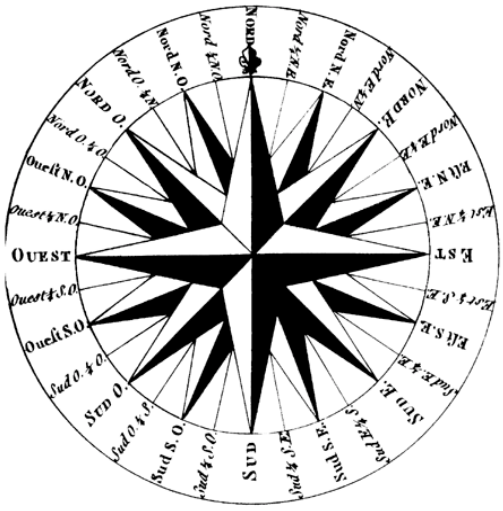
Getty Thesaurus of Geographic Names™

User's Guide to the TGN Data Releases

Release Version 2.0

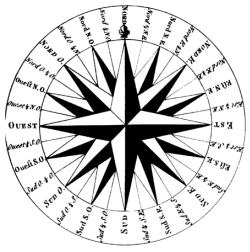


User's Guide to the TGN Data Releases

Release Version 2.0

Compiled and Edited by
Patricia Harpring, Managing Editor
Getty Vocabulary Program

Getty
Thesaurus
of Geographic
Names™



The Getty Vocabulary Program
1200 Getty Center Drive, Suite 1100
Los Angeles, California 90049-1680
http://www.getty.edu/research/tools/vocabulary/
02/12/2001

CONTENTS

PREFACE

About This Guide

This guide describes the release formats for the *Getty Thesaurus of Geographic Names* (TGN), one of the vocabularies produced by the Getty Vocabulary Program. It also provides a description of the history, structure, content, and application of the TGN. It does ***not*** give step-by-step instructions regarding ***how*** to construct a database or interface based on the data files; analysis and a competent programmer will be required of any user who wishes to implement the vocabulary data files. The Getty will not aid in this process.

About the TGN Release Formats

The TGN data releases are intended for users wishing to integrate the TGN data in a custom design that is compatible with their particular documentation program or other system. These files include all name forms (both preferred and variant) in the database, as well as coordinates, notes, dates, and place types.

Regular or extensive use of the Getty Thesaurus of Geographic Names (TGN) is subject to terms of our licenses. To learn about terms of the licenses, contact the Vocabulary Program at vocab@getty.edu. The TGN is available for licensing in three formats created especially for organizations developing their own data management systems. The data files are released annually. The sizes of the files vary depending upon the format, but they will be at least 500 megabytes. Customized versions of these files are ***not*** available.

The data releases use an 8-bit, code-extended ASCII character set. The TGN Relational Files are in relational database format. The TGN:REC files are ASCII records in a flat file format. The TGN:MARC data files contain USMARC authority records. To learn more about the datal go to http://www.getty.edu/research/tools/vocabulary/obtain.html or contact tgn@getty.edu.

The accuracy of the information contained in the database is not warranted in any way by the Getty, and the Getty assumes no responsibility or liability for the results of any use of the information.

.

ACKNOWLEDGMENTS

The development and dissemination of the TGN and the other Getty Vocabularies are possible due to the work of numerous people.

The TGN grows through contributions. Information in the TGN was compiled by the Getty Vocabulary Program in collaboration with many institutions, including the following projects:

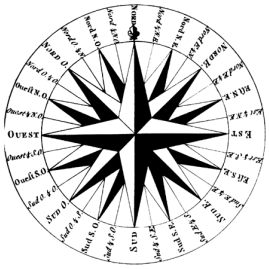
ARLIS: Art Libraries Society of North America
AVERY: Avery Index to Architectural Periodicals
BHA: Bibliography of the History of Art / Bibliographie d'Histoire de l'Art
FDA: Foundation for Documents of Architecture
GCI: Getty Conservation Institute
GRLPA: Getty Research Library Photo Archive
JPGM: The J. Paul Getty Museum
PROV: Getty Provenance Index

We thank the people who have worked to overcome various administrative, logistical, and legal obstacles to make the creation and release of the TGN possible: Murtha Baca, Kathryn Girard, Ken Hamma, Maureen Whalen, Jim Bower, Kathleen McDonnell, and Penney Cobey. We recognize the invaluable guidance of advisors: Henry Millon, John Onians, and Oreste Signore. We acknowledge the commitment, creativity, and superb craftsmanship of people who have provided technical solutions to the production of the TGN: Joan Cobb, Gregg Garcia, Tulio Hernández, April Brown, Joe Shubitowski, Sofie Dondoe, Marty Harris, Vivian Hay, James Lockwood, Ted Jung, and Stephen Toney.

