# Implementing the Enterprise GIS in Transportation Database Design

*J. Allison (Al) Butler and Kenneth J. Dueker*

***Abstract:*** *Earlier articles by the authors described the primary database design approaches that have been and are being used in geographic information system applications for transportation. These articles proposed a theoretical model for an enterprise geographic information system environment for transportation agencies. This follow-up article provides detailed guidance on building a multimodal transportation feature database using relational database concepts.*

## Introduction

The development and use of geographic information systems (GISs) have been areas of activity in transportation agencies for many years. Yet, significant progress toward realizing the promise of GIS as an integrator of data and operations in these agencies has been rare. Legitimate differences in requirements frequently lead to application-specific definitions and representations of transportation features and their geometry. For example, vehicle navigation systems, overweight truck routing, facility location, address geocoding, and emergency management applications all have unique requirements. The challenge is to establish a means of data exchange among these disparate representations, one that leads to overall improvements in accuracy, consistency, and completeness as the users of each system perceive these qualities.

Deploying GIS in transportation (GIS-T) at an enterprise level presents an opportunity to eliminate the traditional application-specific development pattern of information systems by providing a common data structure centered on transportation features. The solution is to embrace the diversity of applications and data requirements within a unifying enterprise data model for GIS-T that allows each application group to meet the established needs while enabling the enterprise to integrate and share data. The primary objective of this model is to allow frequent transaction-based data exchanges and updates, the type that an interactive organization is likely to need. As transportation agencies move toward a more integrated manner of doing business, such as involving design units earlier in the project planning cycle, the need for data to cross former institutional or jurisdictional barriers will become greater.

## Overview

This article will show how a GIS-T database can be constructed to facilitate data exchange. The ultimate goal is to construct a single database that will be used by all applications without the need for formal data exchange mechanisms. The design focuses on a transportation feature table that implements the concepts expressed by the authors in earlier articles (Butler and Dueker 2000, 1998, Dueker and Butler 1999, 1998). These articles present theoretical and implementation aspects of an Enterprise GIS-T Data Model that includes four primary components:

1. a facility inventory comprising jurisdiction, transportation features, event points, linear events, point events, and intersections;
2. a network that includes nodes, links, traversals, and traversal segments;
3. a measurement datum consisting of anchor points, anchor sections, reference objects, and geographic points; and
4. cartography, which includes (in the simplified version discussed here) base map strings, linear event strings, line segments, point symbols, and cartographic points.

The list of components is essentially an "all of the above" response to the existing structures of legacy transportation information systems and current GIS product lines (Butler and Dueker 2000). At the same time, the unbundled approach of the system allows system designers to pick the elements that are needed for a particular application (Dueker and Butler 1998). For example, Want to draw a map of an existing highway inventory? It is necessary to:

1. Choose a GIS platform that supports dynamic segmentation.
2. Establish a one-to-one relationship between each transportation feature map object and its corresponding feature in the inventory. This is accomplished by constructing the object and then attaching the transportation feature identifier in the inventory database.
3. Assign measures for beginning and ending points for linear objects representing linear transportation features.
4. Use the dynamic segmentation function to create a cartographic object for each linear event or to position point events.

Another example would be to route an oversize/overweight truck. Assuming that the facility inventory has all the data needed to select route segments, the steps listed above will need to be done as well as the following:

1. Construct a node table of all points where a route segment can begin, end, or change from one transportation feature to another. Each node must have a unique identifier.
2. Create a link table containing all of the valid links. Each link must have a unique identifier and the related record must include the terminating node identifiers. An ordered-pair approach allows the ability to distinguish the direction of travel.
3. Extract a set of linear and point event table records containing relevant attributes, such as bridge width, height, and load capacity.
4. Join the event table extract(s) with the link table to produce a traversal segment table.
5. Apply minimum criteria to the traversal segment table to produce a list of segments that can be used for the trip.
6. Use the resulting table to create adjacent node entries for each node record. These entries will allow the node table to serve as a turntable at intersections and to access points. Eliminate all other node table records or mark them as invalid for this trip.
7. If desired, create an impedance factor for each traversal segment based on such considerations as traffic volume, posted speed limits, and/or turning movement volumes. This factor can be used to select the most suitable path from among the alternatives available.
8. Use software functionality to find the “best” path based on stated criteria, such as the shortest distance or lowest impedance.
9. Consider the designation of universal traversals that will be used for all long-distance travel, such as rural interstate highways. This practice will simplify the pathfinding process by limiting the work needed to finding a way onto and off of the long-distance facility. It will also help prevent the pathfinding process from accidentally routing a vehicle onto a less desirable facility due to data errors and omissions.

The two sets of guidelines listed above are simplified; in practice, the issues involved are a bit more complex. For example, most facility inventories are structured to use a single logical centerline for each roadway. The corresponding cartographic object is a line down the middle of the road or in the median of a divided highway. If the facility inventory stores data according to side of the road, then a mechanism that lets the user accommodate this structure during dynamic segmentation will need to be established. (We are not aware of any GIS software product that handles this process automatically.) Below are a few options to address this data structure and to provide a cartographic appearance that more closely matches feature geometry for divided highways:

1. Change the facility inventory to utilize a paved-course or travel way centerline (i.e., a logical centerline for each side of a divided roadway). This approach results in a set of three centerlines, one for undivided roads, one for the right-side path, and one for the left-side path.
2. Construct the same line objects as would be needed for Option 1 but do not change the facility inventory. Instead, presort the inventory by side of road into three groups (left, right, and both sides).
3. Continue to use the logical centerline to connect to the facility inventory but do not use it for map making. Instead, use paved-course centerlines related to the logical centerline through a foreign key. This approach is the most satisfying option since it does not require any changes to the facility database or preprocessing of the data; however, it does require some custom coding to transfer the results of dynamic segmentation to the correct adjacent paved-course centerline.
4. Continue to use the logical centerline to connect to the facility inventory and write a custom routine that produces parallel lines for divided highways. This only works for small-scale maps where the difference between the position of the generated parallel lines and the true paths will not be a problem.

