



Economic and Social Council

Distr.: General

3 July 2007

Original: English

Ninth United Nations Conference on the Standardization of Geographical Names

New York, 21 - 30 August 2007

Item 12 (c) of the provisional agenda*

Toponymic data files:

Toponymic data transfer standards and formats

Proposed U.S. National Standard for identifying attributes of named physical and cultural geographic features

Submitted by the United States of America**

* E/CONF.98/1.

** Prepared by Dwight Hughes (USA).

PROPOSED U.S. NATIONAL STANDARD FOR IDENTIFYING ATTRIBUTES OF NAMED PHYSICAL AND CULTURAL GEOGRAPHIC FEATURES

Dwight Hughes (USA)

The U.S. Geological Survey (USGS) Geographic Names Project and the U.S. Bureau of Census submitted to the American National Standards Institute (ANSI) a proposal to establish a national standard addressing minimum identifying attributes for selected named geographic features. As defined and stored in the USGS Geographic Names Information System (GNIS) database, those attributes are: 1) the unique, permanent, feature record identifier, 2) the official feature name, and 3) the official feature location. The institute accepted the proposal, and on June 15, 2007, a draft standard was submitted for formal review. It is expected that the standard will be approved before the end of calendar year 2007.

The purpose of the standard is to facilitate accurate and consistent representation of geographic feature data for the nation as defined by data owners and authoritative sources, and to promote efficient management and exchange of such data among numerous critical, but overlapping and potentially conflicting data sets. The standard compiles with the mission and statutory requirements of the United States Board on Geographic Names. The standard supports the primary mechanisms for storage, management, display, and exchange of geographic feature data, particularly among geographic information systems, relational databases, and Internet-based applications and services.

The Geographic Names Information System database is maintained by the U.S. Geological Survey with thirty years of data to store and disseminate the official geographic feature names and locations recognized by the Board for federal use. The GNIS contains information about physical and cultural geographic features of all types in the United States, associated areas, and Antarctica, current and historical, but not including roads and highways. The database holds the federally recognized name of each feature and defines the feature location by State, county, USGS topographic map, and geographic coordinates.

The database assigns a unique, permanent feature record identifier, the Feature ID as the standard key for accessing, integrating, or reconciling feature data across multiple data sets. Other attributes that are not a part of this standard include names or spellings other than the official name, feature designations, feature classification, historical and descriptive information.

Additions and corrections to the feature data in the GNIS are accepted for consideration from any source, and upon validation, are entered into the database. The GNIS collects data from a broad program of partnerships with federal, state, and local government agencies and other authorized contributors. The latest information technology tools and

methodologies are applied to ensure that local data are properly represented in the GNIS and through the GNIS to all federal agencies and to the public.

The explosive growth of geographic information systems and internet-based applications and services have generated multitudes of related, overlapping, and potentially contradictory datasets containing geographic feature data, served by multiple, interlocking, and interdependent applications at all levels of government and the private sector.

It is virtually impossible to correlate the many different sets of feature data containing thousands, tens of thousands, or more records based solely on attribute comparisons with any assurance that they mutually address the same features, record for record. Unregulated text fields, particularly feature names, may contain wide variations of spelling, content, presentation, and errors that are outside the control of data owners. Locational attributes are subject to multiple formats, variations and differing interpretations. Geospatial representations frequently exist in multiple versions at varying scales and levels of accuracy. Many geographic features have no definable, official, recognized, or agreed upon boundaries, for example, the eighty percent of communities in the United States that are unincorporated, or a mountain for which there is no objective feature extent.

Therefore, there is no known mechanism for ensuring at the national level that related datasets contain and are serving mutually consistent feature data or that state and local feature data are accurately represented in federal datasets. For over a hundred years, the USGS National Base Map series provided the only consistent, national geospatial representation available, but for many critical purposes, printed maps have been rendered obsolete by technology.

The implications of incorrect, inaccurate, or contradictory feature data appearing simultaneously in multiple Internet applications are serious and potentially catastrophic with regard to national security, emergency preparedness and response, regional and local planning, site selection and analysis, cartographic applications, environmental problem-solving, tourism, and all levels of communication. The proposed standard is intended to help correct this deficiency.

The standard is intended to be inclusive and to serve the widest possible scope of features within the interests of all parties. It does not attempt to define or to place limitations or requirements on geographic feature types, relationships, data ownership, permanence, scale, or other feature attributes or display characteristics except as stated therein.