I want to express my special appreciation to Vocabulary Program staff and temporary editors who have edited the database, including the following people: Alison Chipman, Robin Johnson, Jennifer Goodell, Christi Richardson, Michelle Kaufman, Shelley Triggs, Kim Burnham, Erin Coburn, Nathaniel Bleekman, Laïla Lalami, Jennifer Wood, Nanda Prato, Andrea Roth, Victoria Herd, Kezia Schulhof, Mali Griffen, Kharon Hathaway, Keshni Kashyap, Karen Tongson, and Christina Yamanaka. These editors have been enthusiastic researchers and indefatigable laborers who have spent long, tedious hours bent over maps, books, and computer monitors.

Patricia Harpring
Managing Editor
Getty Vocabulary Program
October 23, 2000

Getty
Thesaurus
of Geographic
Names™



CHAPTER 1
INTRODUCTION

OVERVIEW

The Getty Vocabulary Program, working closely with the Getty Standards Program, builds, maintains, and disseminates vocabulary tools for the visual arts and architecture. Development of the Getty Vocabularies began in the mid-1980s. The vocabularies produced by the Getty are the *Art & Architecture Thesaurus*® (AAT), the *Union List of Artist Names*® (ULAN), and *the Getty Thesaurus of Geographic Names*™ (TGN). The AAT is a controlled vocabulary of around 125,000 terms for describing and retrieving information on fine art, architecture, decorative art, and material culture. The ULAN is a database of around 220,000 names, biographical data, and bibliographical information about artists and architects, including a wealth of variant names, pseudonyms, and language variants. The TGN is a hierarchical database of around 1,000,000 vernacular, English, and historical names, coordinates, place types, and other relevant information about places.

ABOUT THE TGN

Contributions and Scope

The Getty vocabularies are compiled resources; they are not comprehensive. The vocabularies grow through contributions; contributors are typically Getty projects or other museums, libraries, archives, and bibliographic and documentation projects that catalog art objects, visual surrogates, or information about the visual arts and material culture. The scope of the AAT, ULAN, and TGN is limited to terminology important for the description of art, architecture, and material culture. At the same time, the Getty vocabularies do not provide all the types of terms necessary to describe art and architecture; terminology for iconography, events, and objects, people, and concepts that are unrelated to art are outside the scope of the Getty vocabularies. The vocabulary program is collecting the names of institutions that may be interested in contributing terms in electronic format to the vocabularies. If you are interested in becoming a contributor, please contact us. For AAT, send mail to aat@getty.edu; for ULAN send mail to ulan@getty.edu; for TGN, send mail to tgn@getty.edu.

Structured Vocabularies

The Getty vocabularies are “structured vocabularies.” Structured vocabularies are collections of terms organized in a way that specifies the relationships between terms and concepts for the purpose of facilitating access to key information. The Getty’s vocabulary tools cover geographic names, artist names, and concepts related to art and art history. The Getty vocabularies are used worldwide as standards for cataloging and describing art, architecture, and material culture. In local environments, the vocabularies can serve as access points or in “search assistants” to help users refine, expand, and enhance their searches in order to retrieve more meaningful results.

Structured vocabularies have traditionally been the province of catalogers and indexers, but in recent years vocabularies have played an increasingly important role in the retrieval of information from varied sources. More and more institutions and individuals are making material available electronically, resulting in a sea of data that is difficult to navigate without tools like vocabularies. These tools help information seekers obtain better results.

The development of vocabularies and common standards is an important component in protecting the long-term value of data, and to permit individuals and institutions to share information. For standards and resources to be useful, they must respect the disparate requirements of various intellectual disciplines and dispersed audiences and interest groups, including museums, libraries, archives, scholars, information specialists, and others who are concerned with art, architecture, and material culture. Therefore, the Getty seeks the advice and collaboration of other institutions and specialists in the various disciplines. The goal is to reach agreement so that the resulting standards may constitute a common voice that will speak forcefully for many cultural interests and be heard by policy-makers in the public and private sectors.

