KANSAS WATER UTILITY DATA STANDARD

September 2008

Kansas GIS Technical Advisory Committee

Water Utility Data Sub-Committee

Kansas Water Utility Data Standard

Table of Contents

Kansas Water Utility Data Standard

1. Introduction

1.1 Mission and Goals of the Kansas Water Utility Data Standard

This document provides guidelines by which public water suppliers may acquire water utility information that is uniformly developed, maintained and cartographically represented and thereby integrated with other geospatial data. Digital representations of water utility features are the focus of this standard, but similar concepts of uniformity, currency, and completeness also apply to manual mapping efforts. This standard is intended to ensure that water utility data are accurate, reliable and accessible to Kansas public water suppliers, as well as the Kansas Division of Emergency Management and the Kansas One Call system.

The Water Utility Data Standard adopts the Kansas GIS Vision Statement as follows:

To shape the growth of GIS through open communication, education, and cooperation in order to:

- *Optimize data accuracy, reliability, and accessibility*
- *Meet the needs of the technical and non-technical user community*
- *Support the decision-making process*

In coordination with other Kansas Geospatial Data Standards this document echoes the following objectives:

- *Create an attitude of cooperation*
- *Generate support*
- *Identify common interests*
- *Identify established guidelines for developing and maintaining standards*
- *Identify areas of need for standardization*
- *Identify obstacles and barriers to data sharing*
- *Avoid duplication in creating data*
- *Establish standardized metadata*
- *Ensure data security*
- *Create flexible standards*
- *Catalog existing data*
- *Build a larger community of technical and non-technical users*
- *Develop a geographic data framework for Kansas that is compatible with the framework of the National Spatial Data Infrastructure.*

1.2 Need for Standard

GIS data development, conversion, and maintenance represent significant investments. To ensure the best possible return on these investments, standards are needed to develop and disseminate quality GIS data, to promote data formats that will support quality control and assurance over the long term, and to seek data distribution options that offer a range of products to best meet the variable needs and capabilities of the potential user base.

It is also important to understand the uses that can be made of the data collected. At its most basic level of attribution, the data can be used to produce maps of a public water supply system to show the system infrastructure in relation to physical features on the ground; it can also be used to locate underground infrastructure such as valves and pipelines. All public water supply systems in

Kansas would benefit from this level of information. Refer to Appendix A, "Minimum Requirements for Water Utility Features" for a listing of infrastructure features that meet the most basic needs for a small public water supplier.

However, more complex datasets can be developed that provide both greater attribution and a distribution network. This level of detail is useful in planning and maintenance and is suitable for many of the larger public water suppliers in Kansas. A comprehensive data model is under development. When it is finalized it will be found in this standard under Appendix B.

This water utility data standard includes both the minimum requirements for a simple dataset and the more complex data model. While the minimum requirements do not provide a robust framework for fully integrating GIS technology into the operations of a public water supply system, it will serve the needs of many entities. If there is a desire to improve the functionality of the dataset in the future, it can be transformed into the larger data model. The standard will assist public water suppliers in their efforts to develop a dataset that suits their needs, budget and technical capacity.

1.3 Relationship to Existing Standards

The Water Utility Data Standard integrates with existing Kansas GIS Geospatial Data Standards. Several resources were used to develop these standards, along with the working knowledge of the committee participants. Geospatial data standard documentation and data models that were used to support this effort include:

- *Kansas Public Water Infrastructure Standard, March 2003, Draft;*
- *ESRI Simplified Water Utilities Database;*
- *Water Utility Program (WatUP) Mapping and Database Standard, Draft 1.0- Arkansas Geographic Information Office;*
- *Standard for Water, Wastewater, Storm Drain infrastructure, Levels I and II version 1, Massachusetts Office of Geographic and Environmental Information;*
- *Kansas Geodata Compatibility Guidelines;*
- *Kansas Geospatial Metadata Standard;*
- *Kansas Hydrography Standard;*
- *Kansas Geospatial Data Addressing Standard.*

1.4 Description of Standard

This document is intended to promote data consistency between public water suppliers that produce and maintain water utility data. The Water Utility Data Standard categorizes and defines specific water utility data types, cartographic representations, and feature attribution. In addition, the standard addresses issues common among all digital GIS resources, such as encoding, quality control, data maintenance, and metadata.

