Building an Enterprise Geodatabase and Data Model

Michael Bieberitz, GISP
Michael Benedict, GISP
Sara Hackbarth, GISP
Sarah Seifert

Abstract

In an effort to improve business and mapping efficiency, the Milwaukee Metropolitan Sewerage District (MMSD), a state-chartered, government agency providing wastewater services for 28 municipalities, developed an Enterprise Geodatabase as part of a requirements gathering and planning phase for the MMSD's Enterprise GIS solution. As an essential component of the planning phase, MMSD identified a need to assemble all of its geographic data assets into a single Enterprise Geodatabase Model.

This presentation will describe and illustrate the development process used to assemble a diverse collection of datasets for the Enterprise Geodatabase Model. The process included defining user needs, designing the conceptual, logical, and physical database structure, and the development of a final model using Unified Modeling language (UML) design.

Introduction

In an effort to improve organizational efficiency and bring added value to the way the Milwaukee Metropolitan Sewerage District (MMSD) conducts business, the MMSD had developed a Geographical Information Technology Vision and Implementation Plan that provides a roadmap towards an enterprise implementation of GIS technology. Included in this Vision plan was acknowledgment of the need for development of an Enterprise Geodatabase.

Background

MMSD provides wastewater services for 29 satellite local government entities in a 420-square-mile service area covering 5 southeastern Wisconsin counties. Serving these municipalities requires MMSD to continue developing spatial inventories and applications that meet internal and external needs for planning and design.

MMSD embraced Computer Aided Design (CAD) technology in the early 1990's, when several large scale conversion projects were migrated from hard copy to digital format. In 2003, MMSD began implementing department specific desktop GIS solutions using ESRI technology. Two years later, in 2005, it became apparent that the organization's approach to compartmentalized GIS implementation was shortsighted and lacked a grand plan.

MMSD enlisted the assistance of a consultant to support the creation of a Geographical Information Technology Vision and Implementation Plan. This plan deliverable inventoried GIS assets and analyzed the organizational needs, presenting a high level approach for enterprise GIS implementation. This gave way to a yearlong Phase I planning effort that prepared 11 workgroup specific Software Requirement Specification (SRS) documents and an Enterprise GIS Business Data Model.

MMSD selected HNTB to develop SRS documents for each of the applications and to design the enterprise geodatabase.

Phase I data model design was built off of the efforts of two separate initiatives that were part of this phase of the project: Existing Data Inventory and Application User Needs.

Existing Data Inventory

A full inventory of existing MMSD GIS data was conducted as part of the Phase I planning effort. Inventoried data included raw geographic information utilized by MMSD staff to perform work functions, communicate or provide context. Over 600 GIS datasets were inventoried in a number of different formats including Shapefiles, geodatabases, CAD, MicroStation, Oracle tables, and others. Many of these datasets were found to be duplicates or very similar to other datasets. An effort was made to best identify the most current dataset. This exercise resulted in the identification of 340 unique datasets.

Application User Needs

Data requirements for each of the functional requirements of the applications were collected during the SRS development process. This allowed for data gaps and data needs to be identified and incorporated into the overall geodatabase design. Each application had a “Champion” designated by MMSD. This person helped HNTB both define requirements and user needs. In total, 58 interviews and follow-up interviews were conducted with over 80 users in 20 departments.

Data Model Design

Development of the MMSD Enterprise GIS Business Data Model followed a three-step process. The first stage, Conceptual, was intended to assemble a high-level abstract representation of the GIS layers and to identify any basic relationships between data entities. After completion of that exercise, the Logical model was developed. The Logical model is more detailed than the Conceptual and includes the addition of some attributes. The final stage, development of a Physical model, provides for the greatest detail and specifically defines attributes and their characteristics.