The Purpose of Vocabularies

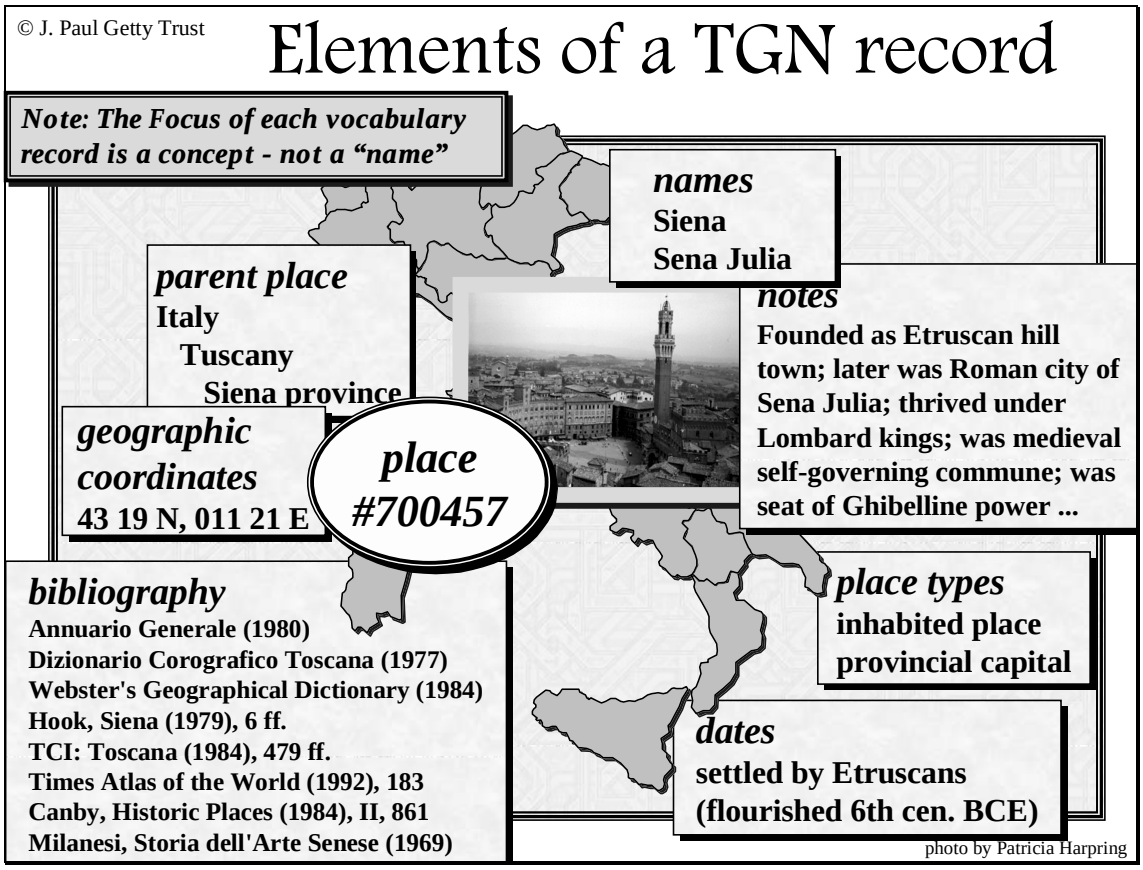
The Getty vocabularies and other structured vocabularies are needed to describe, organize and provide access to information about art, architecture, and material culture. Using a structured vocabulary enables researchers to find *le mot juste*, or the "exact" term for the scholarly description of objects or concepts, as well as the appropriate term for organizing a set of objects or concepts. Structured vocabularies also act as maps that guide people to information. As the number of potential access points to information increases exponentially, vocabularies may provide spelling variants, synonyms, and related terms that may allow access to a vast array of data.

Users of the Vocabularies

The Getty vocabularies are used by a wide variety of professionals, including art librarians, slide librarians, archivists, museum catalogers, scholars, researchers, students, and anyone who needs to describe and/or find information about art.

Ways To Use The Vocabularies

The Getty vocabularies can be used in three ways. They may be used at the data entry stage, by catalogers or indexers who are describing works of art, architecture, material culture, visual surrogates, or bibliographic materials. They may also be used as knowledge bases, providing interesting information to researchers. They may be used as search assistants to provide access in a local environment. Note that the AAT, ULAN, and TGN are copyrighted by the J. Paul Getty Trust, all rights reserved. Companies and institutions interested in regular or extensive use of the vocabularies should explore licensing options by reading about Licenses and Sample Data at http://www.getty.edu/research/tools/vocabulary or by contacting the Vocabulary Program at vocab@getty.edu.


© J. Paul Getty Trust
Elements of a TGN record
Note: The Focus of each vocabulary record is a concept - not a "name"
names
Siena
Sena Julia
parent place
Italy
Tuscany
Siena province
notes
Founded as Etruscan hill town; later was Roman city of Sena Julia; thrived under Lombard kings; was medieval self-governing commune; was seat of Ghibelline power ...
geographic coordinates
43 19 N, 011 21 E
place #700457
bibliography
Annuario Generale (1980)
Dizionario Corografico Toscana (1977)
Webster's Geographical Dictionary (1984)
Hook, Siena (1979), 6 ff.
TCI: Toscana (1984), 479 ff.
Times Atlas of the World (1992), 183
Canby, Historic Places (1984), II, 861
Milanesi, Storia dell'Arte Senese (1969)
place types
inhabited place
provincial capital
dates
settled by Etruscans
(flourished 6th cen. BCE)
photo by Patricia Harpring

CHAPTER 2
RELATIONAL FILES FORMAT

OVERVIEW

The data releases use an 8-bit, code-extended ASCII character set. See *Appendix A: Codes for Diacritics* used in the TGN. The database is at least 500 megabytes in size.

