sense, the concept of integration is much simpler in a feature-based system since by definition, the concept of a feature is holistic and all attributes and relationships are integrated with the feature (Tang *et al*., 1996; Usery, 1996a; 1996b). In a feature model, the spatial data representation of a feature, in either raster or vector geometry, is simply one more attribute of the feature. The actual data integration process thus must occur during the construction of the spatial attributes associated with a feature. For example, a watershed feature will contain as attributes, the digital orthographic image(s) of the watershed, a DEM, a shaded relief image, and a line boundary. Subfeatures of the watershed "contained by" the watershed will include vector lines for the stream network, point locations for sampling stations and other point locations. This model has been well developed and a complete theory supporting it exists with implementations based on vector geometry (Tang *et al*., 1996), raster geometry (Usery, 1994a; 1994b), and a recent USGS implementation of a feature library incorporating multiple geometries and representations (Usery, 2000; Usery *et al*., 2002).

We propose to use this feature model (Figure 2) with the existing feature library (developed under GRA Task 740 in MCMC and Prospectus 2001 project "Feature Extraction from

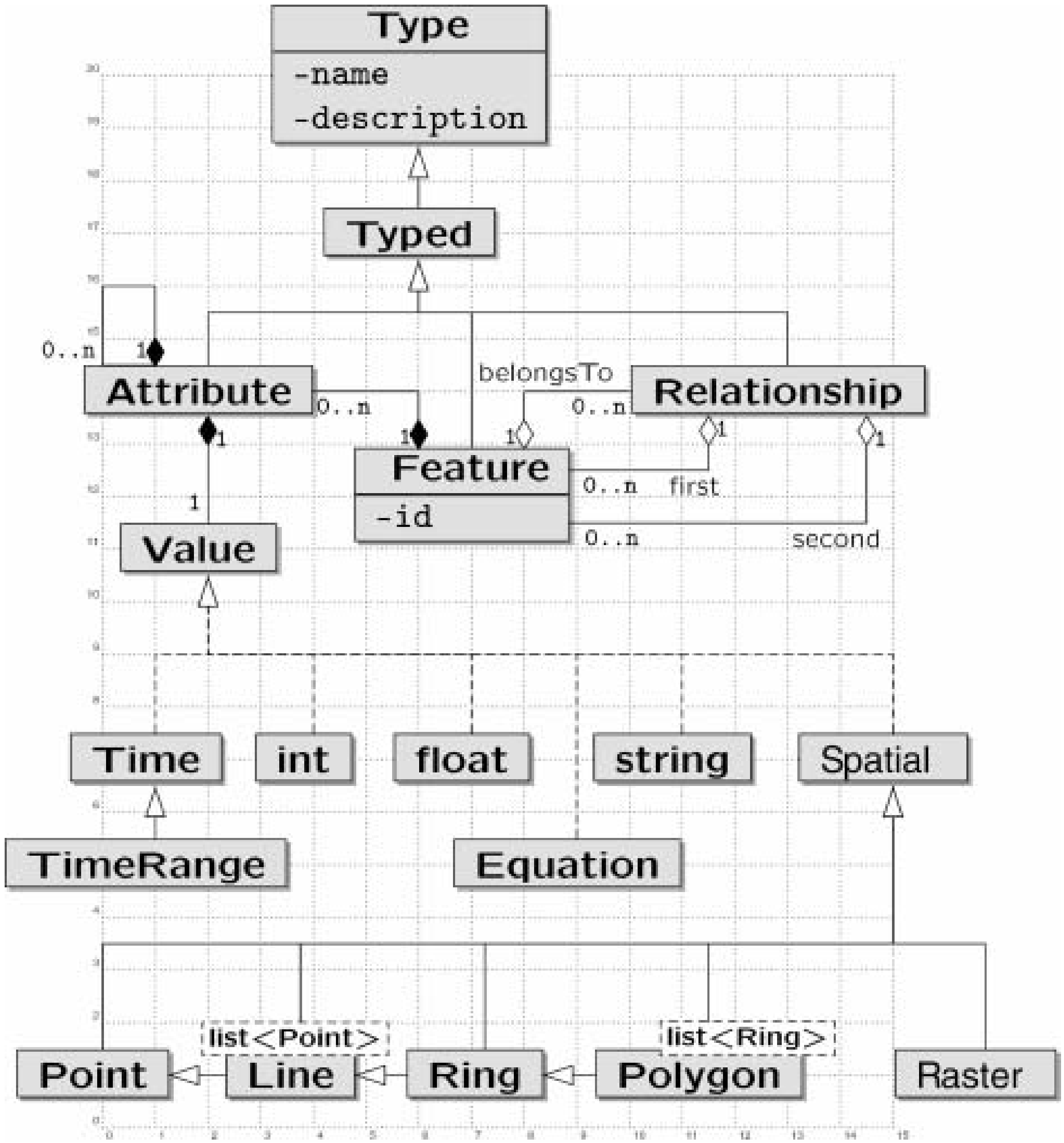

Figure 2. Feature library design to be used for data integration for *The National Map*.

Multimodal Sources to Support *The National Map*) to investigate data integration. The approach will be to use the results from the layer integration procedures above and the implemented databases for St. Louis and Atlanta to extract features for the library. These features will be structured with a fully implemented attribute structure from the available data and then used in a test process for feature display and analysis. The display test will require extracting the features from the library and displaying an integrated *National Map* dataset for St. Louis and Atlanta (Figure 3). For example, we would display selected hydrography features with shaded relief as one test. A second test would be to display selected features with associated transportation attributes including symbology with the digital orthographic image. These displays will be a text of the feature library, the integration of the attributes, and the results of the data integration and fusion operations developed in the layer approach.