Whatever option is chosen, the important starting point is to develop a set of business rules to guide the project. These business rules will include decisions about location referencing methods. We continue to support the use of spatially based linear location referencing systems (linear LRSs) for linear transportation features, such as highways, city streets, transit routes, and railroads. Multidimensional and temporal LRSs should be overlaid on the spatially linear LRS for these features. However, the LRS is not the most important set of business rules to be developed. That status is given to the rules governing transportation feature identification. It does not matter how a feature and attributes are measured if the user cannot find the transportation feature to which the attributes apply.

## Transportation Feature Identification

A transportation feature is an element of the transportation system that may be uniquely identified in the real world and for which attributes are provided (Butler and Dueker 1998). Since the central element of the multimodal Enterprise GIS-T Data Model is the transportation feature, there must be a way to uniquely identify the feature across all modal components. The model defines transportation features by type within a jurisdictional framework that serves no other purpose than to provide a common spatial context for expressing the location and extent of the facility. From an implementation perspective, more detailed specification is required, one that accommodates both the needs of the database and the public user.

A relational database management system needs a unique identifier for each record in a table. The design proposed below uses a relational foundation consisting of tables (relations) connected to each other by foreign keys (relationships). In most cases, these identifiers are integer numbers that serve no other purpose than to uniquely identify a single record and they are difficult for users to remember. For example, it is much easier to remember that a particular transportation feature is called SR 153 in Hamilton County, Tennessee than that its primary key identifier is 286449. The identifier known to the general user is called a public key. A form that we have suggested (Butler and Dueker 1998) for the transportation feature identifier public key would produce something like:

RDTN067ST00153000

where

RD is the transportation feature of type road;
TN is the State of Tennessee as a United States Postal Service (USPS) abbreviation;
067 is the Federal Information Processing Standard (FIPS) code for Hamilton County;
ST is the designating authority, this illustrates a method for identifying a state department of transportation (DOT);
00153 is the road identification number; and
000 is the sequence number for the section of 00153.

A working group under the auspices of the Federal Geographic Data Committee's Surface Transportation Subcommittee with support from the Bureau of Transportation Statistics has developed an alternative approach using many of the same concepts (FGDC 2000). This proposal is an extension to the National Spatial Data Infrastructure (NSDI) for uniquely identifying Framework Transportation Segments. Our sample public key expressed as a transportation segment identifier using this proposed standard could be:

47001.S.001530000

where

47 is the FIPS code for the State of Tennessee;
001 is the designating authority of the state DOT (undefined in the proposal but unique within the state);
S is the segment type of transportation feature; and
067153023 is the road identification number for section 23 of State Road 153 in Hamilton County (067). This number is unique among the designations made by the authority for this feature type (undefined in the proposal but unique within the state).

Regardless of the format used, the public key should be separate from the internal database key because the public key may change but the internal key must not—unless necessary to reflect the replacement of one feature with another. Used in this way, the primary mechanism for data sharing is reduced to simply adding a new column in a data table to store the public key. Once this equivalency has been established, future data exchanges become relational joins of an internal table with one provided by another person. Cartographic conflation would not be used, although it may have been of assistance in the original equivalency determination.

Two points need to be emphasized here. First, the proposed identification method is for data interchange or sharing and is not necessarily for internal databases. Each agency providing or using transportation data could continue to use identification unique to the agency. The proposed public identifier is for "translating" one identification system to another. For this system to work, each agency would need to establish a new field in their file to store the public key identifier, and then enter the appropriate identifier for each transportation feature.

This brings up the second point. It is very likely that each user will have different transportation features. This means that there will not be a one-to-one correlation between the transportation features defined using the public identifiers and those present in user databases. These differences are accommodated in both example methodologies listed above by the use of a three-digit segment sequence number. Differences can be accommodated as long as the segments comprise the same main feature. This means that all users will need to agree on transportation feature endpoints but not on intermediate break points. A state DOT may break transportation features at county lines, while an individual county or city breaks them at every intersection. The overall transportation feature identifier would be the one that ends in "000." As long as there are fewer than 999 segments, the proposed methodologies can accommodate any segmentation mechanism. Only at the highest level do the participants need to agree on transportation feature definition. Such agreements are required for any continuing data exchange process, so this is not really an extra burden imposed by the proposed methodologies.

## Linear Location Referencing System Datum

Before describing the database design, some guidance on establishing a linear LRS datum is in order. The datum is the framework within which a location is specified. Transportation agencies generally utilize an implicit linear LRS datum. The term "implicit" is used to mean that an origin and path are specified, but actual datum requirements and features for measurement accuracy are not usually expressed. We contend that an explicit datum consisting of a network of anchor points (well-defined physical locations) and anchor sections (the roadway segments between anchor points) must be established to provide the temporal and spatial measurement consistency required for reliable data exchange. Anchor points have location references as mandatory attributes. Anchor sections have direction and length as mandatory attributes. Such a datum design must provide a set of rules for defining, selecting, and locating anchor points and anchor sections, and for measuring the length of anchor sections. "Of particular concern are the identifiability and recoverability (persistence) of anchor points" (Vonderohe and Hepworth 1996,