The TGN was compiled using a custom-built editorial system that allowed the Vocabulary Program 1) to merge records from separate contributors that represent the same place, 2) to build and rearrange hierarchies, and 3) to accept, edit, and manage contributions to the database.[1]

There are 14 relational tables.[2] The logical focus of the TGN is a place, information about which forms an intellectual record. This intellectual record comprises all information tied to the place through a unique numeric key (*geog_key*). Names, notes, coordinates, dates, and place types for a place are linked by the *geog_key*. Each name is also uniquely identified by a key (*placename_id*), and contributors, bibliography, and dates are linked to a name by that key. The *geog_key* of the immediate parent of a place is located in the *geog_parents* table, and it is with these keys that hierarchies may be constructed. A minimum record for a place includes one name (and the accompanying sort name and key words), one place type, and a link to an immediate parent.

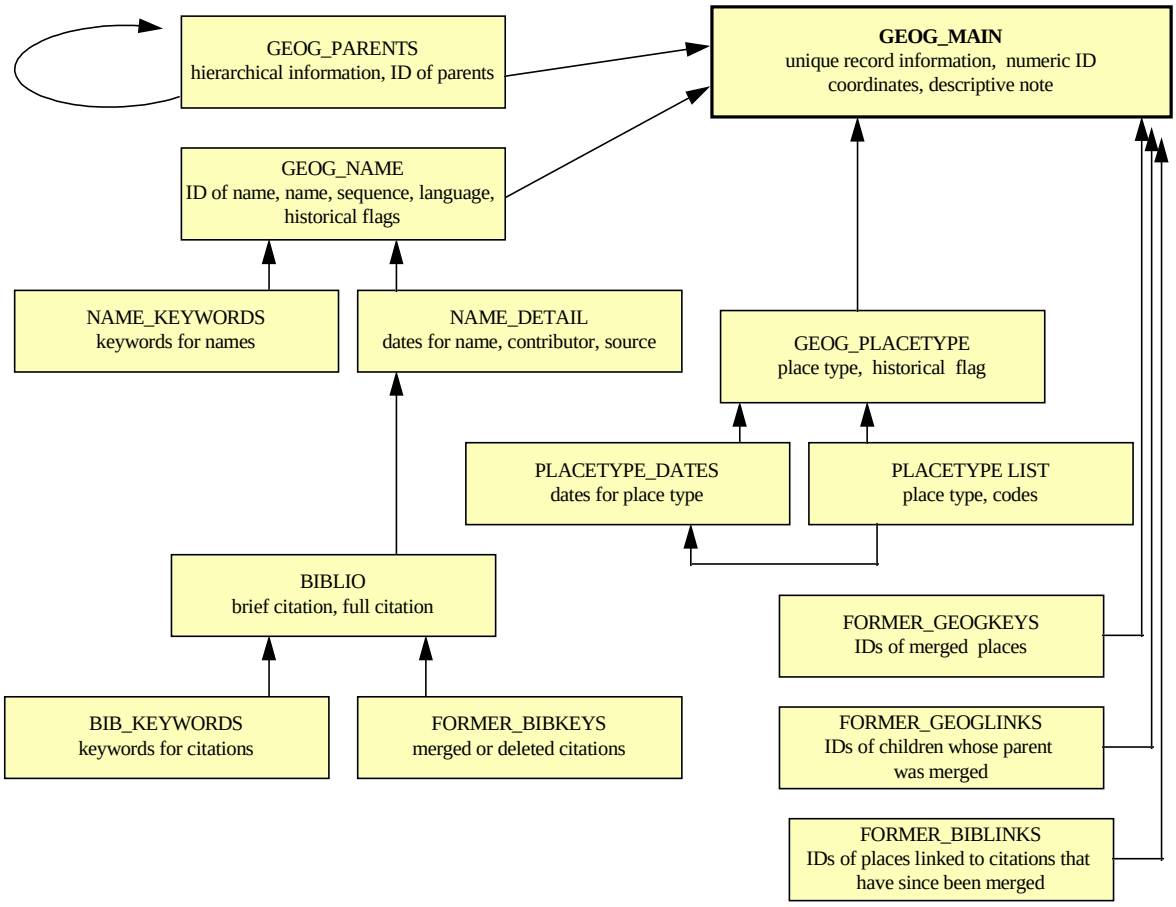
Four tables, *former_geogkeys, former_geoglinks, former_biblinks,* and *former_bibkeys,* store unique keys that have changed since the last release. Unique keys may change because of the nature of a compiled source like the TGN. For example, if there are two places in the database that represent the same place, when these separate records are merged into a single record, at least one of the unique keys must be lost.

