The DLG-E Model Overview

***The DLG-E Model Overview** document describes the Digital Line Graph - Enhanced model and the U.S. Geological Survey/National Mapping Division's implementation of it. This document is intended as a summary of the DLG-E model and should not be regarded as all-encompassing. Furthermore, the DLG-E model, NMD's implementation of the model, and the model content specifications are continuing to evolve.*

12/6/93

1. The DLG-E Data Model

The Digital Line Graph - Enhanced (DLG-E) data model is a model for two-dimensional spatially referenced topologically structured vector data with features, attributes, and relationships. It represents a feature-based perspective of the topologically structured vector-based data model. The first section gives an overview of the model, with the following sections going into more detail.

1.1 Overview

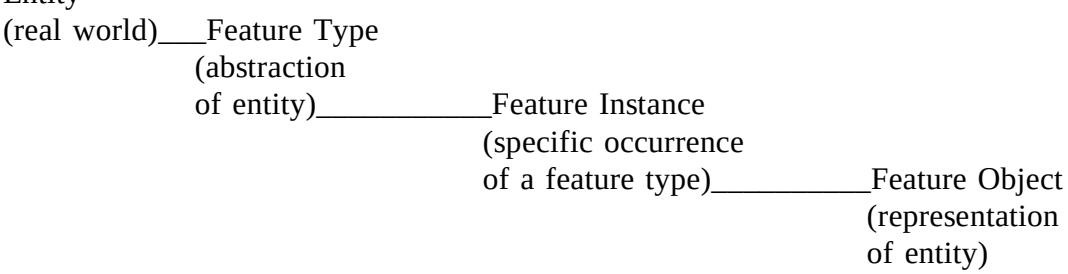
Entities, feature types, feature instances, feature objects, and spatial objects are the fundamental elements of the DLG-E model.

An entity is a real world phenomenon. A bridge is an entity that exists in the real world, and even has parts such as a superstructure and a deck. A lake is an entity, and its shoreline also is an entity. Measurements of the earths' surface are entities, such as contour lines and survey points.

A feature type describes a class of entities defined by common attributes and relationships. It is an abstraction of a set of entity instances that emphasizes characteristics to model in a data representation. Stream, road, spring, and bridge are examples of feature types.

A feature instance is an occurrence of a feature type. Maramec Spring is an instance of the feature type "spring." The Golden Gate Bridge is an instance of the feature type "bridge."

A feature object is the representation of a feature instance. Streams are entities that exist in the real world, generally symbolized on maps with blue lines. The blue lines are the objects, and a specific stream (entity) symbolized by a blue line (object) is a feature instance. The figure below graphically illustrates the relationships between entities, feature types, feature instances, and feature objects.



Whether we recognize it or not, we apply these concepts daily. Consider the situations where an individual studies a map, then points to a specific blue line and says, "That's Mill Creek." The map user has identified the feature instance "Mill Creek" by associating the blue line

(object) with the actual stream (entity). Furthermore, the label, the location of the blue line relative to other objects (geographic relationships), and unique twists and turns in the object (all properties of the entity represented by the object) are clues to the map user that the selected feature instance is Mill Creek.

To represent geographic reality the DLG-E model uses the elemental object types: spatial objects, feature objects, attribute objects, and relationship objects. Both the locational and nonlocational characteristics of the entities are represented by the objects. Locational aspects describe an object's geographic position (using x,y,z coordinates for example) or a geometric characteristic (such as its accuracy). The nonlocational characteristics of an entity include concepts such as its name, hydrographic category, or operational status.

There are four spatial object types of node, point, chain, and polygon. These form the basis for a topologically structured vector based data model. The spatial objects provide locational information when used by feature objects.

The three feature objects (basic, compound, and non-spatial) represent the locational and nonlocational aspects of entities. The locational is represented by using spatial objects and attribute objects. The nonlocational aspects are represented by using attribute objects.

Attribute objects describe both locational and nonlocational characteristics. The locational aspects include things like reliability of position, and frequency of positional change. The nonlocational aspects include things like name, construction material, operational status, and relationships to other feature objects.

