

ATTRIBUTE DEFINITIONS AND CODES

I. PURPOSE GIS attributes contain information associated with geographic features on a map. GIS users, present and future, will most likely need to develop attribute tables for new layers of map information. Existing attribute tables will need to be expanded as well. The Vermont Center for Geographic Information, Inc. (VCGI) encourages Vermont users of geographic information to use the same base attributes and codes so that data can be more easily exchanged and understood by all users.

If you wish to develop a GIS data layer from scratch, please contact VCGI to see what data already exists and for general guidance.

If you are developing new attributes for geographic data, please:

- contact VCGI to see if standard attributes have already been defined,
- follow the guidelines below for developing new attributes,
- submit your attributes and codes to VCGI so that others may use them and for general review.

Individual users can expand on existing attributes as needed. If new attributes are defined and used, please inform so that attributes and codes remain consistent.

II. DEFINITIONS See the VCGI Glossary publication for a full set of GIS definitions.

Data Layer: A separate map layer with a specific theme (e.g., public wells), generally corresponding to an ARC/INFO "coverage".

Coverage: An ARC/INFO or other GIS data layer.

Map Feature: For these guidelines map features are represented as points, lines and areas. Areas are also called polygons. Examples are public wells (points), roads (lines) and wetlands (areas). Each individual feature (or element) corresponds to one *record* in a database.

Attribute: Information associated with each feature in a data layer. For example, depth (feet) is an attribute of public wells. An attribute corresponds to one *field* in a database record.

III. GENERAL Several factors should be considered when defining attributes and

GUIDELINES

codes:

The Data Layer: The general purpose and feature types of the data layer should be well defined. What are the potential applications of the data layer? Has the same data layer been used elsewhere? Contact VCGI or your VGIS Regional Service Center for help if you are just starting out.

Unique ID: Each geographic feature in a database generally has a unique identifier that singles it out from all other features of the same type. For example, each public well would have its own number. The unique identifier is usually the first attribute defined. It may be defined in part by the feature's location, such as by drainage basin. The unique ID should generally be different from the <cover>-ID, described next.

<cover>-ID: ARC/INFO automatically generates a <cover>-ID for each data layer. The <cover>-ID must be unique for each record of information; if not, other attribute values will be lost when **BUILD**ing or **CLEAN**ing the coverage. If you are not using ARC/INFO directly, you do not need to be concerned with the <cover>-ID.

Field width: Each attribute has a width, or number of characters (or bytes) for each entity (or record). For large data sets the attributes should be kept small. In any case, attributes should only be as wide as is needed for the largest possible value. For example, the longest county name in Vermont requires 10 characters, so a width of 10 would be used.

Description: Attributes must be explicitly defined to avoid confusion by other users.

Units: The units must always be given for measured values (e.g., feet or meters), either in the field name or in its description.

Codes: Codes can be used in place of full words or concepts. When possible use a mnemonic code, such as one or two letters which stand for something. Dropping vowels is a common practice. Codes can save considerable storage space in the database.

Possible Values: Many attributes have a specific set of allowable values, such as "1 to 99", or "three characters made up of a capital letter followed by two digits". The range of acceptable values should be explicitly given.

Complexity: Coding schemes should be easily understood by all users,

even when more complex schemes would be computationally more efficient.

IV. DATA LAYER DESCRIPTION

The data layer should be described as fully as needed to develop the attributes (an example is given in Section VI). For GIS layers the description should at least include:

Coverage Name: The computer (or GIS) name for the data layer, limited to 8 characters by DOS. Letters, digits, underscores and dashes may be used; no spaces.

Data Layer: The data layer (or coverage) theme should be clearly described. Be sure to include any exceptional entities which should or should not be included.

Feature Type(s): The type of geographic feature should be specified as point, line or area. Areas are also referred to as polygons. If a coverage has two types of features (points and lines, or lines and polygons), the attributes for each feature type should be in separate lists.