The Water Utility Data Standard is not intended to address all public water supply infrastructure data collection efforts at the scales anticipated by local or private entities. However, this document can be used to establish checklists to assist persons involved in GIS data collection for water utilities and provide background material for those implementing a GIS. This document can also be used to establish guidelines for more refined features which may be appended to this document.

1.5 Applicability and Intended Uses of Standard

The Water Utility Data Standard is intended to support the automation, integration, and sharing of water utility infrastructure data. It is intended to be usable by public water suppliers, all levels of government, and the private sector, to achieve consistency in the digital representation of water utility features.

In preparing this document, care was taken to devise standards that are:

- Simple, easy to understand, and as logical as possible
- Uniformly applicable, whenever and wherever possible
- Flexible and able to accommodate future additions
- Dynamic in terms of continuous review

The Standard is not intended to be a substitute for an implementation design. An implementation design requires adapting the structure and form of these definitions to meet specific application requirements.

1.6 Standard Development Procedures

1.6.1 Participants

A core group of participants was involved in the development of this standard. In this process, comments and suggestions were solicited from a larger group representing water industry personnel, engineering firms, surveyors, state agencies and GIS-related businesses through open forums, conferences, business meetings, and contracts.

1.6.2 Process

The form and content of the Water Utility Data Standard was modeled after the *Kansas Geodata Compatibility Guidelines* template.

1.6.3 Comments and Reviews

Meetings, where comments or reviews were collected, took the form of open forums, public presentations, contract consultations, or small discussion groups.

1.7 Maintenance of Standard

An ongoing maintenance process to meet user needs and to integrate future standard requirements is necessary. It is expected that maintenance responsibilities will be shared by federal, state, and local agencies in compliance with the guidelines established by the Kansas GIS Policy Board.

2. Body of the Kansas Water Utility Data Standard

2.1 Technical/Operational Context

2.1.1 Data Accessibility and Maintenance

Just as quality assurance and reliability are inherent responsibilities for persons developing digital spatial datasets, so too are data access and data maintenance. A spatial data developer has the responsibility to store the data files and make them available, in a useable electronic form, to sanctioned users. Developers also assume the responsibility to update and maintain the spatial data or pass that role onto a reliable data custodian. With the creation of spatial data, the developer assumes the responsibility to document the data content with a conventional metadata format that meets the state's standards.

2.1.2 Reference systems

To support data exchange, all data must carry documentation for the coordinate reference system, projection, datum, and units of measure. Recommendations for reference systems can be acquired from DASC. Where applicable, the most current horizontal and vertical datums should be used. Recommendations for reference systems can be acquired from the state GIS clearinghouse, DASC. Where applicable, the most current horizontal and vertical datums should be used. The Kansas Geospatial Metadata Standard documents the form and content for reporting reference information.

2.1.3 Global Positioning System Data Collection

Global positioning system (GPS) technology can significantly enhance the spatial accuracy of field data collection and geospatial data development. Metadata documentation on GPS equipment, procedures, and processing is recommended to support data exchange and archival for geospatial datasets generated from GPS data. U.S. Environmental Protection Agency (EPA)'s Method Accuracy Database (MAD) structure illustrates the battery of information recommended to document GPS location information. At a minimum, geospatial developers should report the following information, if employing GPS collection techniques:

- Type and accuracy of GPS unit;
- Collection date;
- Collection method;
- Feature description;
- Coordinates;
- Reference system;
- Datum.

2.1.4 Integration of Themes

Data developers should take into account the implications of theme integration when developing spatial datasets. Planning decisions made at the beginning of a project can simplify future use. Choices of scale and extent should be selected with consideration for regional or state collaborations. Guidelines or suggestions for these issues can be obtained through DASC or the Kansas Information Technology Office.

2.1.5 Encoding

Encoding, as addressed in this section, describes how physical or map features are converted into an electronic form. There are two basic encoding data formats, vector and raster.

A vector data structure is a collection of digital graphic elements (points, lines and polygons) that represent the shape of surface features. Descriptive information for each feature (attributes) is usually recorded in a related database table. Vector data structures are regularly used to represent linear datasets, such as water utility infrastructure, to take advantage of the finer positional resolutions possible with this file format.