The single categorical exception to the standard are roads and highways, defined as linear features for the movement of motor vehicles. In addition, the standard generally does not apply to:

- Brand name commercial facilities (unless a landmark).
- Unnamed features locatable only by address or other locative attribute.

- Small infrastructure and utility elements, e.g., utility poles, junction boxes, pumping stations, mile markers.
- Large, unnamed network infrastructure and utility elements, e.g., pipelines, transmission lines.
- Mobile and/or transitory events such as forest fires, hurricanes, or other weather or geologic occurrences.
- Subsurface geologic zones or mineral deposits.
- Postal delivery zones, which are managed by the Postal Service solely for the purpose of delivering the mail.

The above exclusions guidelines are subject to revision. The final authority concerning applicability of any particular feature or feature set to the standard rests with the U.S. Board on Geographic Names or as delegated to the staff of the Geographic Names Project in coordination with the Census Bureau. Considerations for acceptance of proposed data are: 1) utility to and demand from user communities, 2) applicability to the overall purpose and mission of the standard and of the Board on Geographic Names, 3) quality of the data with respect to the validation rules of the Board on Geographic Names, 4) ability of the prospective source to be an effective data partner for data maintenance, and 5) GNIS system physical and personnel resources.

For the purposes of the standard only, the following definitions apply:

geographic feature

geographic entity

A geographic feature is a physical or defined entity on the landscape or seascape, except a road or highway, that requires identification, location, and attribution for the information purposes of government and the public, and having as minimum attributes a feature name and a feature location. Geographic features are characterized and differentiated solely by function, and not by size, composition, structure, extent, relationship to other features, ownership, age, or any other factor. These and other attributes will be associated with the feature in various datasets but do not necessarily or individually differentiate it from similar features. A school and a church operating in the same building, for example, constitute two features and are assigned separate feature identifiers. This definition makes no reference to graphic or digital representations of the feature, although those representations would be attributes of the feature.

historical geographic feature

An historical geographic feature is one that once existed but has ceased to exist and/or is no longer apparent or defined on the landscape. The definition makes no reference to age, current or historical use, condition, appearance, level of human habitation or any other factor. If the feature existed when entered into the Geographic Names Information System database but subsequently ceased to exist, the status of the feature is revised to historical but not otherwise changed. No valid feature record is ever removed from the database unless determined to be a duplicate of another record as a result of entry error. Reference to historical features not in the database and with sufficient documentary evidence are entered into the database with the status of historical.

feature identifier

feature ID

record identifier

The feature identifier is a permanent unique number assigned to a geographic feature for the sole purpose of uniquely identifying that feature as a record in any information system database, dataset, file, or document and for distinguishing it from all other feature records so identified. The number is assigned sequentially (highest existing number plus one) to new records as they are created in the Geographic Names Information System. The number, by design, carries no information or association to the content of the feature record and therefore is not subject to change as attribute values change. Once assigned to a feature, the number is never changed or withdrawn, and never reassigned. The Feature ID can be applied in conjunction with system-unique record identifiers in any database or system, thus providing a national standard common reference identifier across multiple datasets. The Feature ID is stored in the GNIS database as an integer with a maximum of ten digits, and currently is seven digits.

official feature name

The official feature name is a feature name or designation in any language expressible in the Roman alphabet for which the written form and its location are approved by the U.S. Board on Geographic Names by decision or by policy delegation to an authoritative source. The name complies with the Board's Principles, Policies, and Procedures for Domestic Geographic Feature Names and with the Names Data Users Guide, which addresses display and format guidelines regarding spelling, capitalization, diacritical marks, special characters, abbreviations, etc. A geographic feature has one and only one official name. Names are stored in the database as alphanumeric text in AL32/UTF8 format.

variant name

A variant name is a current or historical name or spelling for a geographic entity other than its official name determined by an authoritative source to be useful for reference to, identification of, or location of a geographic feature, and which is supported by documentary or physical evidence or references. The name complies with the Board's Principles, Policies, and Procedures for Domestic Geographic Feature Names and with the Names Data Users Guide.

official feature location

primary point

The official feature location is a single point to which the official feature name is associated in order to ensure positive and unique identification and association, also referred to as the primary point. The location is determined by the authoritative source and is approved or recognized as official by the U.S. Board on Geographic Names by decision or by policy delegation to the authoritative source. The placing of the location point is governed by policies of the Board as defined in the Names Data Users Guide and GNIS metadata, generally at or near the geographic center, but there are exceptions for certain classes of features. For example, the official feature location of flowing bodies of

water (streams, rivers, creeks, etc.) or trending linear features (valleys, gulchs, gullys, hollows, etc.) is at the mouth.