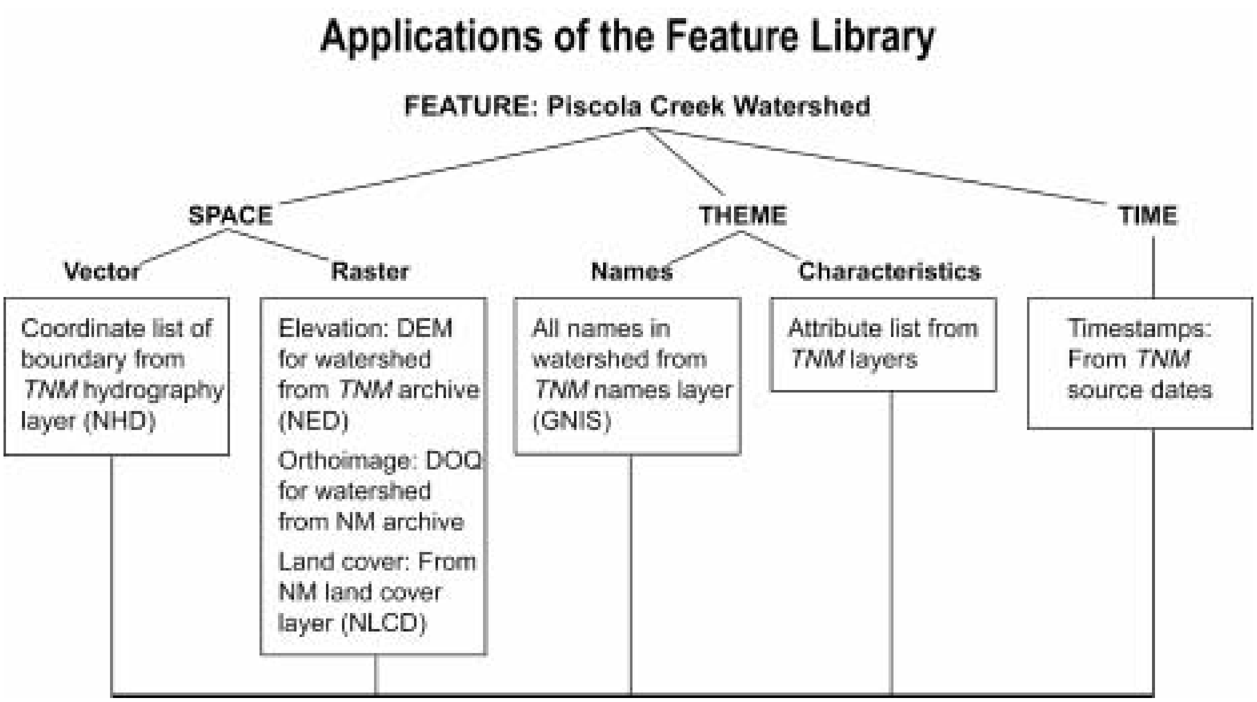

Figure 2. The feature library will access information from *The National Map* (*TNM*) to build the basic features and can be used for update. In initial form, *TNM* will be layer-based with nationwide mosaics of horizontally-integrated layers including the National Elevation Dataset (NED), the National Hydrography Dataset (NHD), the National Land Cover Dataset (NLCD), the Geographic Names Information System (GNIS) and a high resolution image base in the form of digital orthophotos (DOQ). Later versions of *TNM* will be feature-oriented similar to the feature library.

Metadata Use in Automated Data Integration

Both the layer and feature approaches to integration will build and use metadata to support the integration process. The standards developed for metadata can insure accuracy, resolution, and other information are available for any automated procedure (FGDC, 1997a; 1997b; 1997c). The value of metadata for data import, processing, and consistency and to store information on uncertainty has been established (Niemann, 1993; FGDC, 1998) and our experience to date indicates that the feature library can be populated with significant information from metadata stored with datasets such as NHD. The use of metadata in the integration process is essential to any automation operation. From our approach, we can store metadata indicating resolution and accuracy compatibilities and during the integration process, access the metadata to assure compliance and facilitate processing.

Study Sites St. Louis, Missouri, and Atlanta, Georgia, specifically the 133 urban area sites.

PRODUCTS:

1) A theoretical framework for integrating vector and raster data layers based on geometric and topological information.
2) Procedures for using the theoretical framework to establish an integrated *National Map* database from layer-based datasets.