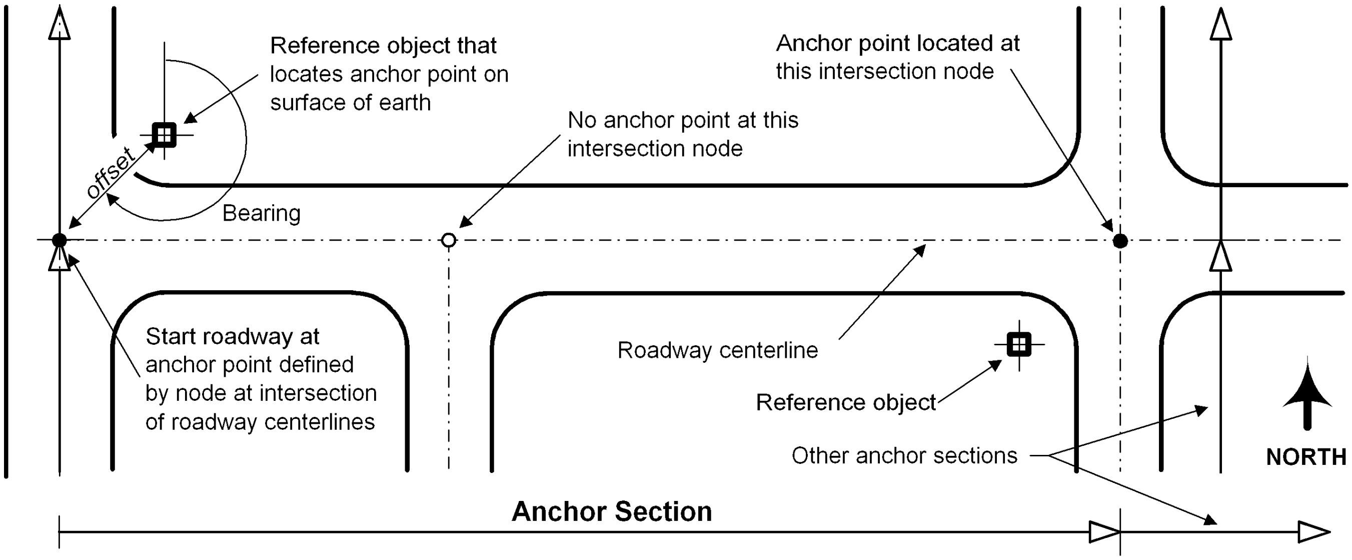

**Figure 1.** Roadway feature with anchor points, anchor sections, and reference objects.

p. 3). Simply put, transportation features and their attributes cannot be left to float around, changing from database to database and from year to year—unless a change is necessary to reflect a true change to the system.

While a linear LRS, such as route milelog, may locate events on linear transportation features, a means is needed to locate those features in the real world and to define the measurement framework. To meet these needs, a system of anchor points and anchor sections must be established (Figure 1). The position of each anchor point would be defined in all desired linear and nonlinear LRSs. Determining the location of anchor points is a good application of the Global Positioning System for GIS-T, since a Global Positioning System location for each anchor point can be used to tie the roadway to other coordinate systems. Anchor sections, which extend from one anchor point to another along the path of a transportation feature, have direction and length as their primary attributes. Anchor points and sections are geodetic objects that establish the geographic datum of the linear LRS. Anchor section length serves as a quality control check for the accuracy of linear LRS measurements. The anchor section follows the logical centerline of the roadway, with anchor points being located along that centerline.

Anchor points may be difficult to precisely identify and capture in the field, as they are usually located on an abstraction of the road (i.e., the roadway centerline). Anchor points need to be tied to reference objects, which are physical locations that a user can easily observe in the field and on maps (Vonderohe and Hepworth 1998). Reference objects can be anything that is not readily movable (e.g., a curb intersection, bridge end, traffic signal pole, or survey marker).

To establish a linear LRS datum, anchor points must minimally be placed at the beginning and end of each roadway. Intermediate anchor points may be located along this base anchor section. The density of anchor points is determined by the degree of positional accuracy desired; the greater the desired accuracy, the greater the number of anchor points needed (Vonderohe and Hepworth 1998). The ability to detect and correct measurement errors varies directly with the number of anchor points and inversely with the average length of anchor section.

To be valid, a datum must be tied to physical, real-world locations that are unambiguously defined. This would seem to eliminate such field references as county lines and other jurisdictional boundaries, since county line and city limit signs may not be properly and/or consistently placed. However, the Tennessee DOT linear LRS, for example, requires the origin to be at the beginning point of the road in the county. The reconciliation of these two needs is to tie the jurisdictional boundary to a reference point that is unambiguously defined (i.e., make the location of the beginning anchor point and transportation feature origin 0.000 at the jurisdictional boundary, but locate the boundary (and origin) as an offset from a reference point). The roadway is thus unambiguously tied to a datum-compatible location.

One or more anchor sections must be created to provide a geographic network reference for each roadway. Vonderohe and Hepworth (1998) provide a way to calculate existing measurement accuracy and to determine the number of anchor points needed to reach a given level of accuracy.

It is important to note the differences between accuracy, precision, and resolution. "Resolution" is the closest proximity of objects that can be represented as being at different locations. For example, the Tennessee DOT roadway inventory uses a resolution of 0.01 mile (52.8 feet), which means that, to be placed at different locations, two objects must be at least half that far apart for measurement rounding. Thus, this linear LRS has a resolution of 26.4 feet. "Accuracy" refers to the closeness with which a set of measurements approximates the true value, which cannot be absolutely known. Anchor points and anchor sections provide a way to reach a specified level of accuracy by controlling systematic error; however, accuracy is ultimately determined by measurement methods. "Precision" refers to the repeatability of these

