North Carolina Department of Transportation (NCDOT)

Base Geospatial Standard

Date	Change Description	Approved By
12/3/2010	Initial document (version1.0)	GIS Implementation Team / GTAC
8/18/2011	Consistency with the GIS Data Policy Standard	GTAC
11/21/2011	Corrections and updates to referenced Standards	GIS Unit

1. INTRODUCTION

Data Governance of a dataset refers to the operating discipline for managing data and information as a key enterprise asset. Key aspects of data governance include such topics as decision making authority (roles and responsibilities), compliance monitoring, policies and standards, data inventories, full lifecycle management, content management, records management (metadata), preservation (archival), data quality, data classification (data model), data security and access, data risk management and data valuation. This standard is the first in a group of standards that defines data governance for the geospatial datasets within the NCDOT

At a high level, the processes and standards described in this document refer to transportation features which can have geographic locations and characteristics. As the NCDOT matures to include geospatial data in its everyday processes, exchange of data is a natural consequence of that maturity. Indeed, it makes sense for the business to efficiently and effectively exchange information. However, data exchange without standards can lead to excessive amount of rework of GIS datasets, and also loss of critical information about the quality of that dataset (related to how it was developed and the accuracy and timeliness of that data). The set of standards, and processes referenced in them, ensures optimal use of business resources in the delivery of information to the business.

1.1. OBJECTIVE

The objective of this standard is to create a baseline for the development and documentation of spatial (or geographic) datasets within the NCDOT. A corresponding objective is to establish a central location, available to the entire organization, that would serve as a resource center for obtaining geospatial standards,

user guides, processes, links to recommended datasets, and coordination with the NCDOT GIS community.

1.2. SCOPE AND APPLICABILITY

This standard defines geospatial data, minimally required elements that encompass recommended and best practices, and backup and archival requirements for business-critical spatial datasets. The primary purpose of this part of the standard is to identify the necessary processes and standards to ensure quality and integrity of geospatial information to meet the needs of business operations within the NCDOT.

The Base Geospatial Standard integrates with the other NCDOT geospatial standards: Geospatial Metadata, Geospatial Data Model, Geospatial Data Acquisition, and Geospatial Data Publication. It is expected that all initiatives or parts of the NCDOT organization that create/produce, maintain or acquire geospatial data for enterprise use within the NCDOT shall satisfy the requirements laid out in these standards. A partial list of geospatial datasets used by the organization can be found in Appendix A.

2. CONFORMANCE

2.1. This standard is part of a group of standards which are required for all geospatial datasets within the NCOT. Conformance to the group of standards includes conformance to the items listed in this base standard, and appropriate other parts for metadata, data model, data acquisition and/or data publication.

2.2. Conformance to this standard ensures that the spatial data used and maintained by the NCDOT delivers business value to the organization, that the data is able to be seamlessly integrated into other business datasets, that the information is maintained and timely in order to support strategic and tactical decisions, and that the data is consistently the best data available.

3. NORMATIVE REFERECES

This standard is meant to work in conjunction with the following NCDOT spatial data standards:

- Geospatial Metadata Content Standard
- Geospatial Data model Content Standard (in development)
- Geospatial Data Acquisition Standard (in development)
- Geospatial Data Publication Standard (in development)

This standard is compatible with the following North Carolina state standards (http://ncgicc.net/Standards/tabid/141/Default.aspx.)

- North American Vertical Datum of 1988
- Statewide Global Positioning System (GPS) Data Collection and documentation standards
- Geographic Data Content Standard for Transportation Roads Data
- Content Elements for Statewide publication of core geospatial parcel data
- Geographic Data Content Standard for Water Distribution and Sanitary Sewer Systems, Version 2.2
- A Standard Classification System for the Mapping of Land Use and Land Cover

This standard also addresses the topic of archival of GIS datasets that is more fully explained in the November 19, 2008 final report to the North Carolina Geographic Information Coordinating Council (GICC) on Archival and Long Term Access Policy (http://www.ncgicc.com/Portals/3/documents/Archival_LongTermAccess_FINAL11_08_GICC.pdf.)

3.1. STANDARDS DEVELOPMENT PROCESS

The process for development of this standard shall be in accordance with policies of the NCDOT GIS Steering Committee.

3.2. MAINTENANCE AUTHORITY

3.2.1 **Level of responsibility:** The GIS Unit and GIS Technical Advisory Committee at NCDOT are responsible for maintenance of this standard.

3.2.2 **Contact Information:** Address questions concerning this standard to:

GIS Manager, NCDOT GIS Unit
NC Dept of Transportation
4101 Capital Blvd,
Raleigh, NC 27604
Telephone: (919) 707-2151
jcfarley@ncdot.gov

4. TERMS AND DEFINITIONS

4.1. METADATA

Metadata is information that describes the content, quality, condition, origin and other characteristics of data. Metadata for spatial data may describe and document its subject matter; how, when, where and by whom the data was collected; availability and distribution information; its projection, scale, resolution, and accuracy; and its reliability with regard to some standard.