3) Example integrated datasets for St. Louis and Atlanta using currently available data, *i.e.*, high-resolution images, land cover, transportation, etc., from the 133 urban area data.
4) A theoretical framework for a feature-based approach to data integration for *The National Map* extracting features from the St. Louis and Atlanta integrated databases.
5) Publications
    a. Data integration using layer based datasets.
    b. A data integration process with examples from St. Louis and Atlanta.
    c. Feature-based data integration for *The National Map*

PROJECT PERSONNEL QUALIFICATIONS:

Principal Investigator

Dr. Usery will serve as the principal investigator responsible for coordination of all project activities. He will manage the overall research activity and be responsible for the theoretical development including the layer integration framework and the feature model approach. He will collaborate with Mr. Finn on the software implementation of the two framework models and with Mr. Helterbrand on the theoretical framework approach using COTS and actual data.

Name E. Lynn Usery
Title & Affiliation Research Geographer, U.S. Geological Survey, MCMC
Associate Professor, University of Georgia

Education B.S. 1974 University of Alabama, Geography
M.A. 1977 University of Georgia, Geography.
Ph.D. 1985 University of Georgia, Geography.

Employment History

1977-1983	U.S. Geological Survey, Rolla, Missouri, Cartographer
1983-1988	U.S. Geological Survey, Rolla, Missouri, Geographer
1988-1993	University of Wisconsin - Madison, Assistant Professor
1994-1997	University of Georgia, Athens, Georgia, Assistant Professor
1997-Present	University of Georgia, Athens, Georgia, Associate Professor
1999-Present	Research Geographer, U.S. Geological Survey, Rolla, Missouri.