A geographic feature may have only one official location regardless of size, extent, composition, structure, or boundaries. The location point is coincident with but in addition to and independent of any other geospatial representation or boundary definition that may be attached to the feature in other datasets. Linear and areal features also may have secondary points as defined in the Names Data Users Guide and GNIS metadata. Locations and secondary points are stored in the Geographic Names Information System as latitude and longitude in decimal degrees to seven places, North American Datum 1983 (NAD83), and are available in geospatial format as needed.

primary state (or state equivalent) and county (or county equivalent)

The primary state and county or their equivalents are those containing the primary point. Linear and areal features may exist in other, secondary states and counties, which also are carried in the GNIS as defined in the Names Data Users Guide and GNIS metadata.

Agencies, departments, or organizations at any level of government in the United States may apply to the Geographic Names Project as an authoritative source for the purposes of entering feature data into the GNIS and assignment of feature identifiers. The Geographic Names Project does not attempt to adjudicate differences concerning jurisdiction, authority, responsibility, ownership, and/or stewardship among agencies for particular geographic areas, features, feature classes, or data sets. If such questions arise (there have been no issues of this kind in over thirty years), the question will be returned to the claimants for resolution among themselves or with a common superior. The final authority concerning applicability of any particular feature or feature set to this standard rests with the U.S. Board on Geographic Names or as delegated to the staff of the Geographic Names Project. Training and user documentation for the processes and automated data partnership tools are available.

The feature identifier (ID), the feature name, and feature location become official and a national standard for that feature within the meaning of the document upon submission by an authoritative source or decision of the Board, validation by the Geographic Names Project data specialists, entry of the record into the database, and assignment by the Geographic Names Information System of the ID. Other attributes are submitted, but are not part of this standard. Data shall comply with the validation rules and guidelines of the Names Data Users Guide.

It is the policy of the United States Government that only official geographic names are to be used on federal maps and in other federal publications or products. It is the policy of the Board that names and locations of most cultural, i.e., not natural, features are determined by the authoritative source and are not subject to formal review and decision by the Board. Locations are to be determined based on verifiable printed or digital documentary, graphic, or image sources or Global Positioning System (GPS). Geocoded locations are not considered sufficiently accurate, but may be accepted in the absence of other sources.

Revisions to feature names and locations, as well as other attributes in the Geographic Names Information System, are submitted by authoritative sources to the Geographic Names Project at any time through standard partnership agreements and multiple mechanisms as agreed. Data shall comply with the validation rules and guidelines of the Names Data Users Guide. Methods include secure web forms, batch files in most standard formats, and automated exchanges utilizing web feature services. Other procedures may be developed as technology advances. Changes are validated and committed by Geographic Names Project data specialists.

In addition to the draft ANSI standard, the GNIS Feature Identifier, official name, and official location are included in the U.S. Federal Geographic Data Committee draft Address Standard, and as top level optional attributes in the U.S. Department of Homeland Security Geospatial Data Model. The GNIS is a critical element in the National Spatial Data Infrastructure and the USGS *The National Map*. It is part of a national gazetteer development project, and the subject of continuing coordination with U.S. National Geospatial Intelligence Agency data collection programs.

The GNIS data are available to government, the public, and web applications through a web query site, web map, feature, and XML services, file download services, and customized files upon request. Other mechanisms may be developed in the future.

Feature names data must be widely, quickly, and reliably available to any system or citizen needing them. In addition, they must be kept current, complete, consistent, and accurate through distributed partnerships using efficient and reliable information technology tools. It is hoped that this standard will contribute to those purposes by:

- Ensuring feature attributes and attribute values from any source apply to the specified feature and to no other through reference to a common record identifier.
- Providing feature names that are complete and correct, standard, and nationally consistent in form and presentation.
- Providing feature locations that are reliable as national locational identifiers, independent of size, extent, or other spatial representations, and based on verifiable documentary, graphic, image, and/or GPS sources.
- Ensuring federal, state, county, local data are properly represented in the official federal database—the GNIS—and available to all.
- Mitigating against incorrect, inaccurate, contradictory feature data appearing simultaneously in multiple applications and layers.

For additional information on the GNIS and related web map and feature services, visit the Geographic Names web site <<http://geonames.usgs.gov>>. Communications concerning the Geographic Names Information System should be addressed to <gnis_manager@usgs.gov>.