A raster data structure uses a grid system to characterize surface features. Each cell in the grid represents a prescribed surface area. Raster data structures are often used to represent continuous surface features, such as temperature or elevation, each file describing one surface characteristic.

Encoding or digitizing controls such as minimum mapping units, scale generalization, ground control, cell resolution, spatial extent, or coordinate system, affect the accuracy of the spatial dataset. Tested and reliable encoding controls are necessary to ensure quality representation of water utility infrastructure features in either a vector or a raster data structure. Options for encoding that might be considered for water infrastructure include:

- Manual digitizing;
- Heads-up digitizing over scanned engineering drawings;
- Heads-up digitizing over digital orthophotography with engineering drawing reference;
- Automated raster/vector conversion of scanned engineering drawing;
- Field data collection with GPS receivers.

Manual digitizing, though the oldest form of vector encoding, is accompanied by some serious disadvantages, including training and personnel costs; variation in the quality and the composition of source media; and lack of decent reference points to register the source media. Heads-up digitizing over scanned engineering drawings suffers from the same source media disadvantages as manual digitizing. The only difference is that the media is scanned into an electronic form. This process stabilizes the shrink-swell concerns but exchanges those problems for image file storage issues and the fallacies of human-interpreted scanned images. Heads-up digitizing over digital orthophotography with engineering drawing reference often requires substantial interpretation of engineering drawings, since many source documents are design drawings and not as-builts. The cell resolution of the orthophotography also impacts the accuracy of the end product. Automatic raster/vector conversion of scanned engineering drawings is still an evolving technology, getting progressively better. However, the quality of the source media is still a major obstacle for automatic conversion software, trying to differentiate between blue lines, creases, and water stains. Field data collection with global positioning units requires time, training, and equipment. Costs rise given that the more features collected; the more time spent in the field, and if only point features are collected then additional processing is required to add water line structures.

A hybrid encoding approach to data development is recommended for public water infrastructure. Upon consultation with local GIS users and engineering firms, the hybrid approach combined two of the encoding methods previously described: heads-up digitizing over digital orthophotography with engineering drawing reference and field data collection with GPS. The hybrid approach acquired point features with GPS receivers. These features were superimposed over digital orthophotography where line features were added in a heads-up mode.

Topology is defined as the spatial relationship between connecting or adjacent spatial features. The accuracy and reliability of the vector or raster data structure topology is impacted by the encoding controls used to develop the digital spatial data.

2.1.6 Resolution

In general practice, geospatial data are collected from map sources at specific resolutions (scales). Data collected at higher resolutions can reduce data duplication, since, with the advent of digital geospatial data, lower resolution data can be extracted from data collected at a higher resolution. The resolution or scale at which data are collected is an important piece of development documentation that must be referenced in the metadata. Resolution information for each data theme should be annotated as well to facilitate comparison and aggregation of datasets produced from different sources.

2.1.7 Accuracy

The attribute accuracy and the spatial accuracy of geospatial data must be documented in the metadata accompanying the geospatial dataset. Metadata elements relating to accuracy are referenced in the Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata (CSDGM). Another reference source for data accuracy on GPS source information is the EPA MAD file structure (see Section 2.1.3).

Attribute accuracy addresses the quality and extent of the information assembled to describe water utility features. Acceptable field values, quality assurance statements, and tolerance factors define the type of applications for which data can be used or for which additional data collection is warranted.

Spatial accuracy is specific to the encoding method employed to derive digital information from source documents or the geodetic controls used to orient the water utility features to real world coordinates. Absolute accuracy reflects how well the geospatial dataset represents utility features and should comply, where applicable, with the National Standard for Spatial Data Accuracy developed by the Federal Geographic Data Committee (June, 1998) which has been adopted into the Kansas Statewide Technical Architecture. Following the guidelines of this standard, accuracy statements to support the hybrid encoding methodology described in Section 2.1.5. (Encoding) were developed and documented in the report, “Kansas Implementation of 'FGDC Geospatial Positioning Accuracy Standards, Part 3: National Standard for Spatial Data Accuracy,'” that accompanies this standard or can be acquired through DASC.