Research Interest And Relevant Experience

Dr. Usery conducts research in geographic information science, including geographic information systems, remote sensing and cartography with publications on theoretical aspects of geographic representation, human cognition of geographic phenomena, map and database projections, automatic feature extraction, and visualization, and applications of geographic information science to precision farming, watershed modeling, and water quality. Dr. Usery currently is a member of the Board of Directors of the University Consortium for Geographic Information Science (UCGIS), President of the Cartography and Geographic Information Society, and the Editor of the journal *Cartography and Geographic Information Science.*

Honors And Awards

2001, American Society for Photogrammetry and Remote Sensing, Leica Geosystems Best Scientific Paper in Remote Sensing.
1995, Connecting Teachers with Technology, Faculty Development Award, University of Georgia.
1986, Dissertation Award, Sigma XI.

Books

McMaster, R. and E.L. Usery, (eds), *The UCGIS Research Agenda*, Forthcoming, John Wiley and Sons, NY.

Usery, E.L., (ed), *1995. Proceedings, Workshop on Mapping and Environmental Applications of GIS Data, International Archives of Photogrammetry*, Volume XXX, Part 4W2, 166 p.

Relevant Publications

Usery, E.L., G. Timson, M. Coletti, 2002. A multidimensional geographic feature system, *GIScience 2002 Poster Summaries*, NOAA.
Seong, J.C., K.A. Mulcahy, E. L. Usery, 2002. "The Sinusoidal Projection: A new meaning for

global image data," *The Professional Geographer*, Vol. 54, No. 2, pp. 218-225.

Usery, E.L. and J.C. Seong, 2001. All Equal Area Map Projections are Created Equal, but Some are More Equal than Others, *Cartography and Geographic Information Science*, Vol. 28, No. 3, pp. 183-193.

Seong, J.C. and E.L. Usery, 2001. Modeling Raster Representation Accuracy Using a Scale Factor Model, *Photogrammetric Engineering and Remote Sensing*, Vol. 67, No. 10, pp. 1185-1191.

Usery, E.L., 2000. Multidimensional representation of geographic features. *International Archives of Photogrammetry and Remote Sensing, Volume XXXIII, Part B4/3, Commission 4*, pp. 240-247.

Seong, J.C. and E. L. Usery, 2001. Fuzzy image classification for continental-scale multitemporal NDVI series using invariant pixels and image stratification method. *Photogrammetric Engineering and Remote Sensing*, Vol. 67, No. 3, 287-294.

Kazmi, S.J.H. and E. L. Usery, 2000. Application of remote sensing and GIS for the monitoring of diseases: A unique research agenda for geographers. *Remote Sensing Reviews*, Vol. 20, No. 1, pp. 45-70.

Mozolin, M., J.C. Thill, and E.L. Usery, 1999. Trip distribution forecasting with multilayer perceptron neural networks: A critical evaluation. *Transportation Research Part B*, Vol 34, pp. 53-73

Usery, E.L., J.C. Seong, B.W. Jun, 1998. Implementing GIS software over the World-Wide-Web. *Proceedings, American Society for Photogrammetry and Remote Sensing Annual Convention*, Tampa, Florida, CD-ROM Publication.

Usery, E.L., 1996. A feature-based geographic information system model. *Photogrammetric Engineering & Remote Sensing*, Vol. 62, No. 7, pp. 833-838.

Usery, E.L., 1996. A conceptual framework and fuzzy set implementation for geographic features. *In* Burrough, P. A. and A.U. Frank, (eds.), *Geographic Objects with Indeterminate Boundaries*, Taylor and Francis, London, pp. 71-85.

Tang, A.Y.S., T. Adams, and E.L. Usery, 1996. A spatial data structure design for a feature-based GIS. *International Journal of Geographical Information Systems*, Vol. 10, No. 5, pp. 643-659.

Usery, E.L., 1996. Membership functions for fuzzy set representation of geographic features. *International Archives of Photogrammetry, Volume XXXI, Part B4*, p. 881-883.

Usery, E.L. and D. Pape, 1995. Extracting geographic features from raster data. *Proceedings, American Society for Photogrammetry and Remote Sensing Annual Convention*, Charlotte, North Carolina, pp. 733-740.

Usery, E.L., 1994. Implementation constructs for raster features. *Proceedings, American Society for Photogrammetry and Remote Sensing Annual Convention*, Reno, Nevada, pp. 661-670.