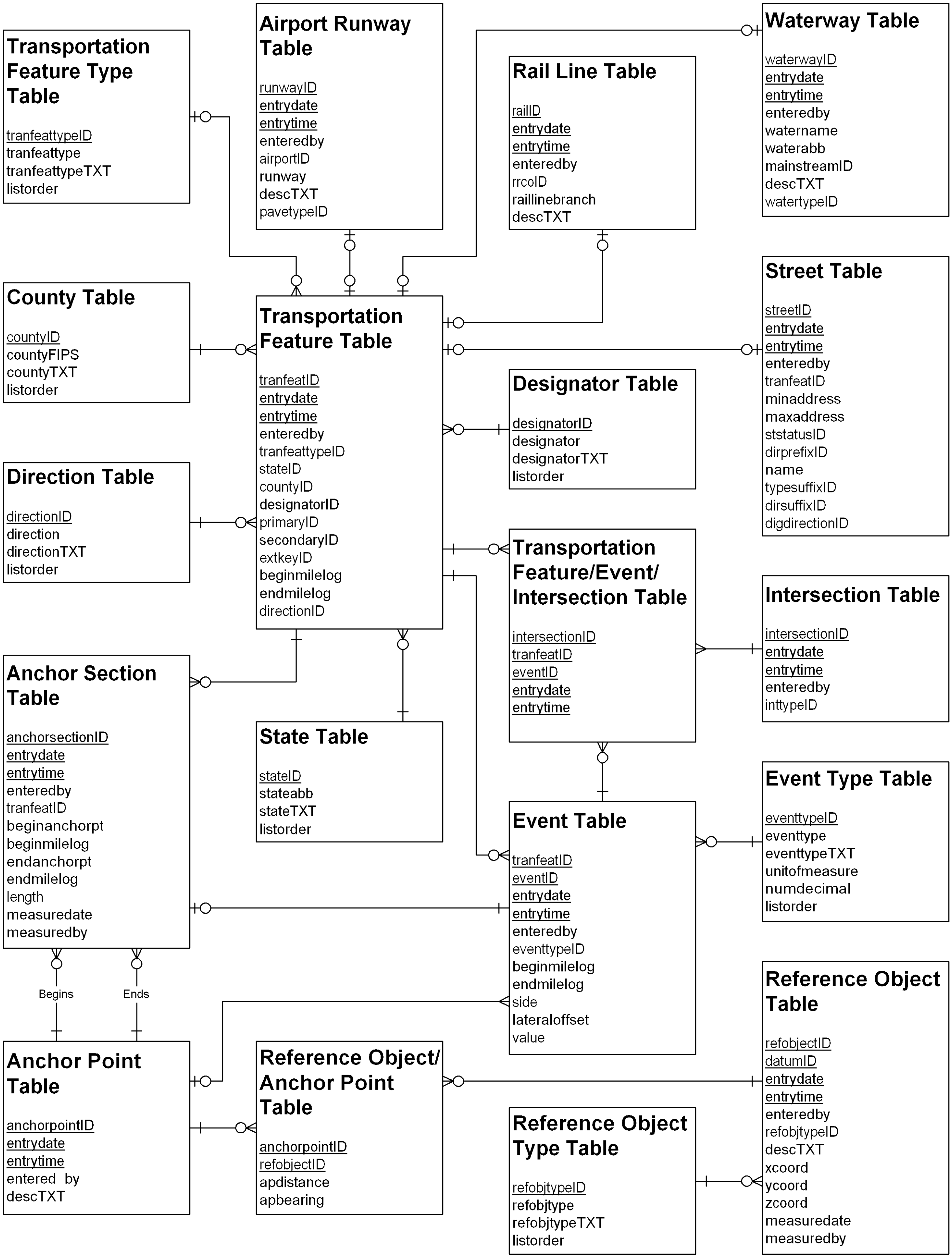

**Figure 2.** The Transportation Feature Database physical data model.

methods. For high precision, all measurements for the location of a feature are numerically similar. When precision is high, errors in measurement can be corrected through a uniform adjustment that compensates for the method's systematic error. When precision is low, errors cannot be readily corrected by adjustment because random error is much larger than systematic error.

An increase in the density and location precision of anchor points results in an increase in linear LRS measurement accuracy. The overall accuracy of the linear LRS is limited by the precision of linear offset measurements from anchor points to locations of interest along a transportation feature. Measurements made with distance-measuring instruments have errors that increase with distance (i.e., are said to propagate). Anchor point accuracy requirements should be determined by looking at both the needs of users and the ability of available field procedures. A multimodal linear LRS datum must meet the highest accuracy, precision, and resolution requirements of any modal database and application that it is intended to support. Applications requiring lower levels of spatial and temporal accuracy and precision can utilize resolutions expressed with fewer degrees of certainty.

## Transportation Feature Database Design

The primary transportation feature-attribute tables needed for a complete multimodal transportation facility inventory form the central element of the proposal (Figure 2). The included transportation features in our sample database design are roadways, airport runways, waterways, railroads, and intersections. The primary key of each table is underlined. The primary keys of data tables are designed to store history through the inclusion of a time stamp (entrydate and entrytime). A field is also provided to record the name of the person who made the entry. The Transportation Feature Table includes a data item called the extkeyID, which is an external key identifier to link this table with an external data table. For instance, the extkeyID could contain the waterwayID for linking a water-based transportation feature record to the Waterway Table. The value of the tranfeattypeID will determine the feature table to which the identifier in extkeyID is related. This approach allows full normalization of the database using look-up tables and simplification of the naming processing. The included tables are described in the following sections.

Being centered on physical transportation features, the design treats utilizing modes as events. For example, a transit route would be a traversal across one or more transportation feature segments (raillines and/or roadways), with each segment defined as a linear event on a transportation feature. A useful design for traversals and other elements of a complete GIS-T database is provided in our previous work (e.g., Dueker and Butler 1998). Time stamps are provided to support temporal applications, such as the evaluation of traffic accidents based on the nature of the roadway network at the time of the events. "Person stamps," in the form of user identifiers, are provided as a managerial means of tracking changes.