Each feature type, i.e. abstraction of an entity, is assigned a list of attributes that represent the characteristics of the entity in its digital representation. For example, feature type "bridge" has attribute types of "name", "cover status", and "deck status." Each attribute type has a set of allowable attribute values, i.e. domain. The domain for attribute type "deck status" is "double decked" and "not decked." A feature instance must have a value for each attribute assigned to its feature type. The attribute values for the Golden Gate Bridge would be name of "Golden Gate Bridge", cover status of "not covered", and deck status of "not decked."

Relationships are the links between objects. They describe not only the topological relationships of the elements of the graph, but also the relationships between feature objects and spatial objects and between feature objects and other feature objects. "Chain is bounded by a node" is an example of a topological relationship between spatial object types. "Feature object is composed of spatial object(s)" is an example of a relationship between feature objects and spatial objects. "Feature object is above feature object" and, "feature object bounds feature object" are examples of feature object to feature object relationships.

This overview of the DLG-E model establishes the conceptual view of a feature-based

perspective of geographic reality, and introduces the components of the model. The next sections further describe the components of the DLG-E model.

1.2 Spatial Objects

The topologically structured vector-based model forms the basis for the spatial aspect of the DLG-E model. The DLG-E model defines a set of 0- (node, point), 1- (chain), and 2-dimensional (polygon) objects. (Section 1.7 contains the definitions of the spatial objects.) The spatial objects portray locational information and participate in topological relationships. For example, a chain object has a string of coordinates to define its shape. A node has a single coordinate to define its position. Every chain is bounded by a starting node and an ending node.

The spatial objects are participants in an aggregate called a "surface." A surface is a collection of spatial objects that form a planar graph and its complementary polygons and isolated points. In graph theory terms, the planar graph component of a surface is a set of planar multi-digraphs. It is a "set" because there can be several (unconnected) graphs; "multi" because there can be loops and multiple chains connecting the same pair of nodes; and, "di" because all chains are directed.

1.3 Feature Objects

A feature object is an element used to represent a feature instance. There are three types of feature objects: basic, compound, and non-spatial. Basic Feature Objects (BFO) are at the basic level that cannot be decomposed into other feature objects. BFOs are composed of like dimension spatial objects. A 0-D BFO is composed of a point and(or) a node; a 1-D BFO is composed of chain(s); and, a 2-D BFO is composed of polygon(s). Compound Feature Objects (CFO) are composed of other feature objects (basic and(or) compound.) Non-spatial feature objects are not composed of spatial objects.

For example, the feature type "stream/river" is represented as a Basic Feature Object of either 1- or 2-dimensions. Each feature instance of a stream is atomic, i.e. not composed of other streams. The feature type "watercourse" is represented as a Compound Feature Object, which can be composed of streams, lakes, canals, and/or estuaries. The feature type "connector" portrays a connection between feature instances without a locational aspect. A "connector" is represented as a non-spatial feature object.

1.4 Dataset-level Objects

Other than spatial objects and feature objects, the DLG-E model contains constructs to form sets of these objects. These dataset-level objects are dataset, theme, and surface.

A dataset is a collection of DLG-E data within a geographic area. It is the complete set of feature objects, spatial objects, attributes, relationships, and dataset-level objects that constitute the digital representation of the geographic domain.

A theme is an association of feature objects. Feature objects are grouped into themes based on feature type or the value of an attribute. All instances of feature type "stream" belong to theme "hydrography." Any instances of feature type "well" with attribute value of "water" or "heat" for attribute "product" belong to the hydrography theme. A feature object belongs to zero, one, or more themes.

A surface is a collection of spatial objects that form a planar graph and its complementary polygons and isolated points over a specified geographic domain, typically corresponding to the domain of the dataset.

A dataset has one or more surfaces and one or more themes. A theme occurs on a single surface. All basic feature objects belonging to a theme have their spatial objects belonging to the same surface. Multiple themes may share the same surface. A surface belongs to a single dataset and contains one or more themes.

In the DLG-E model, the dataset-level objects carry the metadata elements as attributes. The metadata element "source origin" is an attribute of a theme, "vertical datum" is an attribute of a dataset, and "completion date" is an attribute of a surface.