Usery, E.L.,1994. Display of geographic features from multiple image and map databases. *Proceedings, International Society for Photogrammetry and Remote Sensing, Commission IV Symposium on Mapping and Geographic Information Systems, International Archives of Photogrammetry, Volume XXX, Part B4*, Athens, Georgia, pp. 1-9.

Graff, L.H. and E.L. Usery, 1993. Automated classification of basic-level terrain features in digital elevation models. *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 9, pp. 1409-1417.

Michael P. Finn
Computer Programmer/Analyst
USGS, Mid-Continent Mapping Center, Rolla, MO

Mr. Finn will be responsible for software design and development including the integration framework and implementation of prototype algorithms for both layer-based and feature approaches to integration. He will work closely with Dr. Usery to implement the theoretical models in software and Mr. Helterbrand to couple the software developments with COTS and actual data for Atlanta and St. Louis.

<u>Education</u>

B. S. in Geography (Environmental and Natural Resources option)
Minor in Cartography and Map Technology
Southwest Missouri State University, Springfield, Missouri, 1982

M. S. in Civil Engineering (Specializing in Geodetic Engineering)
Virginia Polytechnic Institute & State University, Blacksburg, Virginia, 1991

<u>Work Experience</u>

3 years with DOI, USGS

- As a computer specialist, providing programming support for geographic and cartographic research projects, and developing unique solutions, one-of-a-kind systems, and proof-of-concept systems for research; investigator on two cartographic research projects; leading a team of software engineers designing and developing geospatial applications for research projects; and leading the Programming Support Unit, managing student computer programmers

17 years with US Department of Defense

- 10 years with US Air Force as a Program Analyst leading the Requirements Branch providing requirements management support to the Air Intelligence Agency; as a Computer Specialist leading the Small Computer Customer Support Branch providing maintenance, installation, troubleshooting, and training for systems to users at Patrick AFB and Cape Canaveral AS; and as a Physical Science Technician coordinating the testing of remote sensing systems being developed for the detection of electromagnetic pulses
- 7 years with Defense Mapping Agency as a Cartographer: as a systems analyst managing the Pooled Analytical Stereoplotter System; and as a scientific applications programmer, developing real-time photogrammetric data systems; as a photogrammetric cartographer producing digital terrain models

<u>Accomplishments</u>

Past President, Central Region, American Society for Photogrammetry and Remote Sensing
AF Space Command Civilian Command, Control, Communications, and Computers Systems Professionalism Award in 1993 and 1994
Four DOI Special Thanks for Achieving Results Awards

45th Communications Squadron's Mid-Level Civilian of the Year Award in 1994
Ten Performance Awards
Geography Award for Scholastic Achievement, SMSU, 1980

Publications

Finn, Michael P., E. Lynn Usery, Douglas J. Scheidt, Thomas Beard, Sheila Ruhl, and Morgan Bearden, 2002. "AGNPS Watershed Modeling with GIS Databases," Proceedings Second Federal Interagency Hydrologic Modeling Conference. Las Vegas, Nevada. Jul – Aug

Usery, E. Lynn, Michael P. Finn, and Douglas J. Scheidt, 2002. "Projecting Global Raster Databases," Proceedings Joint International Symposium on Geospatial Theory, Processing, and Applications. Ottawa, Canada. July

Usery, E. Lynn, Michael P. Finn, Daniel R. Steinwand, and Jeong Cheng Seong, 2002. "Projecting Global Raster Databases," Proceedings Geoinformatics for Global Change Studies and Sustainable Development. Nanjing, China. June

Usery, E.L., J.C. Seong, D. Steinwand, and M.P. Finn, 2001. "Methods to achieve accurate projection of regional and global raster databases," *U.S. Geological Survey Open-File Report 01-383*.

Co-Principal Investigator
Wm. Steve Helterbrand
Cartographer
USGS, Mid-Continent Mapping Center, Rolla MO

Mr. Helterbrand will be responsible for the design and development of empirical integration of the St. Louis and Atlanta datasets. This will include program development using COTS for the layer approach to integration and interfacing with Mr. Finn and Dr, Usery to achieve actual integration of the test data.

