United Nations
Group of Experts on
Geographical Names

Working Paper
No. 3

Twentieth Session
New York, 17 – 28 January 2000

Item 9
of the Provisional Agenda

MEETING OF THE WORKING GROUP ON TRAINING DATA FILES GAZETTEERS

Development of "Digital Map 25000 (Geographic Names and Public Facilities")

Submitted by the Japanese Government

Development of "Digital Map 25000(Geographic Names and Public Facilities)"

1. Background

A large number of geographic names are used in daily life in Japan and it is difficult to figure out the number exactly. The collection, arrangement and standardization of geographic names and information on public facilities will not only facilitate economic and social activities but also be of great use for environmental conservation, disaster prevention/disaster counter-measures and cultural heritage preservation.

The Geographical Survey Institute(GSI, the national mapping organization of Japan) initiated a project for producing a multi-purpose geographic name database. This database, coming in CD-ROM form, was envisaged to support various kinds of geographical information systems, manage documents concerning geographic names, and reproduce annotations on maps.

In planning the database, basic attribute information and accurate positional information were considered to be the basic essentials. The following are the details of the prerequisites of the geographic name database.

1. The database should cover the whole country.
2. The database should include a wide variety of geographic information such as administrative names, populated place names, natural feature names, etc.
3. The standards for selecting geographic names and for expressing their positions should be clearly defined.
4. The database should provide us with positions such as longitude and latitude.
5. Attribute retrieval should be simple.
6. Coding in the database should be systematic.
7. Geographic names should be also expressed with *Kana*(syllabary character) for reading.

After examining these matters, we concluded that the geographic name database should be based on the GSI's 1/25,000-scale topographic map series, the biggest scale of topographic maps which cover the whole Japan, and "Digital Map 25000(Geographic Names and Public Facilities)" was produced.

In order to standardize the geographic names on topographic maps and marine charts, collaboration was under taken between the GSI and the Hydrographic Department of the Japan Maritime Safety Agency through the "Liaison Committee on the Standardization of Geographic Names".

2. Outline

"Digital Map 25000(Geographic Names and Public Facilities)" comes with one CD-ROM. The data components are two fold, that is, geographic name data and public facility data. The former is basically a collection of geographic names with their geographic coordinates,

administrative divisions, writings and readings, etc. Data capturing was performed by digitizing 4,352 sheets of the GSI's 1/25,000-scale topographic maps. The latter is a collection of public facilities with their attribute information.

The total number of data included in the CD-ROM is about 570,000. Details are as follows.

Number of annotation data:	about 469,800		
Names of populated areas:		298,100	
Names of natural features:		83,700	
Mountains, hills and peaks:			16,800
Rivers, valleys, lakes and swamps:			51,600
Others:			15,300
Other annotations:		88,000	
(including the annotations for 5,100 public facilities)			
Number of symbols for public facilities:	about 100,700		

A distinctive feature of the geographic name data is that every annotation and public facility symbol is provided with a key point which represents the position of the object of the annotation or the symbol. On the other hand, a distinctive feature of the public facility data is that every public facility symbol is provided with, in addition to its key point, a proper name and an address of the object of the symbol.

3. Data Collection and Data Recording

1) Data Capture and Data Items

An automatic capturing system was used to acquire annotations and public facility symbols delineated on 1/25,000-scale topographic maps. Fortunately, 4,352 sheets of the 1/25,000-scale topographic map series had been already converted into raster format. As for the public facilities, the proper names and addresses were collected directly from the authorities concerned.

2) Positional Information and Positional Accuracy

The data include geographic coordinates of the key points of the geographic features which are expressed with annotations or public facility symbols. The data also include local coordinates of annotations and symbols on the map. Data capturing accuracy is within 0.2 mm at the mapping scale.

(1) Coordinates of annotated geographic features

The positions of annotated or symbolized geographic features are indicated with longitudinal and latitudinal coordinates of their key points.

(2) Coordinates of public facility symbols

In order to identify the positions of annotations and public facility symbols themselves on 1:25,000-scale maps, their positions are indicated with their local coordinates by the normalized coordinate system with its origin (0,0) at the left lower corner of the map neatline and its end (10000, 10000) at the upper right corner.

3) Data in the "Digital Map 2500(Spatial Data Infrastructure)"

Due to its mapping scale, the names of populated places are occasionally not shown in congested areas on 1/25,000 scale topographic maps. In order to fill up these names, about 15,000 annotation data were obtained from the existing "Digital Map 2,500(Spatial Data Infrastructure)", which was digitized from 1/2,500-scale basic maps of the urban areas. Every piece of annotation data is coded with the source map level for easy identification.

4) Data of "*Gaiji*"

The Japanese languge is written with *Kanji*(Chinese character of ideogram), and *Hiragana* and *Katakana* (Japanese syllabary character). *Kanji* has a large number of letters. The Japanese Industrial Standards (JIS) lists 6,355 *Kanji* letters (2,965 listed for the First Level and 3,390 for the Second Level) as the Standard *Kanji*. However, so-called *Gaiji* which are not listed in the JIS Standard *Kanji* are widely used, and the annotations on 1/25,000-scale topographic maps adopt as many as 235 *Gaiji*. To deal with *Gaiji*, we prepare a special font package.