Note that information in certain of the tables and columns is not intended to be displayed to the end-user. Such information includes normalized names intended for retrieval, and date fields intended to index display dates (which are notes).

[1] Contributions will be accepted from selected institutions in an automated format prescribed by the Vocabulary Program. The contribution format and a new Vocabulary Coordination System are scheduled to be in place by Spring, 2001. The original TGN editorial system operated under the Sybase Relational Database Management System and Sybase SQL Server release 4.9.2; it was developed using APT Edit release 4.0.2. TGN ran on Sun SPARC workstations running Sun OS release 4.3.4.

[2] Note that table names and data structure may change with the implementation of the new TGN editorial system and data model, slated for completion in 2001.

Entity Relationship Diagram for the TGN



RELATIONAL FILES DATA DICTIONARY

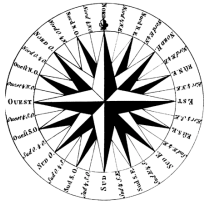
The TGN relational files format contains the following tables:

GEOG_MAIN	Main place record table, containing unique key for the place, coordinates, and descriptive note.
GEOG_PARENTS	Table containing keys for immediate parents of *geog_key*.
GEOG_NAME	Table for names of places, containing unique ID for the name, name string, normalized name, and various flags.
NAME_DETAIL	Table for repeating detail information about each name, including display date, contributor, and link to citation.
GEOG_PLACETYPE	Table for place types associated with the place, including code for place type, and various flags.
PLACETYPE_DATES	Table for repeating detail information about the place type, including display date, start and end dates.
BIBLIO	Table for citations for sources, including key for citation, brief citation, and full citation.
PLACETYPE_LIST	Table for place types, containing code for place type and place type term.
NAME_KEYWORDS	Table containing keywords for the names.
BIB_KEYWORDS	Table containing the keywords for full citations.
FORMER_GEOGKEYS	Table containing mapping of *geog_keys* that have been merged or deleted since last release.
FORMER_GEOGLINKS	Table containing former links to children of *geog_keys* that have been merged or deleted.
FORMER_BIBKEYS	Table containing mapping of *bib_keys* that have been merged or deleted since the last release.
FORMER_BIBLINKS	Table containing *geog_keys* that were linked to citations that have been merged or deleted.

In the following chart, the names of tables appear in small caps (**GEOG_MAIN**). The names of columns are in lower case (geog_key). For each column, the following is indicated: the data type, the length of the data, the allowance of nulls (i.e., whether or not data is required in that column), a description of the data, an example of the data, and an explanatory note, where necessary.

EXAMPLE

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
GEOG_MAIN				Main place record table, containing unique key for place, coordinates, and descriptive note.		
geog_key	int	4	0	Unique key for the place.	7008736	
descr_note	text	500	1	Note describing the history, physical characteristics, and significance of the place.	Founded by Etruscans; ruled by Romans 3rd cen. BC …	Note that no value in this ...



Getty Thesaurus of Geographic Names: *RELATIONAL FILES format* DATA DICTIONARY

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
GEOG_MAIN				Main place record table, containing unique key for place, coordinates, and descriptive note.		
geog_key	int	4 [3]	0	Unique key for the place.	7011179	
descr_note	text	500	1	Note describing the history, physical characteristics, and significance of the place.	Founded by Etruscans; ruled by Romans 3rd cen. BC & Lombards 6th cen.; constitution 1179; medieval seat of Ghibelline faction; rival of Florence; ruled by Charles of Anjou 13th cen., Visconti of Milan early 15th cen., & Medici of Florence after 1557.	Note that no value in this column is currently greater than 500, even though this is a text field.
lat	varchar	5	1	The degrees and minutes of the latitude, which is the angular distance north or south of the equator, measured in degrees and minutes along a meridian.	43 19	Format is two digits, one space, two digits.
lat_dir	char	1	1	Letter indicating if latitude is north or south of equator.	N	Legal values = N or S.
long	varchar	6	1	The degrees and minutes of the longitude, which is the angular distance east or west of	011 21	Format is three digits, one space, two digits.

[3] A 4-byte integer can hold a number as high as 2 billion.