Education
A.S. in Design Drafting, Linn Technical College, Linn MO, 1982

Work Experience
21 Years with DOI, USGS

- 11 years as a technician specializing in data collection
- 10 years developing GIS tools to use digital products developed by the USGS.

Publications:

Starbuck, Michael, William Helterbrand, 1996. Merging Digital Raster Graphics and Digital Orthophoto Quadrangles for use in Digital Map Revision. Proceedings ASPRS / ACSM Annual Convention & Exhibition. Volume III Surveying and Cartography.

LITERATURE CITED:

Arnberg, W., 1981, Integration of map and remote sensing data, *Geografiska Annaler*, 63A3-4, 319-324.

Abel, D. J. and M. A. Wilson, 1990, A systems approach to integration of raster and vector data and operations, In Brassel, K. and Kishimoto, H., (eds), *Proceedings, 4th International. Symposium Spatial Data Handling*, Zurich: 2:559-566.

Chavez, P.S., Jr., 1986, Digital merging of Landsat TM and digitized NHAP data for 1:24,000-scale image mapping, *Photogrammetric Engineering & Remote Sensing*, 52/10, 1637-1646.

Chrisman, N.R., 1990, Deficiencies of sheets and tiles: Building sheetless databases, *International Journal of Geographical Information Systems*, 4/2, 157-168.

Cobb, M. A., Chung, M. J., Foley, H., Petry. F. E., Shaw, K. B. and Miller, H. V., 1998, A rule-based approach for the conflation of attributed vector data, *GeoInformatica*, 2/1, 7-36.

Congalton, R. G., 1991, A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37, 35-46.

Cowen, D., Jensen, J. R., Bresnahan, G., Ehler, D., Traves, D., Huang, X., Weisner, C., and Mackey, H.E., 1995, The design and implementation of an integrated GIS for environmental applications, *Photogrammetric Engineering & Remote Sensing*, 61, 1393-1404.

Ehlers, M., Edwards, G. and Bedard, Y., 1989 Integration of remote sensing with geographic information systems: A necessary evolution, *Photogrammetric Engineering and Remote Sensing*, 55/11,1619-1628.

Ehlers, M., Greenlee, D., Smith, T., and Star, J., 1991, Integration of remote sensing and GIS: Data and data access, *Photogrammetric Engineering and Remote Sensing*, 57/6, 669-676.

Egenhofer, M. and Herring, J., 1991. High level spatial data structures for GIS, In Maguire, D., Goodchild, M.F., and Rhind, D. (eds), 1991, *Geographical Information Systems: Principles and Application, Volume 1*, Longman, London227-237.

ESRI, 2001, *Geodatabase*, Environmental Systems Research Institute, Redlands, CA.

FGDC, 1997a, Framework: Introduction and Guide, Washington: Federal Geographic Data Committee, 105 p.

FGDC, 1997b, Geospatial Metadata, Washington: Federal Geographic Data Committee, 2 p.

FGDC, 1997c, Draft Geospatial Positioning Accuracy Standards, Washington: Federal Geographic Data Committee, 50 p.

FGDC, 1998, The Value of METADATA, Reston: FGDC Secretariat, 4 p.

Flowerdew, R., 1991, Spatial Data Integration, In Maguire, D., Goodchild, M.F., and Rhind, D. (eds), 1991, *Geographical Information Systems: Principles and Application, Volume 1*, Longman, London, 375-387.

Foley, H., Petry, F., Cobb, M. and K. Shaw, 1997, Using semantic constraints for improved conflation in spatial databases, *Proceedings, 7th International Fuzzy Systems Association World Congress*, Prague, 193-197.

Jensen, J.R., Saalfeld, A., Broome, F., Cowen, D., Price, K., Ramsey, D., and Lapine, L., 1998, Spatial data acquisition and integration, University Consortium for Geographic Science, http://www.ucgis.org/research_white/data.html. Also in McMaster, R. and Usery, E.L., (eds), In Press, *The UCGIS Research Agenda*: John Wiley and Sons, New York.