4. File Structure

The file in the CD-ROM is composed of the following five tables.

1) Annotation Table

The structure of the annotation table is as follows.

Items		Data format	Contents
Coding of a 1/25,000-scale topographic map		Integer	Code of a topographic map to which an annotation belongs.
Name of a 1/25,000-scale topographic map		Text	Name of a topographic map to which an annotation belongs.
Standard Grid Square Code		Integer	Code of the grid to which the key point of an annotation belongs.
Administrative coding of the key point of an annotation		Integer	Code of the local administrative body to which the key point of an annotation belongs.
Annotation number		Integer	Unique number given to an annotation on a topographic map
Classification coding	First level coding	Integer	Code classified according to the characteristics of an annotation. Three levels of classification.
	Second level coding	Integer	
	Third level coding	Integer	
Annotation letter		Text	Letters used for an annotation on a topographic map
Reading of annotation letters		Text	Reading of an annotation in *Hiragana*
Blank		Text	Blank
Longitude of the key point of an annotation		Floating decimal point	Longitudinal coordinate of the key point of an annotation

Latitude of the key point of an annotation	Floating decimal point	Latitudinal coordinate of the key point of an annotation
ID of a public facility	Integer	ID in the public facility table given to an object annotated as a public facility
Flag for a historic spot, scenic spot and natural monument	Integer	Flag to distinguish a historic spot, scenic spot and natural monument

2) Annotation Coordinate Table

The structure of the annotation coordinate table is as follows.

Items	Data format	Contents
Coding of a 1/25,000-scale topographic map	Integer	Code of a topographic map to which an annotation belongs.
Annotation number	Integer	Unique number given to an annotation on a topographic map
Placement number	Integer	Serial numbers given to each split string of annotated letters
Placement of a letter string	Text	A string of letters identified by the placement number
X coordinate at the lower left of a letter string	Integer	X coordinate in the normalized coordinate system at the lower left corner of the first letter in a horizontal letter string, or the lower left corner of last letter in a vertical letter string
Y coordinate at the lower left of a letter string	Integer	Y coordinate in the normalized coordinate system at the lower left corner of the first letter in a horizontal letter string, or at the lower left corner of the last letter in a vertical letter string
X coordinate at the upper left of a letter string	Integer	Blank for a horizontal letter string, or X coordinate in the normalized coordinate system at the upper left corner of the first letter in a vertical letter string
Y coordinate at the upper left of a letter string	Integer	Blank for a horizontal letter string, or Y coordinate in the normalized coordinate system at the upper left corner of the first letter in a vertical letter string
X coordinate at the lower right of a letter string	Integer	X coordinate in the normalized coordinate system at the lower right corner of the last letter in a horizontal letter string, or blank for a vertical letter string
Y coordinate at the lower right of a letter string	Integer	X coordinate in the normalized coordinate system at the lower right corner of the last letter in a horizontal letter string, or blank for a vertical letter string

3) Table to Which an Annotation Belongs

The structure of the table to which an annotation belongs is as follows.

Items	Data format	Contents
Coding of a 1/25,000-scale topographic map	Integer	Code of a topographic map to which an annotation belongs
Annotation number	Integer	Unique number given to an annotation on a topographic map
Administrative coding to which an annotated object belongs	Integer	Codes of all administrative bodies on a map to which an annotated object belongs

4) Table of Public Facility Symbols

The structure of the table of the public facility symbols is as follows.

Items	Data format	Contents
Coding of a 1/25,000-scale topographic map	Integer	Code of topographic map to which an annotation belongs
Symbol number	Integer	Unique number given to a symbol on a topographic map
Standard Grid Square Code	Integer	Code of the grid to which the key point of an annotation belongs.
Administrative coding of the key point of a symbol	Integer	Code of the local administrative body to which the key point of a symbol belongs
Symbol coding	Integer	Code to identify the type of a symbol
X coordinate at the center of a symbol	Integer	X coordinate in the normalized coordinate system at center of a symbol
Y coordinate at the center of a symbol	Integer	Y coordinate in the normalized coordinate system at center of a symbol
ID of a public facility	Integer	ID of a public facility in the public facility table

5) Public Facility Table

The structure of the public facility table is as follows.

Items		Data format	Contents
ID of a public facility		Integer	Unique number given to a public facility
Administrative coding of the key point of a public facility		Integer	Code of the local administrative body to which the key point of an annotation or a symbol belongs
Classification coding	First level	Integer	Codes which classify public facilities. Second level code in the annotation table corresponds to the First level code in this table; and Third level code corresponds to the Second level code,respectivelly.
	Second level		
Name of a facility		Text	Proper name of a public facility
Location		Text	Address of a public facility; City/town/village name is dropped.
Standard Grid Square Code		Integer	Code of the grid to which the key point of a public facility belongs.
Longitude of a key point		Floating decimal point	Longitudinal coordinate of the key point of a public facility
Latitude of a key point		Floating	Latitudinal coordinate of the key point of

	decimal point	a public facility

5. Software for Simple Display

The CD-ROM comes with simple display software for data confirmation and retrieval.

6. Revision of the Data

The data are to be revised annually.