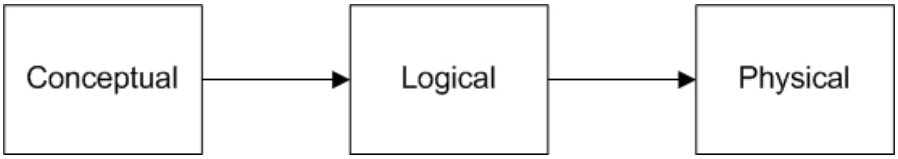

Figure 1 – Data Model Design Process

Conceptual Design

The conceptual stage of the data model design started with the identification and grouping of similar features (GIS layers) into categories, or thematic groups. MMSD staff identified 10 major thematic groups. Grouping individual features allows for a manageable data model and a higher level of preferred standards on which MMSD staff can rely upon. Within each thematic group, tabular and geospatial information is stored as a collection of related entities.

The table below lists the thematic groups that were identified, along with their descriptions.

Thematic Group	Thematic Group Description
Aerial	Aerial thematic group includes aerial images of the MMSD area throughout the years. Examples include aerials from 1936,1995,2000,etc.
Base Data	Base Data thematic group includes layers frequently used for GIS purposes. Base layers provide additional context for GIS analysis and cartographic products.
Boundaries	Boundaries thematic group includes boundaries maintained by MMSD or obtained by MMSD to suit a specific business purpose.
Conveyance	Conveyance thematic group includes assets that define a distribution system used to carry wastewater/stormwater from one location to another. Although Conveyance generally refers only to the conduits and channels involved in the distribution system; for the purposes of this model, this group will include other infrastructure associated with the distribution system.
Geotechnical	Geotechnical thematic group includes layers related to earth material and below-ground investigation.
MMSD Base	MMSD Base thematic group refers to layers maintained by MMSD and are not part of MMSD core business but are commonly used for GIS purposes.

Thematic Group	Thematic Group Description
Plants	Plants thematic group includes layers related to MMSD structures built to treat wastewater before discharging it into the environment.
Real Estate	Real Estate thematic group includes layers related to the acquisition and maintenance of property interests.
Water Quality	Water Quality group includes layers involved in the measurement of characteristics about water, usually in respect to its suitability for a particular purpose.
Watercourse	Watercourse thematic group includes defined channels within which water flows, either continuously or in season. A watercourse is continuous in the direction of flow and may extend laterally beyond the definite banks to include overflow channels contiguous to the ordinary channel. Layers in this group are owned or maintained by MMSD.

Table 1 – Thematic Groups and Descriptions

The Conveyance thematic group was further defined by four thematic sub-groups. Sub-groups more specifically group similar spatial or attribute features within a broader thematic group.

Thematic Group	Sub-group	Sub-group Description
Conveyance	Equipment	Equipment sub-group contains the features in the wastewater system that do not have an associated geometry or position. Instead, their locations are determined by their relationships with other explicitly positioned features such as network structures. Equipment items do not participate in the geometric network.
	Facility	Facilities sub-group contains features that are used in the transmission and distribution of wastewater and are commonly used to join various water lines together. Facilities have geometric positions and participate in the geometric network.
	Feature	Feature sub-group contains spatial entities in water, sewer, and stormwater systems that do not participate in the geometric network
	Line	Line sub-group includes the assorted types of pipes used in the transmission and distribution of water. Line features do participate in the geometric network.

Table 2 – Thematic Sub-groups and Sub-group Descriptions

Logical Design

The logical design level provided MMSD with an important stage for helping to visualize how the final geodatabase design would look. Important data relationships, shapes, and business attributes were identified.