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
				the Prime Meridian at Greenwich, England.		
long_dir	char	1	1	Letter indicating if latitude is east or west of the Prime Meridian.	E	Legal values = E or W.
lat_num	dec (5,3) or float		1	Latitude expressed as decimal fractions of degrees.	43.316	Whole degrees are represented by a two-digit decimal number ranging from 0 through 90, with any decimal fraction of a degree separated by a decimal point; latitudes south of the equator are designated by a minus sign.
long_num	dec (6,3) or float		1	Longitude expressed as decimal fractions of degrees.	011.350	Whole degrees are represented by a three-digit decimal number ranging from 0 through 180, with any decimal fraction of a degree separated by a decimal point; longitudes west of the Prime Meridian are designated by a minus sign.
GEOG_PARENTS				Table containing keys for immediate parents of *geog_key*.		Places may have mulitple parents.
geog_key	int	4	0	Unique key for the place.	7011179	
parent_key	int	4	0	Unique key for a parent of the place.	7003168	
pref_flag	varchar	1	0	Flag indicating if parent is P = Preferred, or N = Non-preferred.	P	Legal values = P or N.
GEOG_NAME				Table for names of places, containing unique ID for the name, name string, normalized name, and various flags.		
geog_key	int	4	0	Unique key for the place.	7011179	
placename_id	int	4	0	Unique key for this name (unique across entire database).	47413	All names have unique IDs, including homonyms.

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
seqn	int	1	0	Sequence position in which the name should be ordered in displays.	0	Values = 0-x.
place_name	varchar	122	0	Name string.	Siena	
sort_name	varchar	122	0	Normalized name; all caps with punctuation, diacritic codes, numbers, and spaces removed.	SIENA	To be used for sorting and retrieval.
lang	char	1	0	Flag indicating if name is V = Vernacular, or O = Other language.	V	Legal values = V or O. In restructuring of data for new system, this will be expanded to handle multiple languages.
age_flag	char	1	0	Flag indicating if the name is C = Current, H = Historical, or B = Both.	C	Legal values = C, H, or B.
pref_flag	char	1	0	Flag indicating if this is the P = Preferred name, or V = a Variant name.	P	Legal values = P or V.
other_flags	char	10	0	Flags indicating if name is ADJ = adjectival name form, ADJE = preferred English version of adjectival name form, ENG = preferred English name, FIPS = code indicated for the place by the Federal Information Processing Standards, USP = U.S. Postal abbreviation for a state, ISO = 3-letter ISO code for nation, ISO2 = 2-letter ISO code for a nation, ISON = numeric ISO code for a nation, DIS = special display name to be used when place is displayed as a broader context in a horizontal string	N	Legal values = ADJ, ADJE, DIS, ENG, FIPS, ISO, ISO2, ISON, USP.
NAME_DETAIL				Table for repeating detail information about each name, including display date, contributor, and link to citation.		
geog_key	int	4	0	Unique key for the place.	7011179	
placename_id	int	4	0	Unique key for this name.	47413	

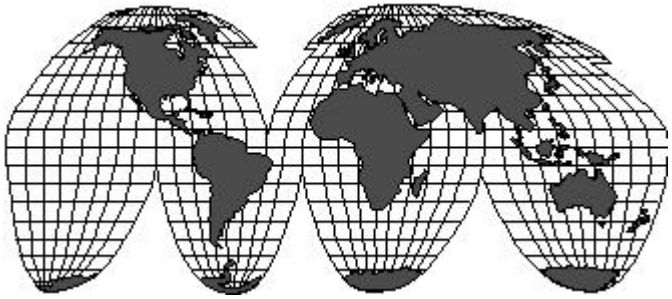
COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
bib_key	int	4	1	Unique key for the source where this name was found.	9006009	
display_ date	varchar	125	1	Note describing the time period when the name was used, and other information about the name.	used since 13th cen.	
start_date	varchar	30	1	Date indicating the earliest actual or estimated year delimiting the span of time indicated in the display date.	1200	Dates BCE are preceded by a minus sign.
end_date	varchar	30	1	Date indicating the latest actual or estimated year delimiting the span of time.	9999	Dates BCE are preceded by a minus sign. If the name is still used for the place, this date is "9999."
bib_page	varchar	40	1	Volume (if applicable) and page number, plate number, etc., for the source where the name was found.	V, 123	
contrib	varchar	30	0	Acronym, initials, or brief name of the project or institution that contributed the name.	VP	
GEOG_PLACETYPE				Table for place types associated with the place, contains code for place type, and various flags.		
geog_key	int	4	0	Unique key for the place.	7011179	
seqn	int	1	0	Sequence position in which the place type should be ordered in displays.	0	Values = 0-x.
pref_flag	char	1	0	Flag indicating if the place type is P = Preferred, or N = Non-preferred	P	
placetype_ code	int	4	0	Numeric code for the place type.	83002	For a given *geog_key*, a place type may be linked only once.
age_flag	char	1	0	Flag indicating if the place type is C = Current, H = Historical, or B = Both	C	