4.2. DATA MODEL

A data model is a computer representation of the data used by the business, along with processes that the business performs. It includes such items as spatial representation of business objects (e.g., a road, imagery), attributes; integrity rules and relationships; cartographic portrayal; and metadata requirements. A common model governs how well software applications will perform using an organization's data.

3.5 VECTOR SPATIAL DATA

Spatial data can be represented generally in two ways: Vector and raster. Vector data is a representation of the data in a geometric form, such as point, line or area, and referenced to a location on the earth. Most real-world objects (such as a landmark, a road or a lake) are represented as a vector.

3.6 RASTER SPATIAL DATA

Spatial data can be represented generally in two ways: Vector and raster. Raster data is represented as a collection of grids with values, and referenced to a location on the earth. Most images are represented as raster.

3.7 LINEAR REFERENCED DATA

Geospatial data that is stored by using a relative position along an already existing line feature. In linear referencing, location is given in terms of a known line feature and a position, or measure, along the feature. Methods of referencing include (1) milepost along a unique route, (2) direction and distance offset from the intersection of two unique routes, (3) percent of total distance along a unique road segment, (4) direction and distance along a unique road segment. Linear referencing is an intuitive way to associate multiple sets of attributes to portions of linear features.

3.8 REFERENCE SYSTEM

A system for identifying positions on the globe. This is often constructed with a grid that either refers to the earth's latitude and longitude (graticule), or a planar

equivalent that divides grid lines by a fixed length from a predefined point of origin. It is common practice to use coordinates to reference a location, using both the latitude/longitude reference system, as well as the North American Datum reference system.

3.9 ARCHIVAL

The long term collection and maintenance of data snapshots retained permanently that can be utilized to help manage long term risk (i.e. regulatory/ legal requirements) while allowing ongoing access to authentic historical data for the purposes of analysis.

3.10 BACKUP

The capture of active datasets with the intention of providing a means to restore changing records that have been deleted or destroyed. The purpose of a backup is to manage short term risk and address disaster recovery. Typically these snapshots are only retained for a few days or weeks before being overwritten by newer snapshots.

5. REQUIREMENTS

5.1. SPATIAL REFERENCE:

All enterprise spatial data shall use a standard spatial reference and be spatially referenced. The references listed below are NCDOT standards.

5.1.1. North American Datum 1984: The standard spatial reference for NCDOT vector and raster data is North American Datum 1984. Measurement units are in feet. The NCDOT's LRS is referenced to this common spatial reference.

5.1.2. Linear Reference Standard: Most data related to the NCDOT's business operations is referenced to the road network, commonly called the Linear Reference System (LRS). Data that uses this reference shall use a linear feature referenced to the NCDOT standard reference of the North American Datum 1984.

5.1.3. Other Reference Systems: Addresses and ownership identification (Tax parcel ids) are other commonly used references within the NCDOT. In general, address references shall be transformed to the LRS or standard spatial reference before they are used by the business. Coordination with the GIS unit is recommended to accomplish the transformation.

5.2. REPRESENTATION OF FEATURES:

All enterprise spatial data shall use a standard spatial representation. Features referenced below are NCDOT standards. These do not refer to the cartographic representation, which describes how the data is represented on a map, not how they're represented in a database.

5.2.1. Road network: The standard spatial representation for the NCDOT road network is as a vector linear feature.

5.2.2. Administrative Boundaries: The standard spatial representation for the administrative boundaries is as a vector area feature. Examples of this type of feature are municipal, county, state, and census boundaries.

5.3. REPRESENTATION OF CHARACTERISTICS:

All enterprise spatial data related to the road network shall follow a standard representation of characteristics. The NCDOT standard for representing most characteristics referenced to a road are as linearly referenced events. That means that the data is usually kept in tables that reference line features. There are four standard methods for referencing characteristics along the road network: (1) milepost along a unique route, (2) direction and distance offset from the intersection of two unique routes, (3) percent of total distance along a unique road segment, (4) direction and distance along a unique road segment. Another NCDOT standard for representing characteristics is known as Coordinate reference.

5.3.1. Route milepost reference: Includes the unique route number and a milepost value for the feature. Example: A crash site is located at milepost 18.9 on route 16. The number of lanes on route 24 is 2 from milepost 26 to milepost 78.

5.3.2. Intersection offset: direction and distance offset from the intersection of two unique routes or road segments, and an offset along a particular route or road segment. If the characteristic is of a linear nature (e.g. an event that starts at one point and continues for some distance), then a distance along the road network is included. Examples: a sign would be a point event, located 13 feet east of the intersection of routes 1 and 2, along route 2. Number of lanes would be a linear event, located 145 feet west of the intersection of routes 3 and 4, along route 4, running for 1.5 miles.

5.3.3. Percent along a road segment (G1FTseg): Includes the road network segment and a percentage distance along that segment. Could be either a point or a linear event. Examples: A sign is located 34% along road segment 19. The number of lanes along road segment 20 is 4, starting at 18.1% and ending at 69.9%.