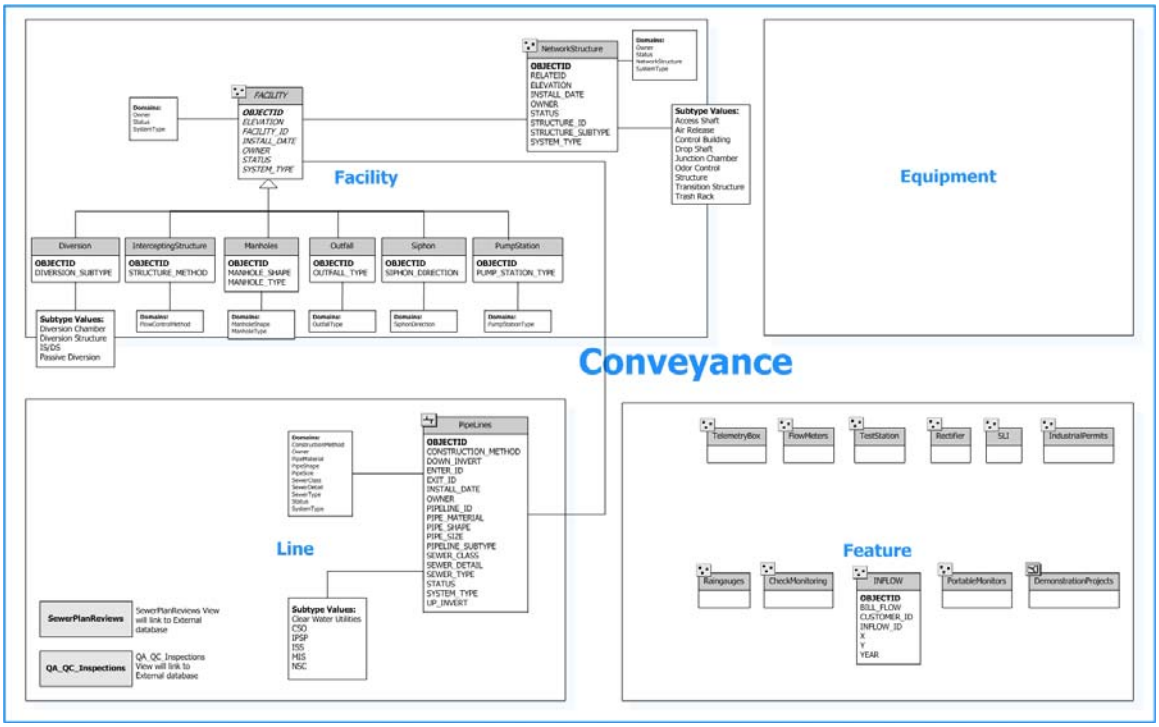

Figure 2 – Logical Design for Conveyance Thematic Group

Physical UML Design

A decision was made to build and maintain the physical design of the enterprise geodatabase with Unified Modeling Language (UML) in Microsoft Visio. UML provided the highly rigid structure that was appealing for modeling MMSD GIS data.

The following geodatabase schema elements were modeled in UML:

- Feature Datasets
- Feature Classes
- Stand-alone Tables
- Fields
- Domains
- Subtypes
- Geometric Networks
- Relationship Classes

The following geodatabase elements were not modeled in UML and are maintained separately:

- Raster Catalogs
- Raster datasets
- Topology
- Annotation
- Spatial Reference
- Metadata