The sample tables illustrate the nature of attribute fields that can be included. No intent should be inferred from the absence of a particular attribute. Where justified by their potential usefulness, tables and fields have been included that offer benefits in implementing the design. For example, we included look-up tables for such defined domain variables as designator and direction. Not all tables need to be utilized; many are included here to illustrate the multimodal flexibility of the proposal. Field names have been selected for their mnemonic value, but are not otherwise critical.

### Transportation Feature Table

The Transportation Feature Table contains the data needed to describe each feature in the transportation network. There will be one record for each physical facility on the base map. The table uses tranfeatID, plus the date and time the record was created, as the primary key to identify each record. The descriptive data include the beginning and ending milepoints, a standard name, a separate external key (usually that of the data source), and the direction of travel. The design assumes that all included transportation features will be of the linear type. Nonlinear features may be referenced to adjacent linear features. For example, an airport terminal may be tied to a point on the accessing roadway.

The table below provides additional details regarding each included data item. In this example, the jurisdiction domain consists of counties in a single state, so a county line represents a forced end to each linear feature.

| Data Item | Meaning | External Key |
|---|---|---|
| entrydate | the date that the record was created | |
| entrytime | the time that the record was created | |
| enteredby | the user identification (ID) of the person creating the record | |
| tranfeatID | the unique numeric identifier for a transportation feature | |
| tranfeattypeID | the unique identifier for a transportation feature type | Yes |
| stateID | the unique identifier for the record containing the USPS state code | Yes |
| countyID | the unique identifier for the record containing the FIPS county code | Yes |
| designatorID | the unique identifier of the type of agency defining the feature | Yes |
| primaryID | the unique primary route identifier created by the designator (all related features will carry this same primary identifier) | Yes |

| Data Item | Meaning | External Key |
|---|---|---|
| secondaryID | the unique secondary route identifier created by the designator (realignments, ramps, and service roads will carry a secondary identifier other than "000", which indicates the existing mainline) | Yes |
| beginmilelog | the milelog measure for the feature's origin | |
| endmilelog | the milelog measure for the feature's terminus | |
| directionID | the unique identifier of a direction code | |

The fields tranfeattypeID, stateID, countyID, designatorID, primaryID, and secondaryID would be combined to create a single public key for accessing the data without knowledge of the internal key (tranfeatID). This public key conforms to our previously published naming convention proposal (Butler and Dueker 1998). Other forms for this key are supported by the table's design.

The public key fields could be combined into a single field, but the grouping of selections, such as selecting all the roads in a given county, would be more difficult due to the need to parse the combined value. External attribute tables, such as the Event Table in Figure 2, utilize the internal primary key of the Transportation Feature Table as the necessary foreign key connection. To be placed in geographic order, records could be sorted by either the internal or external key and beginmilelog.

It is important to note that the actual key field values are not stored here. They are kept in look-up tables, the population of which defines the domain for each public key component. The unique record identifier for the applicable public key component value is stored in the Transportation Feature Table. This design concept allows domain value meanings to change over time without requiring changes in the Transportation Feature Table. For example, if the decision is made at some point to utilize the two-digit FIPS state code instead of the initial USPS abbreviation, only the State Code Table needs to be updated, assuming that the identifier for each state (stateID) remains unchanged. It would be a simple matter to concatenate and expand the stored identifiers to produce a single public key field in a data extract for exchange with another user.

The public key fields do not form the internal primary key of the table, which is simply tranfeattypeID plus the time stamp fields (entrydate and entrytime). The time-stamp fields provide the option to store changing transportation feature descriptions over time. The concept could be extended to include two similar fields that stored the valid time period for the record (i.e., from a beginning date to an ending date). These values are independent from the time stamp shown, which is intended to store the time that the record was last updated.

## Transportation Feature Type Table

The type of transportation feature being described in a Transportation Feature Table record is determined by the value of tranfeattypeID. The Transportation Feature Type Table stores the domain for this field and serves as a mechanism for generating a pick list or pull-down menu of choices for data entry and reporting. To support this function, a listorder field is supplied that allows a database administrator to sort records in this table for presentation to the user. This field is alphanumeric to allow alphabetic sorting. This approach can be used to enter a value of "03h" to place an entry between "03" and "04," so that inserted options do not require renumbering of all subsequently listed choices. The field can also be used to change the ordering of values over time (e.g., to make the more frequently used values appear at the top of the list).

| Data Item | Meaning |
|---|---|
| tranfeattypeID | the unique numeric identifier for a tranfeattype |
| tranfeattype | the code value representing the transportation feature type (e.g., RD or AR) |
| tranfeattypeTXT | the longer description of type (e.g., road or airport runway) |
| listorder | the alphanumeric indicator of list display order |

Only the pick lists relevant to the concepts expressed by this illustrative design—that a single database can be constructed for a multimodal agency's GIS-T infrastructure—are described in this article. Look-up tables implied by the use of such field names as ststatusID and dirprefixID in the Street Table are not included, as the examples listed are sufficient to express the concept.

## Anchor Point Table

Anchor points define the locations on the roadway system that can be readily identified on maps as an aid to conflation, the process of combining two databases (maps) to create a new one. Anchor points also provide the means to define roadway segment locations using a two- or three-dimension coordinate system, such as a State Plane Coordinate System, as a supplement to the linear LRS measures that may be resident in an inventory database.

Anchor point locations are stored in this table as a text description, such as "the intersection of Maple Street and Broad Avenue." The physical location of this intersection is defined by one or more reference objects and the bearing and distance from the object(s) to the anchor point. The location of each anchor point in the linear LRS of roadways is specific to the roadway, as defined by one or more anchor sections. Thus, the linear LRS location of an anchor point is part of the Anchor Section Table.

| Data Item | Meaning |
|---|---|
| anchorpointID | the unique numeric identifier for an anchor point |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |

| | |
|---|---|
| enteredby | the user ID of the person creating the record |
| descTXT | the description of the anchor point location |

Fields could be easily included to store other items of interest such as anchor point type (intersection, bridge end, county line, etc.) and the date that the anchor point was established.

