Integration of *The National Map*

Michael P. Finn [a*], E. Lynn Usery [a], Michael Starbuck [a], Bryan Weaver [b], and Gregory M. Jaromack [a **]

[a] U.S. Geological Survey, Mid-Continent Mapping Center, 1400 Independence Road, Rolla, Missouri 65401, (mfinn, usery, mstarbuck)@usgs.gov
[b] University of Georgia, Department of Geography. GG Building, Athens, Georgia 30603-2502, bweaver@uga.edu

COMMISION IV, WG IV/7

ABSTRACT:

The United States Geological Survey (USGS) is seeking partnerships to ensure that current, complete, consistent, and accurate geographic base information is readily available and useful to the U. S. (Groat, 2003). The partnerships will be used to help create *The National Map*, a seamless, continuously maintained set of base information that will serve as a foundation for integrating, sharing, and using other data easily and consistently (USGS, 2002). Data integration, a significant problem for *The National Map*, is being examined to develop a conceptual framework based on resolution, geometric accuracy, and topological consistency of data layers. We will apply this conceptual framework to five of *The National Map* layers: digital orthoimages, elevations, land cover, hydrography, and transportation. The framework developed from this approach will allow an examination of a feature approach to integration based on a model previously developed and implemented as a feature library (Usery et al, 2004). We anticipate results leading to an approach based on the conceptual framework, the empirical results, and the use of metadata to drive an automated process.

This study is motivated by the fact that *The National Map* will consist ultimately of integrated datasets, a digital product similar to the comprehensive, integrated lithographic printed map, and that current USGS digital products are single layer and not vertically-integrated. The goal is to develop procedures for automated integration based on metadata. Objectives of our research are to develop a framework for layer integration based on metadata, a framework for feature integration, and demonstrate example results of integration processes for Atlanta and St. Louis (Usery *et al*, 2003). Integrating data themes for *The National Map* is a significant problem consisting of several components including differences in datums, projections, coordinate systems, data models, spatial and temporal resolution, precision, and accuracy.

Datasets include orthoimages from Nunn-Lugar-Domenici 133 priority cities of the Homeland Security Infrastructure Program (Vernon, Jr., 2004); National Hydrography Data (USGS, 1999); National Elevation Data (Gesch *et al*, 2002); transportation (from Digital Line Graphs (USGS, 1996), 133 priority cities, or state departments of transportation); and the National Land Cover Dataset (Vogelmann *et al*, 2001). Of the list of spatial database issues presented by Laurini (1998), these are the most pertinent to our dataset integration problem:

- Diversity in spatial representations
- Diversity in global projections
- Diversity in values for the same items located at different sites
- Diversity of spatio-temporal sampling
- Variability of definitions over time and space
- Discrepancies in coordinate values
- Discrepancies in boundary alignment
- Variability in content quality
- Variability in data maintenance procedures
- Discrepancies in spatial metadata

Methods of layer integration being studied include determining compatible resolutions and accuracies and use of metadata to automatically combine appropriate datasets; determining possible transformations that integrate datasets of incompatible resolutions and accuracies; and determining limits of integration based on resolution and accuracy. Keates (1982) pointed out some basic cartographic transformations that relate to our integration problems, from sphere to plane (projection), three-dimensional to two-dimensional, and generalization. A general theoretical and conceptual framework is needed to accommodate at least five distinct forms of data integration (Jensen *et al.*, 1998):

- In situ measurement-to-in situ
- In situ measurement-to-foundation
- Vector-to-foundation map

---

[*] Corresponding author.
[**] Now with RCM Technologies, Omaha, NE

- Image-to-foundation map
- Image-to-foundation image

The National Map requirements for data integration include all of the types listed above. We are studying several generalization issues: selection (Topfer and Pillewizer, 1966), simplification, symbolization, and induction. This approach consists of two stages: one empirical and the other theoretical. There are two distinct types of integration: horizontal integration and vertical integration. A fundamental step in the integration is merging and situating features.

Beyond the development of automated integration approaches, *The National Map* success depends upon thoughtful consideration of the complexities inherent to spatial partnership programs. Masser and Campbell (1995) highlight some key obstacles to GIS sharing: variation in participant priorities; variation in GIS experience and technical ability; differences in spatial data handling skills; and disagreement among participants regarding data openness, leadership, data standards, equipment and training.