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
PLACETYPE_DATES				Table for repeating detail information about the place type, including display date.		
geog_key	int	4	0	Unique key for the place.	7008736	
placetype_code	int	4	0	Numeric code for the place type.	80132	
display_date	varchar	125	0	Note describing the time period when the place type applied, and other information about the place type.	founded by Etruscans	
start_date	varchar	30	1	Date indicating the earliest actual or estimated year delimiting the span of time indicated in the display date.	-800	Dates BCE are preceded by minus sign.
end_date	varchar	30	1	Date indicating the latest actual or estimated year delimiting the span of time indicated in the display date.	9999	Dates BCE are preceded by a minus sign. If the name is still used for the place, this date is "9999."
BIBLIO				Table for citations for sources, including key for citation, brief citation, and full citation.		
bib_key	int	4	0	Unique key for the source.	9006009	
bib_brief	varchar	40	0	Brief citation for the source; to be used in displays.	TCI: Toscana (1984)	
bib_brief_sort	varchar	40	0	Normalized brief citation; all caps with punctuation, diacritic codes, numbers, and spaces removed.	TCITOSCANA	
citation	text	16	1	Full citation for the source.	Toscana (non compresa Firenze). Guida d'Italia del Touring Club Italiano. Milano: Touring Club Italiano, 1984.	

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
PLACETYPE_LIST				Table for place types, containing code for place type and place type term.		
placetype_code	int	4	0	Numeric code for the place type.	83002	
placetype_search	char	4	0	Character string version of the code; to allow searching by truncation.	83002	
placetype	varchar	64	0	Place type term.	inhabited place	
NAME_KEYWORDS				Table containing keywords for the names.		
geog_key	int	4	0	Unique key for the place.	7011796	
placename_id	int	4	0	Unique key for this name.	47413	
keyword	varchar	40	0	A normalized version of a word contained in the place name; all caps with punctuation, diacritic codes, numbers, and spaces removed.	SIENA	
BIB_KEYWORDS				Table containing keywords for the citations.		
bib_key	int	4	0	Unique key for the source.	9006009	
keyword	varchar	40	0	Normalized word from the citation; all caps with punctuation, diacritic codes, numbers, and spaces removed.	TOSCANA	
FORMER_GEOGKEYS				Table containing mapping of *geog_key*s that have been merged or deleted since last release.		
new_key	int	4	0	Key of merged record.	7011179	
master_key	int	4	0	Key of the record that was the “master” in the merge operation.	1000982	

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
collapse_ key	int	4	0	Key of a record that was a "member" in the merge operation.	4000985	A merge operation involves one master record and one or more "member" records. The new record inherits the defaults of the "master" where there is conflict; the new record may then be edited.
FORMER_GEOGLINKS				Table containing former links to children of *geog_key*s that have been merged or deleted.		
new_key	int	4	0	Key of merged record.	7011179	
collapse_ key	int	4	0	Key of a record that was a "member" in the merge operation, and thus lost children when the merge occurred.	4000985	
old_link_ key	int	4	0	Key of the record that used to be a child of a "member" involved in a merge operation.	4002973	
FORMER_BIBKEYS				Table containing mapping of *bib_keys* that have been merged or deleted since the last release.		
new_key	int	4	0	Key of merged record.	9006009	
master_key	int	4	0	Key of the record that was the "master" in the merge operation.	9000098	
collapse_ key	int	4	0	Key of a record that was a "member" in the merge operation.	9000954	A merge operation involves one master record and one or more "member" records. The new record inherits the defaults of the "master," though the new record may then be edited.
FORMER_BIBLINKS				Table containing *geog_key*s that were linked to citations that have been merged or deleted.		
new_key	int	4	0	Key of merged record.	9006009	

COLUMN	*TYPE*	*Len*	*Nulls*	*DESCRIPTION*	*EXAMPLE*	*NOTE*
collapse_ key	int	4	0	Key of a record that was a “member” in the merge operation.	9000954	
old_link_ key	int	4	0	*Geog_key* of the place record that used to have this *bib_key* linked as a source.	7002973	



CHAPTER 3
REC FORMAT

OVERVIEW

The data releases use an 8-bit, code-extended ASCII character set. See *Appendix A: Codes for Diacritics* used in the TGN. The database is at least 500 megabytes in size.