### Anchor Section Table

An anchor section begins and ends at anchor points. Each anchor section provides a highly accurate field measure (length) for the roadway segment. The number of anchor points and anchor sections are determined by the level of field measurement accuracy required. Anchor section records provide the linear LRS milelog measure for each terminating anchor point, the distance along the measured route between them, and the date and by whom the length measurement was made. One could extend the list of anchor section attributes to include how the measure was made and similar descriptive items.

| Data Item | Meaning |
|---|---|
| anchorsectionID | the unique numeric identifier for an anchor section |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| roadwayID | the unique numerical identifier for a roadway |
| beginanchorpt | the identifier of the anchor point at the section origin |
| beginmilelog | the milelog measure for the anchor section origin |
| endanchorpt | the identifier of the anchor point at the section terminus |
| endmilelog | the milelog measure for the anchor section terminus |
| length | the anchor section length, in units of 0.001 mile |
| measuredate | the date that the length measure was taken |
| measuredby | the name of the person taking the length measure |

### Reference Object Table

One or more reference objects may be used to locate each anchor section. The initial base map is not expected to contain reference objects, but the data table needs to be created as part of the original database so that appropriate records may be added over time. Reference objects may be of various types, such as survey monuments or iron pins. A given reference object may be defined in more than one datum or measurement system.

| Data Item | Meaning |
|---|---|
| refobjectID | the unique numeric identifier for a reference object |
| datumID | the datum in which the reference object location is defined |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| refobjtypeID | the unique numeric identifier for the reference object type |
| descTXT | the text description of the reference object and its location |
| xcoord | the *x*-plane coordinate measure for the reference object location |
| ycoord | the *y*-plane coordinate measure for the reference object location |
| zcoord | the *z*-plane (height) coordinate measure for the reference object location |
| measuredate | the date that the location measure was taken |
| measuredby | the name of the person taking the location measure |

### Reference Object Type Table

Reference objects may be of several types, each representing a kind of real-world object that may be used as a reference object (e.g., iron pins, concrete monuments, and traffic signal poles). The listorder variable is alphanumeric and can be used to sort items on a pull-down pick list.

| Data Item | Meaning |
|---|---|
| refobjtypeID | the unique numeric identifier for a reference object type |
| refobjtype | the code value representing the reference object type (e.g., LP) |
| refobjtypeTXT | the type code meaning (e.g., light pole) |
| listorder | the alphanumeric indicator of the list display order |

### Reference Object/Anchor Point Table

A single reference object may be used to locate multiple anchor points, or one anchor point may be positioned using multiple reference objects. The resulting many-to-many relationship requires a resolution or associative table. To reflect expected field survey methods, this table also stores the bearing and distance from the reference object to the related anchor point, instead of *x,y* measures.

| Data Item | Meaning |
|---|---|
| anchorpointID | the unique numeric identifier for an anchor section |
| refobjectID | the unique numeric identifier for a reference object |

| | |
|---|---|
| apdistance | the distance in units of 0.001 foot to the anchor point |
| apbearing | the compass bearing to the anchor point using the reference object as the pivot point and true north as the zero point of rotation (see Figure 1) |

## Event Table

Events are the characteristics and subordinate features of a linear transportation feature. Characteristics include such things as the functional class, number of lanes, pavement type, speed limit, and shoulder type. Subordinate features include intersections, bridges, signs, and guardrails. The example Event Table design accommodates temporal references with a date and time stamp. It accommodates divided highway attribution through a side field; a similar approach could be used to handle data by lane. Potential values for side are left, right, and both. The data model shows anchor points and anchor sections being described as events. Reference objects could be stored as events using the lateraloffset and side fields to show a distance and direction from the transportation feature to the object. Point and linear events have a beginning milelog measure, while only linear events have an ending milelog measure. The eventID field is proposed to be unique only within the transportation feature; although it could be universally unique across the entire database, this is not required. The value of tranfeatID is a partial foreign key to connect the event to its own transportation feature. In application, software could search for the most recent Transportation Feature Table record or one consistent with the time stamp in the Event Table record. As with the Transportation Feature Table, additional fields could be incorporated into the design to store the period of time for which the record was applicable.

| Data Item | Meaning |
|---|---|
| tranfeatID | the unique numeric identifier for a transportation feature |
| eventID | the unique numeric identifier for an event |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| eventtypeID | the identifier for an event type |
| beginmilelog | the milelog measure for the event origin |
| endmilelog | the milelog measure for the event terminus; [null] for point events |
| side | the side-of-road indicator (left, right, or both) |
| lateraloffset | the distance from a logical centerline to an object located off the roadway |
| value | the value of the specified attribute |

## Event Type Table

Events may be of several types (e.g., functional class, number of lanes, pavement type, speed limit, shoulder type, intersection, bridge, sign, or guardrail). In the example database design, each event type has its own parameters for units of measure and number of decimal places. The number of decimal places is needed because the value field in the Event Table is alphanumeric; therefore, it cannot be stored in a numeric format that includes decimal information. This means that a data extraction process will need to use the numdecimal value for the event type to create the proper appearance. An alternative is to store an explicit decimal point in the value field.

| Data Item | Meaning |
|---|---|
| eventtypeID | the unique numeric identifier for an event type |
| eventtype | the code value representing the event type (e.g., post-mounted sign) |
| eventtypeTXT | the type code meaning (e.g., post-mounted sign stored as MUTCD type) |
| unitofmeasure | the units of measure for the event type |
| numdecimal | the number of decimal places in the field, if numeric |
| listorder | the alphanumeric indicator of list display order |