1.5 Relationships

Spatial objects, feature objects and dataset-level objects can participate in relationships. Currently, all relationships involve two participants. Relationships are classified by the types of the participants. The categories of relationships in DLG-E are topological, feature, and dataset-level. For each category of relationships the following table lists participants and examples.

Category	Pattern	Example
Topological	Spatial RELN Spatial	Chain is-bounded-by Node
Feature	Feature RELN Spatial	Stream/River is-composed-of Chain
	Feature RELN Feature	Lake/Pond is-bounded-by Shoreline
Dataset	Dataset RELN Dataset	Surface contains Theme
	Dataset RELN Feature	Dataset has Feature
	Dataset RELN Spatial	Surface contains Spatial

1.6 Attributes

Attributes are the characteristics to model in the digital representation. Attributes describe both locational and nonlocational characteristics. The locational aspects include things like reliability of position, and frequency of positional change. The nonlocational aspects include things like name, construction material, operational status, and relationships to other feature objects.

All DLG-E model components can have locational and non-locational attributes (except for spatial objects which can not have non-locational.) Feature objects, dataset-level objects, relationships, and attribute values can have attributes. For example, feature type "lake/pond" has an attribute "elevation." The dataset-level object, dataset, has an attribute "state coverage", theme has "source origin", and surface has "processing software system." The relationship "flows to" could have an attribute "flow rate" (for illustration only; currently no relationships have attributes.) The attribute values for "elevation" can have attribute "stage" to further describe the elevation value. The attribute value "water" can have attribute "water characteristics" to further describe the properties of the water.

Every attribute has a domain which defines the set of valid values for the attribute. The domain can be defined or it can be enumerated. For example, the attribute "number of lanes" has a defined domain of integer values from 1 to 9, inclusive. The attribute "name" has a defined domain of any alphanumeric string. The domain for attribute "operational status" is an enumerated domain of "abandoned", "operational", "under construction", and "reclaimed."

Each feature type (e.g. lake/pond, stream/river, etc.), relationship type (e.g. flows to, is above, etc.), attribute value (e.g. water, rock, etc.) and dataset-level object has zero or more attributes assigned. For example, the feature type "lake/pond" has assigned attributes of "name", "impoundment status", "water characteristics", "elevation", and "hydrographic category." The relationship type "flows to" has no assigned attributes. The (integer) attribute values for "elevation" when applied to feature type "lake/pond" has assigned attributes of "stage" and "source." The attribute value "water" when a value of attribute "product" as it applies to feature type "well" has attribute "water characteristics." The theme object has assigned attributes of "processing organization", "source origin", and others.

Many elements share common attributes like "name" or "operational status." However, the context in which an attribute is used often restricts a domain, especially if enumerated. For example, both "strip mine" and "road" have the same attribute "operational status." In the context of a "strip mine" value "reclaimed" is allowed, but for "road" it is not.

All assigned attributes must have a value for every instance. There are a set of null attribute values that help satisfy this requirement. The value "not applicable" means that for an

instance there is no relevant value for a particular attribute. The value "unspecified" means a value is possible but not known, and it is not required. The value "unknown" means a value is not known, but it is required. (The value "unknown" will normally be resolved prior to the release of the data.)

The domain of values for an attribute may use multiple data types. For example, the attribute "width" can be an integer or a null value. The attribute values for "elevation" can be integer or floating point depending on the feature type to which it applies.

Depending on context, attributes can have multiple values. For example, the attribute "product" in the context of feature "mine" can have multiple values that all apply to a single feature instance.

To summarize the attribute model of DLG-E, there is a set of attributes with domains. Feature objects, dataset-level objects, relationships, and attribute values can have attributes. Each model element is assigned zero or more attributes. Each attribute's domain may be restricted depending on the context. Assigned attributes must have a value for every instance. (The defined set of "nulls" can be used.) In certain contexts, attributes can have multiple values that apply to a single instance.