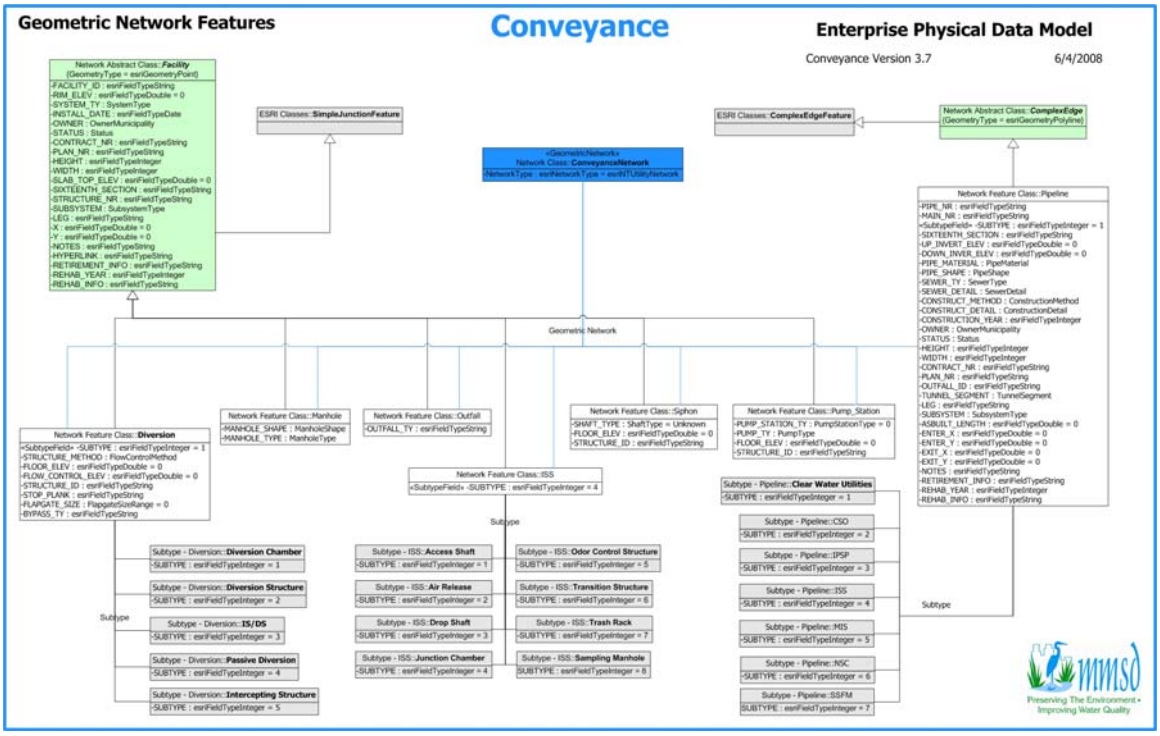

Figure 3 – Physical Design for Conveyance Geometric Network

Geodatabase System Design

The MMSD Enterprise geodatabase system consists of three geodatabases maintained in an Oracle environment.

Geodatabase	Description	Thematic Groups
Raster	Aerial and other imagery.	Aerial
Versioned	GIS layers maintained by MMSD.	Boundaries
		Conveyance
		Geotechnical
		MMSD Base
		Plants
		Real Estate
		Water Quality
		Watercourse
Non-Versioned	Base data and static GIS layers.	Base Data

Table 3 – Geodatabase Environment

Each of the 10 thematic groups was consolidated into one of three ArcSDE geodatabases. Only data targeted for the versioned geodatabase is maintained by MMSD using the UML model.