## State Table

The U.S. Postal Service has established a set of state abbreviations that we have endorsed as a geographic naming standard for the transportation feature public key. This look-up table stores the relevant values. An alternative would be to use the FIPS state codes, as proposed by the current NSDI initiative.

| Data Item | Meaning |
|---|---|
| stateID | the unique numeric identifier for a state |
| stateabb | the abbreviation for the state (e.g., TN) |
| stateTXT | the name of the state (e.g., Tennessee) |
| listorder | the alphanumeric indicator of the list display order |

## County Table

The FIPS catalog includes a standard on numeric county designations for use in computer applications. This look-up table stores the relevant values. A multi-state agency might include the state component of the FIPS county code. States have a two-digit designation, and the county designations are three digits in length.

| Data Item | Meaning |
|---|---|
| countyID | the unique numeric identifier for a FIPS county code |
| countyFIPS | the code value representing the county (e.g., 065 or 47065) |

| | |
|---|---|
| countyTXT | the name of the county (e.g., Hamilton County, TN) |
| listorder | the alphanumeric indicator of the list display order |

## Direction Table

The Direction Table stores the domain values for the direction of linear measurement on transportation features. Linear LRS measures generally follow the federal standard of direction (i.e., west to east and south to north). Other approaches may be used, including bi-directional methods for divided highways.

| Data Item | Meaning |
|---|---|
| directionID | the unique numeric identifier for a compass direction of travel |
| direction | the code value representing the direction (e.g., 3 or SB) |
| directionTXT | the full name of the direction (e.g., southbound) |
| listorder | the alphanumeric indicator of the list display order |

## Designator Table

The naming convention proposed by us in an earlier article and the subsequently proposed NSDI transportation feature identification process include a component for the designating agency. Using such a component in the public key precludes the need for a universal agent that ensures unique values for identifiers. (There is, of course, the issue of making sure that designator codes are unique, a subject beyond the scope of this article.) With the designating agency component, these agencies may act independently to ensure that the transportation feature identifiers are globally unique. Standards can be developed for the construction of an agency designation code such that they are automatically unique, such as by using their federal tax identifier.

| Data Item | Meaning |
|---|---|
| designatorID | the unique numeric identifier for a designator of transportation feature identifiers |
| designator | the code value representing the designator (e.g., 074) |
| designatorTXT | the name of the designator (e.g., Hamilton County TN Highway Department) |
| listorder | the alphanumeric indicator of the list display order |

## Airport Runway Table

The database design in Figure 2 includes five transportation feature types. These tables extend the database to include information retained by individual modal offices in a multimodal agency. In some cases, such as that of the Airport Runway Table, no changes can be made in the legacy modal table to point back to the corresponding Transportation Feature Table record's identifier. Instead, one would be required to use runwayID in the Airport Runway Table as a foreign key to select the correct record in the Transportation Feature Table using extkeyID and tranfeattypeID.

| Data Item | Meaning |
|---|---|
| runwayID | the unique numeric identifier for an airport runway |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| airportID | the identifier for an airport |
| runway | the compass direction pair for the runway name (e.g., 090-270) |
| descTXT | the long description of the facility (e.g., main east-west runway) |
| pavetypeID | the identifier for the runway pavement type |

## Rail Line Table

The Rail Line Table stores information about railroad tracks, potentially including light rail and other public transit forms. Many railroads use linear LRS measures, so this table could be expanded to include them. Another option would be to include terminating facilities for each rail segment, such as a switch or block signal.

| Data Item | Meaning |
|---|---|
| railID | the unique numeric identifier for a railroad track |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| rrcoID | the identifier for a railroad company |
| raillinebranch | the railroad service area (e.g., Chesapeake Subdivision) |
| descTXT | the long description of the facility (e.g., Mainline from Signal 342 to Switch 9477) |

Although this article does not fully explore the application, the concepts could be equally applied to public transit on shared and dedicated roadways. For example, bus routes could be expressed as linear events and/or as transportation features in their own right. Route segment travel time would be a useful attribute.

## Waterway Table

Waterways are transportation facilities and impediments to surface travel; they may be present in a multimodal transportation database for either reason. The Waterway Table could serve as a look-up table for selecting a feature crossed by a structure in a bridge inventory.

| Data Item | Meaning |
|---|---|
| waterwayID | the unique numeric identifier for a waterway or water body |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| watername | the name of a waterway or water body (e.g., Tennessee River) |
| waterabb | the short name to put on maps (e.g., Tennessee R.) |
| mainstreamID | the identifier for the waterway into which this feature flows (e.g., Ohio River) |
| descTXT | the long description of the facility (e.g., Tennessee River within Chickamauga Lake) |
| watertypeID | the identifier for the water type |

### Street Table

Unlike the Airport Runway Table, the Street Table has been designed to include tranfeatID as a direct foreign key to point to the related Transportation Feature Table record. We have taken a decidedly local government approach in showing a Street Table, using address numbers as a linear LRM. The four street name component fields (dirprefixID, name, stypesuffixID, and dirsuffixID) follow the standard endorsed by the National Emergency Number Association for E-911 messages. The dirprefixID, dirsuffixID, and adddirectionID fields could point to a single directionID field in the Direction Table. A subordinate Block Table could also be included.

| Data Item | Meaning |
|---|---|
| streetID | the unique numeric identifier for a named street |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| tranfeatID | the identifier for the transportation feature |
| minaddress | the minimum address number for the street |
| maxaddress | the maximum address number for the street |
| ststatusID | the identifier for the street status (to indicate open, closed, planned, etc.) |
| dirprefixID | the identifier of the street name direction prefix, if any |
| name | the root street name (could include street type prefixes common in names derived from foreign languages) |
| typesuffixID | the identifier of the applicable street type |
| dirsuffixID | the identifier of the direction suffix, if any |
| adddirectionID | the identifier of the direction of increasing address numbers |