Data Source(s): The source of the attribute data should be clear, including the date of the information. Different attributes may have different sources, so the attribute source may need to be included with individual attribute definitions. Include a bibliographic reference to a document and the scale of the source data if appropriate.

V. ATTRIBUTE DEFINITIONS

The definition for each attribute should include the following information. An example is given in Section VI.

Item or Field Name: The actual name used by the computer, with no spaces. Underscores can be used, but dashes are not compatible with dBase names. The display of attribute values in tables may be improved if short names are used for narrow attributes.

Generic names like "ID" and "CODE" should not be used because of potential conflicts between names when data layers are overlaid. In particular, item names "AREA", "PERIMETER" and "LENGTH" should be avoided as these names are generated by ARC/INFO.

Attribute: Several words describing the attribute.

Attribute Source: Used when different attributes come from different sources.

Description: Describes in detail (1) the contents of the attribute, (2) the possible values for the attribute, (3) all codes and their meanings (listed or footnoted), and (4) any notable exceptions. Upper, lower or mixed case should be specified for character fields. Often the description can be quite simple, but it should be complete.

Type: The type of data to be stored. The type can be used for both INFO and xBase, unless otherwise noted. Several special purpose INFO types are not listed here.

- C = Character (includes digits and punctuation)
- I = Integer (positive, negative or zero), INFO only; may be represented as type N with no decimal places, as in xBase
- N = Real Number (decimal points are part of the width), xBase only.
- B = Binary format integer, INFO only. The width is the number of bytes (generally 2 or 4).
- F = Binary floating point real number (INFO only). The width is the number of bytes (generally 4 or 8).
- L = Logical (True or False), xBase only. Logical items in INFO should be 1 character items, set to 'Y' or 'N'.
- M = Memo field, for variable lengths of text, xBase only. Useful for notes and descriptions. In INFO, memo-type fields can be handled by having a separate documentation textfile.
- D = Date

Width: The number of characters or bytes used by the attribute. See "Type". Note that the decimal point for type N must be included in the width.

Output Width: The number of spaces used for output. This is only required for INFO types B and F.

Decimal Places: The number of decimal places for types N and F, if used.

Example: An example of a typical value, if needed.

VI. EXAMPLE

SAMPLE ATTRIBUTE DEFINITIONS FOR WATERFALLS, CASCADES AND GORGES

Data Layer: Waterfalls, Cascades and Gorges from ANR publication (see Data Source)
 Coverage Name: WATCASGO
 Feature Type: Point locations
 Data Source: All attribute information is contained in "Waterfalls, Cascades and Gorges of Vermont", 1988, VT Agency of Natural Resources (unless otherwise noted).

<u>Attribute Identifier</u>	<u>Item Name</u>	<u>Description</u>	<u>Type</u>	<u>Width</u>	<u>Decimal Places</u>	<u>Example</u>
	WCGID	basin#report# (basin = 1 to 17)	I	5	0	17999
Type	WCGTYPE	W, C or G (upper case)	C	1		W
Name	SITENAME	Site name (upper/lower)	C	25		Lana Falls
Length (ft)	WCGLENGTH	Horizontal	I	4	0	9999
Width (ft)	WIDTH	Horizontal	I	3	0	999
Total drop (ft)	TOTDROP	Vertical	I	3	0	999
Slope (degrees)	SLOPE	Avg, degrees	I	2	0	99
Wall height (ft)	WALLHT	Vertical	I	3	0	999
Rare plants	RAREPLANT	Y or N	C	1		Y
Flow regulated	FLOWREG	Y or N	C	1		N
Privacy	PRIVACY	see Codes	I	1	0	5
Tidiness	TIDY	see Codes	I	1	0	3
Swimmability	SWIM	see Codes	I	1	0	4

Codes

PRIVACY: 1 = not secluded or wild
 2 = somewhat secluded
 3 = wild
 4 = very secluded
 5 = private

TIDY: 1 = clean
 2 = fairly clean
 3 = a mess

SWIM: 1 = good bathing
 2 = fair swimming
 3 = good swimming
 4 = great swimming