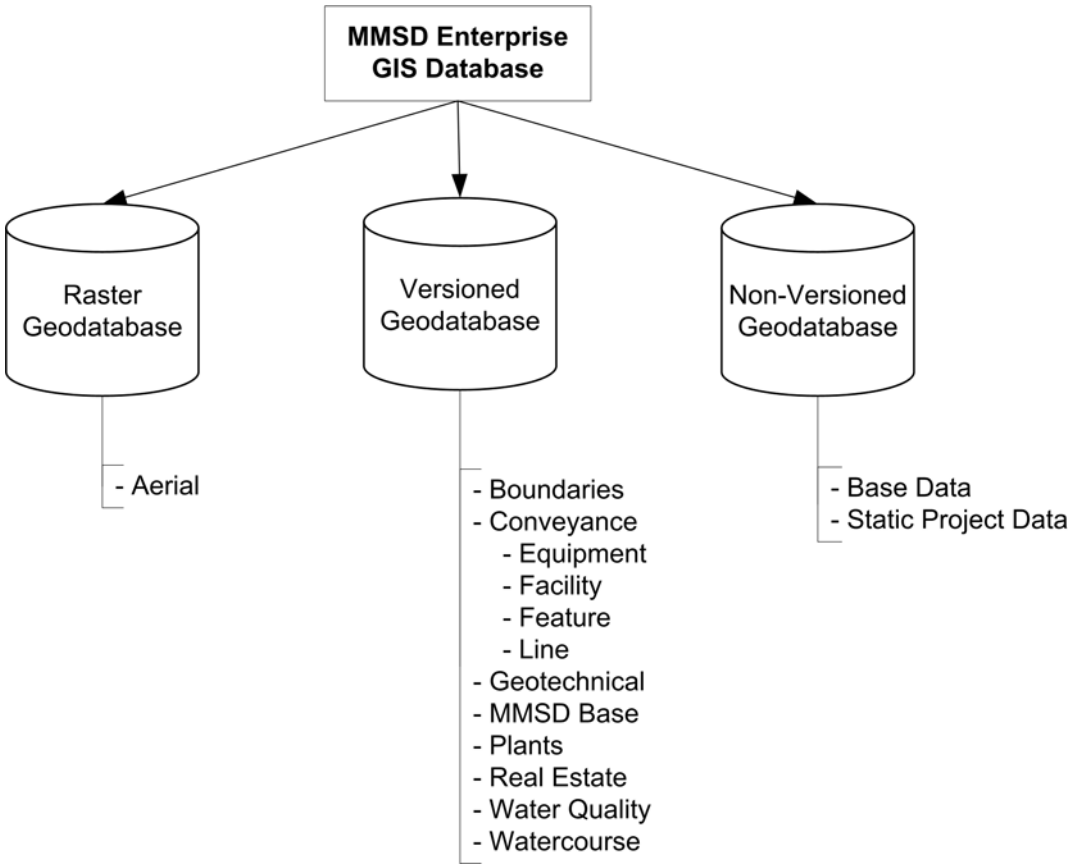

Figure 4 –Geodatabase System Design

Conclusion

Phase I, which included enterprise geodatabase design, was completed in October of 2007. MMSD began Phase II, GIS Vision Implementation in April of 2008. Phase II executes a GIS Implementation based upon the detailed guidelines resulting from the completed Phase I project. MMSD selected HNTB for Phase II Implementation.

As a component of Phase II, a data conversion effort will be undertaken to populate the enterprise geodatabase. A change management procedure has been implemented for requesting and managing changes to the MMSD Enterprise GIS Data Model. User data requirements will continually change over time. Because of this, an enterprise GIS data model needs to be flexible enough to evolve as business needs change. At the same time, it is vital that this process be managed in an effective way to ensure database integrity.

The enterprise geodatabase will be fully populated as a result of a related data conversion effort. This task will be completed by the end of 2008. The remainder of the Phase II project is expected to be completed by April 2010.

References

The ESRI Support Center maintains documents that show how to design a geodatabase in UML with Visio. The documents below were used to help assemble the enterprise geodatabase:

CASE Tools Tutorial
http://support.esri.com/index.cfm?fa=knowledgebase.documentation.viewDoc&PID=43&MetaID=658

Designing Geodatabase with Visio
http://support.esri.com/index.cfm?fa=knowledgebase.documentation.viewDoc&PID=43&MetaID=658

Introduction to CASE Tools
http://support.esri.com/index.cfm?fa=knowledgebase.documentation.viewDoc&PID=43&MetaID=658

Repository or XMI to Geodatabase
http://support.esri.com/index.cfm?fa=knowledgebase.documentation.viewDoc&PID=43&MetaID=571

Portions of the MMSD Enterprise GIS Data Model build off of existing industry-specific data models. ESRI maintains several data model templates that are maintained by academic and industry leaders. A complete list of currently maintained data models for select industries is available on the ESRI website.

ESRI Data Models
http://support.esri.com/index.cfm?fa=downloads.dataModels.matrix

The most commonly referenced data model that was referenced for developing the MMSD Enterprise GIS Data Model was the Water Utilities Data Model. Specifically, the Stormwater Data Model and the Water Utilities Data Model Reference Book.

Water Utilities Data Model
http://support.esri.com/index.cfm?fa=downloads.dataModels.filteredGateway&dmid=16

Author Information

Michael Bieberitz, GISP
Senior GIS Developer
HNTB Corporation
111 North Canal St, Ste. 1250
Chicago, IL 60606-7252

Michael Benedict, GISP
Senior Systems Analyst
Milwaukee Metropolitan Sewerage District
260 W. Seeboth Street
Milwaukee, Wisconsin 53204

Sara Hackbarth, GISP
Facilities Information Supervisor
Milwaukee Metropolitan Sewerage District
260 W. Seeboth Street
Milwaukee, Wisconsin 53204

Sarah Seifert
Data Coordinator
Milwaukee Metropolitan Sewerage District
260 W. Seeboth Street
Milwaukee, Wisconsin 53204