### Intersection Table

We propose that intersections have a separate table. This design approach provides a simple means for storing information regarding intersections, such as traffic control and turn restrictions, in a normalized database design. A normalized design precludes the need for redundant storage of intersection information for all intersecting street segments. The design also provides a “poor man’s” topology for pathfinding. Using the intersection type field, the meaning of intersection can be extended to include any type of junction, from limited-access highway interchanges to bridges and railroad grade crossings. The intersection identifier is proposed to be universally unique. Since an intersection is also an event on the two or more involved linear transportation features, the various many-to-many relationships of intersections and transportation features would be stored in the Transportation Feature/Event/Intersection Table, a resolution or associative table.

| Data Item | Meaning |
|---|---|
| intersectionID | the unique numeric identifier for an intersection |
| entrydate | the date that the record was created |
| entrytime | the time that the record was created |
| enteredby | the user ID of the person creating the record |
| inttypeID | the identifier for the intersection type |

### Transportation Feature/Event/Intersection Table

This table is an associative or resolution table that stores the many-to-many relationship between intersections and the events that tie them to transportation features. The table presented here consists only of foreign keys; however, they can include attributes. We have chosen to include the complete key of the Event Table but only a partial key for Intersection and Transportation Feature table records. This choice means that query software will need to pick the Transportation Feature and Intersection table records that best matches the timing of the Event Table record. An obvious alternative is to include the complete primary key of all three tables.

| Data Item | Meaning |
|---|---|
| intersectionID | the identifier for an intersection |
| tranfeatID | the identifier for a transportation feature |
| eventID | the identifier for an event |
| entrydate | the date event that the record was created |
| entrytime | the time event that the record was created |

## Conclusion

This article demonstrates an explicit path to meet the primary requirements for a multimodal transportation database that can serve the needs of an enterprise GIS-T infrastructure. Related guidance can be derived from articles noted in the list of references. Our intent is to show that it is possible to construct such an infrastructure regardless of the current state of information systems. Whether it is GIS- or computer-aided design-based, any successful enterprise solution will need to adopt an "all of the above" approach that includes all existing and potential applications. The key is not in the cartography, programming language or software platform currently used by various offices, but in the design of the database that connects them to each other.

The proposed Enterprise GIS-T Data Model has been shown to include the various data models now used by the many functional units of a multimodal transportation agency (Butler and Dueker 2000). For the sake of simplicity, only the elements of the model that serves linear transportation facilities are included here. Previous articles by the authors describe a more robust model that includes the area and point features, including nontransportation facilities.

The Enterprise GIS-T Data Model supports isolated data exchanges of component pieces, such as cartography or feature characteristics, by unbundling the elements of the presently dominant integrated topological vector data model. Relatively simple and inexpensive software currently exists to meet the majority of transportation agency needs. The Enterprise GIS-T Data Model isolates the data from software technology by allowing data-centric systems to be constructed in all currently available product lines that support dynamic segmentation. This article demonstrates one way in which the model can be implemented.

## About the Authors

**J. Allison (Al) Butler** is Director of GIS for Hamilton County, Tennessee, where he leads a growing community GIS enterprise. Previously, Butler was with the Florida DOT, and he serves as a part-time consultant to other state and local transportation agencies. He is certified by the AICP and the ASPRS, and earned a B.B.A. from the University of Georgia (1974). The author may be contacted at Hamilton County, Tennessee, 117 E. 7th St., Suite 300, Chattanooga, TN 37402. alb@exch.hamiltontn.gov.

**Kenneth J. Dueker** is the Director of the Transportation Studies Center at Portland State University, where he previously served as the Director of the Center for Urban Studies. Dr. Dueker holds a B.S. from the University of Washington (1960), an M.S. from the University of Washington (1963), and a Ph.D. from the University of Washington (1967). His research interests include GIS applications and transportation planning. The author may be reached at the Center for Urban Studies at Portland State University, Portland, OR 97207-0751. Duekerk@aol.com.

## References

Butler, J.A. and K. Dueker, 2000, *A Primer on GIS-T Databases.* Portland, OR: Center for Urban Studies, Portland State University. http://www.upa.pdx.edu/CUS/PUBS/contents.html

Butler, J.A. and K. Dueker, 1998, *A Proposed Method of Transportation Feature Identification.* Portland, OR: Center for Urban Studies, Portland State University. http://www.upa.pdx.edu/CUS/PUBS/contents.html

Dueker, K. and J.A. Butler, 1999, *A Framework for GIS-T Data Sharing.* Portland, OR: Center for Urban Studies, Portland State University. http://www.upa.pdx.edu/CUS/PUBS/contents.html

Dueker, K. and J.A. Butler, 1998, GIS-T Enterprise Data Model with Suggested Implementation Choices. *Journal of the Urban and Regional Information Systems Association,* 10 (1), 12-36.

Federal Geographic Data Committee, Ground Transportation Subcommittee, 2000, *NSDI Framework Transportation Identification Standard—Draft Number 3*. Washington, DC: US DOT Bureau of Transportation Statistics. http://www.bts.gov/gis/fgdc/web_intr.html

Opiela, K.S.,1997 *Research Results Digest No. 218: A Generic Data Model for Linear Referencing Systems.* Washington, DC: Transportation Research Board.

Vonderohe, A., and T. Hepworth, 1998, A Methodology for Design of Measurement Systems for Linear Referencing. *Journal of the Urban and Regional Information Systems Association*, 10 (1), 48-56.