Lynch, M. and Saalfeld, A., 1985, Conflation: Automated map compilation -- A video game approach, *Proceedings, AutoCarto 7*, Washington, 343-352.

McKeown, Jr., D.R., 1987, The role of artificial intelligence in the integration of remotely sensed data with geographic information systems, *Geoscience And Remote Sensing*, Ge-25/3,330-348.

Niemann, B., 1993, Integration of metadata and space-time-depth variabilities and uncertainties, *Proceedings, Auto-Carto 11*, Minneapolis, 567-576.

Quattrochi, D., and N. S. N. Lam, 1991, Perspectives on integrating multiscale, multitemporal remote sensing data with GIS, *Proceedings, Special Session on Integration of Remote Sensing and GIS, ASPRS/ASCM*, Baltimore.

Saalfeld, A., 1988, Conflation: Automated map compilation, *International Journal of Geographical Information Systems*, 2/3, 217-228.

Star, Jeffery L., John E. Estes, and Kenneth C. McGwire (eds.) (1997) *Integration of Geographic Information Systems and Remote Sensing.* Cambridge University Press. Cambridge: New York.

Tang, A.Y.S., T. Adams, and E.L. Usery, 1996. A Spatial Data Structure Design for a Feature-Based GIS, *International Journal of Geographical Information Systems*, 10(5): 643-659.

Usery, E.L., 1994a. Implementation Constructs for Raster Features, *Proceedings, American Society for Photogrammetry and Remote Sensing Annual Convention*, Reno, Nevada, pp. 661-670.

Usery, E.L., 1994b. Display of geographic features from multiple image and map databases. *Proceedings, International Society for Photogrammetry and Remote Sensing, Commission IV Symposium on Mapping and Geographic Information Systems*, *International Archives of Photogrammetry, Volume XXX, Part B4*, Athens, Georgia, pp. 1-9.

Usery, E.L., 1996a. A Feature-Based Geographic Information System Model, *Photogrammetric Engineering and Remote Sensing*, 62(7): 833-838.

Usery, E.L., 1996b. A Conceptual Framework and Fuzzy Set Implementation for Geographic Features, *In* Burrough, P. A. and A.U. Frank, (Eds.), *Geographic Objects with Indeterminate Boundaries*, Taylor and Francis, London, pp. 71-85.

Usery, E.L., 2000. Multidimensional representation of geographic features. *International Archives of Photogrammetry and Remote Sensing, Volume XXXIII, Part B4/3, Commission 4*, pp. 240-247.

Usery, E.L., G. Timson, M. Coletti, 2002. A multidimensional geographic feature system, *GIScience 2002 Poster Summaries*, NOAA.

USGS, 2001, U.S. Geological Survey, *The National Map: Topographic Mapping for the 21st Century*. Reston, VA.

Welch, R., 1995, Emerging technologies for low cost, integrated GPS, remote sensing and GIS applications, *Proceedings, Cambridge Conference for National Mapping Organizations*, Cambridge, England, 6 p.

Welch, R. and Ehlers, M., 1987, Merging multiresolution SPOT HRV and Landsat TM data. *Photogrammetric Engineering & Remote Sensing*, 53/3, 301-303.

Welch, R. and Homsey, A., 1997, Datum shifts for UTM coordinates, *Photogrammetric Engineering and Remote Sensing*, 63/4, 371-375.

MISCELLANEOUS:

The proposed work will leverage results from a previously funded Prospectus project, "Feature Extraction from Multimodal Sources to Support *The National Map,*" and a project internally funded by MCMC, "Multi-Dimensional Theory and Multiple Representation of Geographic Features." The proposed work will also take advantage of ongoing database developments for *The National Map* and attempt to improve those databases through data integration.