Relative accuracy qualifies the positional integrity between and among water utility features. Affected by the level of absolute accuracy, relative accuracy is also influenced by projection distortion. (Note: the National Standard for Spatial Data Accuracy (NSSDA) developed by the Federal Geographic Data Committee in June 1998 has superseded the National Map Accuracy Standard established by the U.S. Office of Budget in 1947.)

2.1.8 Edge Matching

As implied by the term, edge matching is the process of matching continuous spatial features across a contiguous spatial dataset tiling scheme. This is becoming less of an issue as GIS software and computing technology are better able to handle large continuous datasets. River and stream channels extending across tile boundaries are examples of continuous spatial features, whereas an example of a spatial dataset tiling scheme is the 7.5-minute USGS quadrangle boundaries. If tiling is necessary, data developers will want to address edgematching.

The quality of the edge matching process is affected by the accuracy and encoding standards established for the geospatial dataset being joined together. Quality control tolerances for edge matching should be consistent with the absolute and relative accuracies established for the geospatial dataset (see Section 2.1.7). Exceptions or variances to these tolerance factors (if any) should be noted in a metadata document.

2.1.9 Feature Identification

A feature is a real world phenomenon that is whole and not divisible into phenomena of the same kind. In general, basic features are unique and not made up of other features. Compound features are composed of basic features. Feature categories include points, lines, and polygons.

Each feature, either basic or compound, should be uniquely identified. The use of existing or complementary identification systems is a fundamental concept of standards development to provide a link among spatial and tabular datasets. Where possible, the use of existing identification systems or the development of logical extensions is recommended. At present, a review of the state GIS core databases, housed at the state GIS clearinghouse, will help identify contact names to query about feature identification systems. It is the responsibility of the developer to either seek out existing identification or develop identification systems for their data development efforts.

2.1.10 Attributes

Attributes describing public water infrastructure features provide essential data to the public water supplier. The features and attributes referenced in Appendix B, “Comprehensive Water Distribution Data Model,” suggest one schema to describe water utility infrastructure. The attributes for these features should, at a minimum, include a statement regarding how accurately this information represents the public water infrastructure features in terms of completeness, logical consistency, and currency.

2.1.11 Transactional Updating

Transactional updating shall be a function of the data custodian. Documenting changes to a feature shape or attribution builds data reliability. The data developer is also responsible for establishing an update cycle, whether contingent upon available funding or part of a regular maintenance schedule. This schedule, along with the last revision date, should be incorporated into the metadata documentation.

It is recognized that documentation to substantiate changes on feature shapes or attributes can become too burdensome for even the most sophisticated information systems. It is recommended that basic guidelines be established by the data developer for the following aspects of transactional updating: date of change; identification of full or partial change; type of change, either shape or attribute; purpose for change; identification tag to link original and revised data; entity and person(s) requesting the change.

2.1.12 Records Management

By its very nature, a water utility is in a constant state of flux, and correspondingly, so is the geospatial dataset representing that system. While transaction updating covers changes occurring in the artificial system, record management addresses changes in the constructs used to organize and maintain the geospatial database. The "SDE Migration Prototyping for Geometric Network Geodatabase based on Simplified Water Distribution Data Model" case study, accompanying this document, describes archive options associated with the geodatabase format. Regularly scheduled archives or versions help ensure data continuity as changes occur to database constructs, such as file structures, identification schemes, transaction functions, file categories or naming conventions. Historical records management is critical during the formative development of a spatial dataset and is a routine aspect of continued data maintenance. Data developers should anticipate requests for historical versions of their data and retain archive copies that can be made available upon request. The state GIS coordinator and DASC are also encouraging data developers to archive geospatial datasets with that agency as an option for contingency off-site data recovery.

2.1.13 Metadata

Metadata summarize the development history, data content, maintenance process, accuracy statements and applicable use of a spatial dataset. Metadata completed for water Utility datasets should comply with the Content Standard for Digital Geospatial Metadata (CSDGM) and be maintained and updated by the data developer.