5.3.4. Distance along a road segment (G1FTseg): Includes the road network segment and a distance along that segment. Could be either a point or a linear event. Examples: A sign is located 1.5 miles along road segment 19. The number of lanes along road segment 20 is 4, starting at 2.3 miles continuing for 3 miles.

5.3.5. Coordinate : Location of a feature referenced to a grid coordinate system. Commonly referred to as "lat/long" or "X/Y". Usually this is for a point feature, such as a sign. Example: The sign is located at Northing 35°40'15" and Westing -77°25'30" .

5.4. DATA QUALITY

All enterprise spatial data shall include descriptions of the quality of data as it refers to the items below. These descriptions shall be documented in the metadata for each spatial dataset.

5.4.1. POSITIONAL ACCURACY:

All enterprise spatial data shall have a known spatial accuracy. Positional accuracy is a statement of how closely the location of a feature represents a true position on the ground. An important component of positional accuracy is a statement of precision, whereby the exactness of the data is indicated.

5.4.2. ATTRIBUTE ACCURACY

Attributes are facts tied to the Earth's surface. Attribute accuracy is the closeness of attribute values to their true values. This applies to accuracy of continuous attributes such as elevation and accuracy of categorical attributes such as number of lanes or pavement type. An important component of attribute accuracy is a statement of precision, where values are measured.

5.4.3. COMPLETENESS

Completeness is defined as the degree to which the features and their attributes represent all the instances of the entire dataset. Missing data (incompleteness)

can affect logical consistency needed for correct processing and/or analysis of the data. A statement of completeness shall be documented in the metadata.

5.4.4. LOGICAL CONSISTENCY

A spatial dataset is logically consistent when it does not violate business rules and constraints for the dataset. Logical consistency is best expressed in the rules and constraints defined for a spatial data model. Data that meets rules and constraints (both spatial and non-spatial) can be used with a high level of confidence for business applications and analysis. Without consistency, additional time and effort will need to be expended to allow software to handle inconsistencies in ways that do not propagate or increase the errors. Logical consistency shall be documented in the metadata and reference the geospatial data model for the dataset.

5.4.5. LINEAGE

Lineage refers to the origin and processing history of a dataset. It includes the name of the organization that produced the data so that its policies, procedures and methods can be evaluated to see if they were biased in representing the surface of the Earth or its features. Lineage of the dataset shall be documented in the metadata.

5.5. METADATA

Metadata shall be accurate and complete, compliant to the NCDOT Metadata Content Standard for Geospatial Data (pending approval). Ownership of the dataset shall be clearly identified within the metadata, as will any publication constraints. Acquisition method shall be identified in the metadata. An automated tool to verify compliance with the metadata standard (Metadata Parser program), shall be provided to the user community by the GIS Spatial Data Management Group. Contact the GIS Unit for access to that tool.

5.6. GEOSPATIAL DATA MODEL

Data models used as part of the development of the datasets will be delivered along with the dataset. Documentation of the data model will follow the NCDOT Geospatial Data Model Content Standard (in development).

5.7. ACQUISITION

The method of acquisition of the dataset and maintenance plan will follow requirements specified in the NCDOT Geospatial Data Acquisition Standard (in development). This standard covers the following items:

5.7.1. Acceptable formats for development and acquisition of geospatial datasets.

5.7.2. Recommended sources for non-NCDOT data that may be of interest to the NCDOT business units.

5.7.3. Dependencies upon a dataset or any dependencies that a spatial dataset may have on other datasets will be clearly identified.

5.7.4. Identification of geospatial datasets critical for business operations

5.7.5. Identification of geospatial datasets as part of other applications, including documentation of workflow and dataflow for that application

5.8. PUBLICATION

Sharing of data will be in accordance with the specifications in the NCDOT Geospatial Data Publication Standard (in development). The NCDOT is subject to all state laws regarding sharing of spatial datasets. This standard covers the following items:

5.8.1. Methods of publication

5.8.2. Formats for publication

5.8.3. Restrictions or requirements for internal or external sharing of the spatial datasets.

5.8.4. Requirements for publication of datasets as services to be consumed by web browsers.

5.9. LOCATION AND RECOVERABILITY

The physical location of the dataset will be identified, as well as documentation of any backup and restore procedures. Any spatial dataset considered critical to the business functions of NCDOT will be clearly identified and a backup and recovery plan verified and implemented to ensure that NCDOT meets it's public obligations. In addition to backup and recovery, a plan for long term archival should be included

in the documentation associated with the dataset. (For more information on state activities related to archival of spatial datasets, please refer to the document in paragraph 3.0 above.) In lieu of an archival capability, backup datasets will be stored as archives, based on business needs and best practices. Geospatial datasets related to the road network will be backed up quarterly as part of the publication cycle.

6. BEST PRACTICES

It is highly recommended that developers of spatial datasets within the NCDOT work closely with the GIS unit in the organization to submit a draft of documentation and datasets for acceptance before final delivery.