Budget:

FY 2003 PROJECT BUDGET

Project Title: **Automated Data Integration in Support of *The National Map***

Principal Investigator(s): **E. Lynn Usery, Michael P. Finn, and Steve Helterbrand**

PROJECT DIRECT COSTS:

	Region Performing Work	
LABOR EXPENSES	Central Region	TOTAL
Govt. Salary & Benefits	215,613.27	215,613.27
Subtotal Labor Costs:	215,613.27	215,613.27
OTHER EXPENSES		
International Travel	4,000.00	4,000.00
Domestic Travel	4,000.00	4,000.00
Supplies	2,000.00	2,000.00
Training	2,000.00	2,000.00
Equipment	25,000.00	25,000.00
Indirect Services	40,000.00	40,000.00
Subtotal Other Expenses:	77,000.00	77,000.00
Subtotal Direct Costs:	292,613.27	292,613.27

PROJECT INDIRECT COSTS:

Assessment Rate Applied	11%	
Indirect Costs	14,907.27	14,907.27
Subtotal Indirect Costs:	14,907.27	14,907.27
TOTAL PROJECT COST:	307,520.53	307,520.53

FY 2004 PROJECT BUDGET

Project Title: **Automated Data Integration in Support of *The National Map***

Principal Investigator(s): **E. Lynn Usery, Michael P. Finn, and Steve Helterbrand**

PROJECT DIRECT COSTS:

	Region Performing Work	
LABOR EXPENSES	Central Region	TOTAL
Govt. Salary & Benefits	226,393.93	226,393.93
Subtotal Labor Costs:	226,393.93	226,393.93
OTHER EXPENSES		
International Travel	4,200.00	4,200.00
Domestic Travel	4,200.00	4,200.00
Supplies	2,100.00	2,100.00
Training	2,100.00	2,100.00
Indirect Services	42,000.00	42,000.00
Subtotal Other Expenses:	54,600.00	54,600.00
Subtotal Direct Costs:	280,993.93	280,993.93

PROJECT INDIRECT COSTS:

Assessment Rate Applied	11%	
Indirect Costs	15,652.63	15,652.63
Subtotal Indirect Costs:	15,652.63	15,652.63
TOTAL PROJECT COST:	296,646.56	296,646.56

FY 2005 PROJECT BUDGET

Project Title: **Automated Data Integration in Support of *The National Map***

Principal Investigator(s): **E. Lynn Usery, Michael P. Finn, and Steve Helterbrand**

PROJECT DIRECT COSTS:

	Region Performing Work	
LABOR EXPENSES	Central Region	TOTAL
Govt. Salary & Benefits	237,713.63	237,713.63
Subtotal Labor Costs:	237,713.63	237,713.63
OTHER EXPENSES		
International Travel	4,410.00	4,410.00
Domestic Travel	4,410.00	4,410.00
Supplies	2,205.00	2,205.00
Training	2,205.00	2,205.00
Subtotal Other Expenses:	13,230.00	13,230.00
Subtotal Direct Costs:	250,943.63	250,943.63

PROJECT INDIRECT COSTS:

Assessment Rate Applied	11%	
Indirect Costs	16,435.26	16,435.26
Subtotal Indirect Costs:	16,435.26	16,435.26
TOTAL PROJECT COST:	267,378.89	267,378.89

PROJECT BUDGET JUSTIFICATION

The travel funds will be used t support of the principal investigators to attend symposiums on data integration and to report results of the research at appropriate conferences. The supplies funds will support acquisition of books, journals, and other publications as well as computer supplies. The supplies funds will also be used to support publication costs, such as page charges and reprints. Training funds will be used to support project staff training with specific software to be used on the project including development toolkits and application packages.

The equipment budgeted is needed to support all project developments. Each of the PI's will require equipment for this project and the programming staff, including a computer specialist and two fulltime students require equipment. High-end equipment will also be needed to handle the empirical datasets for St. Louis and Atlanta.

Signed:

/s/ E. Lynn Usery, Michael P. Finn, Wm. Steve Helterbrand

Principal Investigator(s)

/s/ Kari J. Craun

Cost Center Chief

This form should indicate signatures and be forwarded electronically to bmiller@usgs.gov by Cost Center Chief.