2.1.14 Map Production

With current technology, map products are assuming many forms. Paper map products have been joined by a variety of web and composition map options. Irrespective of map distribution method, data developers still need to meet some basic obligations, such as:

- Provide a range of map formats to support varying user needs and capabilities;
- Support standardized cartographic symbology to minimize misinterpretation;
- Provide templates for users unfamiliar with cartographic conventions;
- Where possible, control viewable thresholds to match the scale of digital datasets used in the map composition.
-

2.2 Data Characteristics

The "Minimum Requirements for Water Utility Features" (Appendix A) represents the basic level of information that is currently collected for most small public water suppliers.

A more robust, complex data structure is under development and will be found in Appendix B, Comprehensive Water Distribution Data Model. This data structure will be a collaborative effort of GIS user community and the large public water suppliers who are, or will be, using the data structure.

2.2.1 Minimum Required Water Utility Geospatial Features

See Appendix A, Minimum Requirements for Water Utility Features

2.2.2 Minimum Required Water Utility Attribute Elements

See Appendix A, Minimum Requirements for Water Utility Features

2.2.3 Optional Water Utility Geospatial Features

See Appendix B, Comprehensive Water Distribution Data Model (under development)

2.2.4 Optional Water Utility Attribute Elements

See Appendix B, Comprehensive Water Distribution Data Model (under development)

Appendix A

Minimum Requirements for Water Utility Features

Although most public water suppliers are interested in digitally mapping their water utility infrastructure, many do not have the capability of developing and maintaining a fully functional distribution network. These minimum requirements provide data that can be used to produce maps of a public water supply system to show the system infrastructure in relation to physical features on the ground; it can also be used to locate underground infrastructure such as valves and pipelines.

Geospatial Features:

Cleanouts	Meters	Valves
Hydrants	Other	Mains

Attribute Elements:

Cleanouts Feature Type: Point				
Column Name	**Type**	**Length**	**Domain Name**	**Column Definition**
CleanOutID	Text	20	N/A	Primary Key. A unique user defined identifier for each record or instance.
Comments	Text	100	N/A	Special information about the record.

Hydrants Feature Type: Point				
Column Name	**Type**	**Length**	**Domain Name**	**Column Definition**
HydrantID	Text	20	N/A	Primary Key. A unique user defined identifier for each record or instance.
Comments	Text	100	N/A	Special information about the record.

Mains Feature Type: Line				
Column Name	**Type**	**Length**	**Domain Name**	**Column Definition**
LineID	Text	20	N/A	Primary Key. A unique user defined identifier for each record or instance.
LineType	Text	13	Main_Line	Values that differentiate the general use of a water pipe.
LineSize	Numeric	9	Main_Size	Allowable input values for diameter of water pipe.
LineMaterial	Text	12	Main_Material	Allowable input values for construction material of water pipe.
Comments	Text	100	N/A	Special information about the record.

Meters Feature Type: Point				
Column Name	**Type**	**Length**	**Domain Name**	**Column Definition**
MeterID	Text	20	N/A	Primary Key. A unique user defined identifier for each record or instance.
MeterSize	Text	9	Met_Size	The manufacturer’s designated size.
HouseNum	Text	50	N/A	Address.
Comments	Text	100	N/A	Special information about the record.

Valves Feature Type: Point				
Column Name	**Type**	**Length**	**Domain Name**	**Column Definition**
ValveID	Text	20	N/A	Primary Key. A unique user defined identifier for each record or instance.
ValType	Text	23	Val_Type	Allowable values for type of valve.
Line	Text	15	Val_Line	Allowable input values for description of the line to which the valve is attached.
ValveSize	Numeric	9	Val_Size	Allowable input values for diameter of valve.
Comments	Text	100	N/A	Special information about the record.

Other Feature Type: Point				
Column Name	**Type**	**Length**	**Domain Name**	**Column Definition**
OtherID	Text	20	N/A	Primary Key. A unique user defined identifier for each record or instance.
OtherType	Text	9	Oth_Type	Allowable input values for other infrastructure.
Comments	Text	100	N/A	Special information about the record.