* * * * * *

The DLG-E model is independent of the content-level details, i.e. the set of feature types, the set of attribute types, and the set of relationship types. The description of the DLG-E model in the previous sections included many examples. However, the examples are not part of the model. The concept of "feature type" is part of the model, but "lake/pond" is not part of the model--it is part of the content. The concept of "relationship type" is part of the model, but the specific type of relationship "flows to" is not. "Flows to" is part of the definition of the content.

1.7 Summary of DLG-E Model Terms

Attribute - A characteristic of an entity or model component (i.e, name, location, status, physical appearance).

Attribute Type - the label for a set of attributes which may be common to a "type" of thing; also includes a definition of the characteristics being captured.

Basic Feature Object [BFO] - A feature object at the basic level that cannot be decomposed into other feature objects. A 0-D BFO is composed of a point and/or node; a 1-D BFO is composed of chain(s); and, a 2-D BFO is composed of polygon(s).

Chain - A one-dimensional spatial object that is a directed connection between two nodes (not necessarily distinct) with geometric location specified by a sequence of two or more coordinate pairs (or triplets.)

Compound Feature Object [CFO] - A feature object composed of other feature objects (basic and(or) compound.)

Dataset - Digital/spatial information pertaining to a geographic area of any size or shape.

Dataset-level Object - A term referring to the dataset, theme, and surface model components.

Entity - A real world phenomenon, like a bridge, stream, lake, shoreline, contour, or survey point.

Feature - A real world phenomena and its digital representation [Entity + Feature Object]

Feature Type - A class of entities defined by common attributes and relationships.

Feature Object - A construct used to represent a feature instance and its attributes and relationships. (see also Basic Feature Object, Compound Feature Object, and Non-spatial Feature Object.)

Feature Instance - An occurrence of a feature defined by unique characteristics.

Node - A zero-dimensional spatial object that is the junction of two or more chains or is an end point of a chain.

Non-spatial Feature Object - A feature object that has no spatial representation, i.e. it is not composed of spatial objects.

Object - The digital representation for a feature instance. The DLG-E model uses seven types: four spatial objects (node, point, chain, polygon) and three feature objects (basic, compound, non-spatial).

Point - A zero-dimensional spatial object that is not along the path of any chain.

Polygon - A two-dimensional spatial object that consists of the area interior to a closed circuit of chains excluding holes. A polygon may have a representative coordinate associated with it.

Relationship - A link between spatial, feature, or dataset-level objects.

Relationship Type - the label for a set of relationships which have similar meanings and share a set of common attribute types.

Spatial Object - An object that portrays locational information and is used to represent the topology of a surface.

Surface - A collection of spatial objects that form a planar graph and its complementary polygons and isolated points over a specified geographic domain.

Theme - An association of feature objects based on feature type or the value of an attribute of a feature.

2. Implementation of DLG-E

This section will discuss how National Mapping Division (NMD) of the USGS is approaching the implementation of the DLG-E model. The product based on this model will be the next generation of NMD's topological-vector-based digital product.

2.1 The Production System

The DLG-E Production System is being developed to implement the DLG-E model. One of its main objectives is for the collection software to be content-independent. The addition of a new feature type or an attribute should not cause a change in the software.

This is not at all obvious from the user-interface perspective. It is implemented using X and the menus are filled with content-specific choices. So where do the content-specifics come from? There is a DLG-E Content Standards Database (DSDB) that provides the information to the collection software (see 2.2). Although the content is not "hardcoded" into the collection software, the software has access to the DSDB through a Standards Interchange File. The content of this file is used by the collection software to assist the cartographic technician in the data collection process.

Other objectives of the implementation are to make the software portable between UNIX hardware platforms and the data portable between other hardware platforms. The DLG-E Portable Data File (DPDF) is an ASCII file set that was designed to move the internal data structures of the processed data between unlike hardware platforms.

2.2 Content Standards Database

The DLG-E Content Standards Database (DSDB) is the repository of NMD's content specification for the DLG-E product. The objective in developing the DSDB is to support all aspects of DLG-E data collection. This includes generation of data collection instructions (i.e. NMD's "templates"), supplying content specifics to the collection software, providing a basis for on-line help, supporting interactive training tutorials to help employees learn the content, and other related activities.