Organizational impediments present some fundamental project realizations: the cost of spatial data sharing is significant; successful development of *The National Map* is dependent upon an up-front understanding of *The National Map* stakeholder community needs and capabilities; and long-term implementation must find a balance between mandated compliance with standards for participants and local control of data management. The integration of data based on layers is necessary to handle our current generation products and the initial implementations of *The National Map*. Once clear project authority is assigned and leadership established, making the transition from the grand vision statement to a feasible implementation plan will require a thorough evaluation of the geospatial community needs, at all levels of government.

Examining the transportation theme as an example of the complexity of integration issues, it is difficult to envision how the USGS can employ the current strategy over the entire nation and meet transportation requirements of *The National Map* and National Spatial Data Infrastructure stakeholders. Weaver (2004) recommends a phased approach to project implementation: determine transportation requirements and national capabilities; develop data integration alternatives; cost-benefit analysis of alternatives; and physical implementation. The long-term solution to integrating national street centerline data is to develop a distributed database of national extent that meets the scale requirements of local governments. Weaver (2004) suggests a long-term conceptual data flow for *The National Map* Transportation layer:

- Stage 1: Data Providers.
- Stage 2: Provider Schema Federation.
- Stage 3: Central Transportation Database.
- Stage 4: Cartographic Rendering.
- Stage 5: *The National Map* Product.
- Stage 6: Other Data Theme Relations
- Stage 7: Geospatial One Stop Relation (an intergovernmental project managed by the U. S. Department of the Interior in support of the President's Initiative for E-government, Geospatial One Stop builds upon its partnership with the Federal Geographic Data Committee to improve the ability of the public and government to use geospatial information to support the business of government and facilitate decision-making).

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). Preliminary results indicate that data integration of layers for *The National Map* can only be accomplished with datasets that are compatible in resolution and accuracy. Further, it appears that mathematical transformation can automate integration with limited ranges of scales, but cannot correct generalization differences between datasets.

REFERENCES

Gesch D, Oimoen M, Greenlee S, Nelson C, Steuck M, Tyler D, 2002. The National Elevation Dataset. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 1. American Society for Photogrammetry and Remote Sensing. Bethesda, Maryland.

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.

Groat, C. G., 2003, *The National Map* – A Continuing, Critical Need for the Nation. *Photogrammetric Engineering and Remote Sensing* **69** (10) 1087-1090. American Society for Photogrammetry and Remote Sensing, Bethesda, MD.

Jensen, J.R., Saalfeld, A., Broome, F., Cowen, D., Price, K., Ramsey, D., and Lapine, L., 1998, Spatial data acquisition and integration, In McMaster, R. and Usery, E.L., (eds), In Press, *The UCGIS Research Agenda*: John Wiley and Sons, New York.

Keates, J. S., 1982, *Understanding Maps*. Longman: London

Laurini, R. 1998, Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability. *International Journal of Geographic Information Science* **12,** 373-402.

Masser, I., and Campbell, J., 1995, Information sharing: the effects of GIS on British local government. In Onsrud H. J. and Rushton, G., (eds), *Sharing Geographic Information:* Center for Urban Policy and Research, New Brunswick, NJ. 230-249.

Topfer, F. and Pillewizer, W., 1966, The Principles of Selection. *The Cartographic Journal*, **3**, 10 – 16.

USGS, 1996, *US GeoData Digital Line Graphs*. U.S. Geological Survey Fact Sheet 078-96, Reston, VA.

USGS, 1999, *The National Hydrography Dataset*. U.S. Geological Survey Fact Sheet 106-99, Reston, VA.

USGS, 2002, *The National Map: Topographic Mapping for the 21st Century*. U.S. Geological Survey Fact Sheet 018-02, Reston, VA.

Usery, E. Lynn, Michael P. Finn, and Michael Starbuck, 2003. Data Integration of Layers and Features for *The National Map*. *American Congress on Surveying and Mapping Annual Conference,* July. Phoenix, AZ. ACSM.

Usery, E.L., G. Timson, M. Coletti, 2004. "Multidimensional Representation of Geographic Features," *International Journal of Geographic Information Science*, in review.

Vernon, Jr., Daniel E., 2004. Geospatial Technologies in Homeland Security. *EOM, Earth Observation Magazine*, Vol 13, No. 1 (Jan). GTIC America, Inc. Frederick, MD.

Vogelmann JE, Howard SM, Yang L, Larson CR, Wylie BK, and Van Driel N, 2001. Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper Data and Ancillary Data Sources. *Photogrammetric Engineering and Remote Sensing*, Vol. 67, No. 6. American Society for Photogrammetry and Remote Sensing. Bethesda, Maryland.

Weaver, B., 2004. Implementation of *The National Map* Road Database. *American Congress on Surveying and Mapping Annual Conference,* April. Nashville, TN. ACSM.