Domain Tables:

Domain Table Name Main_Line	Definition
value list - line type	Allowable input values for type of water pipe.
Value	**Definition**
Main	Main line
Hydrant	Hydrant Line

Domain Table Name Main_Size	Definition
value list - line diameter	Allowable input values for diameter of water pipe.
Value	**Definition**
0.25	0.25 inch diameter
0.50	0.50 inch diameter
0.75	0.75 inch diameter
1.00	1 inch diameter
1.25	1.25 inch diameter
1.50	1.50 inch diameter
1.75	1.75 inch diameter
2.00	2 inch diameter
2.50	2.5 inch diameter
3.00	3 inch diameter
4.00	4 inch diameter
5.00	5 inch diameter
6.00	6 inch diameter
8.00	8 inch diameter
10.00	10 inch diameter
12.00	12 inch diameter
14.00	14 inch diameter

15.00	15 inch diameter
16.00	16 inch diameter
18.00	18 inch diameter
20.00	20 inch diameter
21.00	21 inch diameter
22.00	22 inch diameter
24.00	24 inch diameter
28.00	28 inch diameter
30.00	30 inch diameter
32.00	32 inch diameter
36.00	36 inch diameter
Other	Other diameter

Domain Table Name Main_Material	**Definition**
value list – line material	Allowable input values for construction material of line
Value	**Definition**
ABS	Acrylonitrile butadiene styrene
AC	Asbestos Concrete
CI	Cast Iron
C900	C900 PVC
CU	Copper
DI	Ductile Iron
GI	Galvanized Iron
PE	Polyethylene
PVC CL160	Poly Vinyl Chloride Class 160
PVC CL200	Poly Vinyl Chloride Class 200
PVC CL250	Poly Vinyl Chloride Class 250
PVC Sch 40	Poly Vinyl Chloride Schedule 40
PVC Sch 80	Poly Vinyl Chloride Schedule 80
PVC Oth	Poly Vinyl Chloride Other or Not Specified
S	Steel
YM	Yellow Mine
Other	

Domain Table Name Met_Size	**Definition**
value list - meter size	Allowable input values for diameter of valve.
Value	**Definition**
3/8"	3/8 inch diameter
1/2"	1/2 inch diameter
5/8"	5/8 inch diameter
3/4"	3/4 inch diameter
1"	1 inch diameter
2"	2 inch diameter
Other	other diameter

Domain Table Name Val_Type	**Definition**
value list - valve type	Allowable input values for type of valve.
Value	**Definition**
Gate	Main line valve
Air Release	Air Release Valve
Butterfly	Butterfly Valve
Curb Stop	Curb Stop
Pressure Reducing	Pressure Reducing Valve
Other	

Domain Table Name Val_Line	**Definition**
value list - valve line	Allowable input values for description of the line to which the valve is attached.
Value	**Definition**
North	North line
South	South line
East	East line
West	West line
Bypass	Bypass line
Check	Check valve line
Hydrant	Hydrant valve line
Main	Main line
Private	Private line
Service	Service line
Other	

Domain Table Name Val_Size	**Definition**
value list - valve diameter	Allowable input values for diameter of valve.
Value	**Definition**
1.00	1 inch diameter
1.50	1.5 inch diameter
2.00	2 inch diameter
2.50	2.5 inch diameter
3.00	3 inch diameter
4.00	4 inch diameter
5.00	5 inch diameter
6.00	6 inch diameter
8.00	8 inch diameter
10.00	10 inch diameter
12.00	12 inch diameter
16.00	16 inch diameter
20.00	20 inch diameter
24.00	24 inch diameter
30.00	30 inch diameter
36.00	36 inch diameter
Other	

Domain Table Name **Oth_Type**	**Definition**
value list - other infrastructure	Allowable input values for other infrastructure.
Value	**Definition**
Source	Raw water source (well, surface intake, spring)
Interconnect	Connection between water systems for delivery of raw or treated water
Tower	Storage tank
Pit	Any pit or vault containing infrastructure controls.
Station	Pump, booster, chlorination, telemetry or other station
Plant	Treatment plant
Outfall	Drinking water treatment plant wastes that outfall to surface waters of the State
Waste_Trt	Treatment of plant wastes, such as softeners and filtering (lagoon, mechanical)
Other	

Appendix B
Comprehensive Water Distribution Data Model

The Comprehensive Water Distribution Data Model provides a robust framework for fully integrating GIS technology into the operations of a public water supply system. It includes significant attribution and a distribution network for those public water suppliers interested in managing infrastructure records and spatial information associated with their water utility. This model is currently under development.