The Content Standards Database contains (or will eventually contain) the following types of information:

- set of feature types
- set of attribute types
- set of theme types
- set of relationship types
- which attributes apply to which feature types

- which features types can belong to which theme type
- representation of feature instances using feature objects
- attribute domains and units
- attributes on attribute values (i.e. functional dependencies)
- definitions of object types
- definitions of enumerated values
- and more...

The Content Standards Database design is still under development. And the content with which the database is populated is still evolving. However, there were sufficient portions of the DSDB in place to support a prototype release of a production system (Dec 1992.) The content-population reached a sufficient level to perform data collection for a prototype DLG-E product in 1993.

2.3 DSDB Content Examples

This section contains some examples of features, attributes, relationships and representation rules to illustrate the contents of the DSDB.

Consider a feature type of "lake/pond." It is described by attributes "hydrographic category", "impoundment status", "elevation", "name", and "water characteristics." The domains for the attributes when applied to feature type "lake/pond" are in the following table.

Feature Type: Lake/Pond		a standing body of water with a predominantly natural shoreline surrounded by land.					
Attribute	Definition	Domain	MV	Attribute	Definition	Domain	MV
Hydrographic Category	portion of the year the feature contains water	Intermittent Perennial Unknown	N	--	--	--	--
Impoundment Status	whether an area is artificially impounded or natural	Impounded Unspecified	N	--	--	--	--
Elevation	the vertical distance from a given datum	(integer)	Y	**Stage**	height of water surface	Average Water Elevation Date of Photography High Water Elevation Normal Pool Spillway Elevation	N
				Source	derivation of the value	Interpolated Non-interpolated	N
		unspecified		--	--	--	--
Name	proper name, specific term, or expression	(alphanumeric) Unspecified	Y	--	--	--	--
Water Characteristic	distinctive properties of the water	Salt Unspecified	N	--	--	--	--

For the feature type "lake/pond", the table lists the assigned set of attributes, their domains, and whether single or multi-valued. Both "elevation" and "name" can have multiple values, with the others having only single values. The "lake/pond" also has attributes on one of its attribute values. When the attribute value for "elevation" is numeric, it has attributes "stage"

and "source" when applied to feature type "lake/pond."

In the NMD content, each feature type has rules for which feature objects may be used to represent it and within which relationships it can participate. A "lake/pond" is represented as either a 0-D BFO or a 2-D BFO. A 2-D BFO "lake/pond" can participate in a bounded-by relationship with a 1-D BFO "shoreline." 0-D or a 2-D BFO "lake/pond" can participate in network relationships (flows to, flows from, and connects to). As another example, "stream/river" is represented as either a 1-D BFO or a 2-D BFO. A 1-D BFO "stream/river" can participate in network relationships (flows to, flows from, connects to), but not in boundary relationships (bounds, bounded by) or vertical separations (above, below). As a 2-D BFO, it can only participate in network relationships and the boundary relationship "bounded by."

NMD content requires every feature object to belong to at least one theme. Feature types "lake/pond" and "stream/river" have their instances associated with the "hydrography" theme. Some instances may belong to more than one theme. Consider a "lake/pond" instance that is in both the "hydrography" theme and the "land use/land cover" theme. Some feature instances belong to a theme depending on the value of an attribute. For example, an instance of feature type "well" with value "water" or "heat" for attribute "product" belongs to the" hydrography" theme, otherwise it belongs to the "built-up" theme.

The DLG-E model establishes the framework within which the NMD content is defined. The distinction between content and model and implementation can be difficult to discern since all three are still evolving.

2.4 DLG-E Development Continues

The NMD is continuing to define its own specific content with regards to the DLG-E model. Currently, NMD content does not require the full flexibility of the DLG-E model. For example, relationships are allowed to have attributes, but NMD content does not use this option. The DLG-E compound feature object can be composed of other CFOs, but NMD content restricts a CFO to be composed of only basic feature objects.

This document captures a snapshot of the development work relating to DLG-E. Development continues in all aspects. The DLG-E model, NMD's implementation of it, and NMD's content specification for a DLG-E product are all still evolving.