The TGN was compiled using a custom-built editorial system that allowed the Vocabulary Program 1) to merge records from separate contributors that represent the same place, 2) to build and rearrange hierarchies, and 3) to accept, edit, and manage contributions to the database.[4]

The TGN REC format contains the following types of information:

Place Record	
GENERAL RECORD INFORMATION	The unique numeric ID, status, and date of the record.
HIERARCHY FOR DISPLAY	The name of the place together with its parents, displayed in a horizontal string.
MAIN HIERARCHY	Unique ID of the immediate parent of the place at hand, used to build full hiearchies.
SIMPLE HIERARCHY	The names of the administrative parents of the place at hand, to be used for filling in values in flat files.
PREFERRED NAME	The TGN preferred name and associated flags.
NAME DETAIL	Contributors, dates, and sources for the preferred name.
ALTERNATE NAMES	Alternate names for the place, and associated flags(the preferred English name is included among these names).
NAME DETAIL	Contributors, dates, and sources for the alternate name.
PREFERRED PLACE TYPE	The preferred place type for the place, together with flags and dates.
DATES	Dates for the preferred place type.
OTHER PLACE TYPES	Other place types that are applicable to the place, together with flags and dates.
DATES	Dates for the other place types.
COORDINATES	The latitude and longitude for the place.
DESCRIPTIVE NOTE	Note describing the place.
Place Type List	
PLACE TYPE	List of terms and codes for place types.
Bibliography	
SOURCE	List of unique ID, brief citation, and full citation for sources.

TGN:REC uses a mnemonic field labeling convention, each section of the record beginning with a tag. Each TGN REC record is divided in sections corresponding to the chart above. The heading names begin with an asterisk (*), are in title case, and end with a carriage return/line feed <CR><LF> (ASCII decimal character 013 followed by ASCII decimal character 010). The field tags are in all upper case, and

[4] Contributions will be accepted from selected institutions in an automated format prescribed by the Vocabulary Program. The contribution format and a new Vocabulary Coordination System are scheduled to be in place by Spring, 2001. The original TGN editorial system operated under the Sybase Relational Database Management System and Sybase SQL Server release 4.9.2; it was developed using APT Edit release 4.0.2. TGN ran on Sun SPARC workstations running Sun OS release 4.3.4.

correspond to the tags listed in the *Data Dictionary* below. Each field tag is followed by an equal sign (=), and the equal sign is followed by the data. The end of the data is marked by a carriage return/line feed <CR><LF>. Each TGN REC place record begins with the first section heading and carriage return/line feed (***General Record Information<CR><LF>**), and ends with two dollar signs and a carriage return/line feed **($$<CR><LF>**). Each instance of a repeating group will be preceded by the *header for that group. Repeating groups may be nested.

EXAMPLE

```
General Record Information<CR><LF>
STATUS=n<CR><LF>
TGNIDNO=4001715<CR><LF>
DATENT=19980316<CR><LF>
*Hierarchy for Display<CR><LF>
DISPHIER=Bosco della Fontana (Lombardia, Italia, Europe)<CR><LF>
*Main Hierarchy<CR><LF>
FACET=Current World<CR><LF>
PARENT=7003237<CR><LF>
*Simple Hierarchy<CR><LF>
CONTINT=Europe<CR><LF>
NATION=Italia<CR><LF>
FIRSTLEV=Lombardia<CR><LF>
*Preferred Name<CR><LF>
PREFNAME=Bosco della Fontana<CR><LF>
PNAFLAGS=C,V<CR><LF>
PNSORT=BOSCODELLAFONTANA<CR><LF>
PNAID=145113<CR><LF>
*Name Detail
PNCONTR=BHA<CR><LF>
PNASOURC=TCI: Lombardia (1970)<CR><LF>
PNASRCID=9004901<CR><LF>
PNAPAGE=552<CR><LF>
*Name Detail
PNCONTR=VP<CR><LF>
PNASOURC=TCI: Lombardia (1987)<CR><LF>
PNASRCID=9005691<CR><LF>
PNAPAGE=788<CR><LF>
*Preferred Place Type<CR><LF>
PREFPTP=forest<CR><LF>
PPTCODE=21642<CR><LF>
*Coordinates<CR><LF>
LAT=45 00<CR><LF>
LATDIR=N<CR><LF>
LONG=010 40<CR><LF>
LONGDIR=E<CR><LF>
LATNUM=45.000<CR><LF>
LONGNUM=10.666<CR><LF>
*Descriptive Note<CR><LF>
DESCNOTE=One of the few remaining forests in the Po delta.<CR><LF>
$$<CR><LF>
```

The TGN REC format contains the following fields:

THE PLACE RECORD

GENERAL RECORD INFORMATION
- *STATUS record status*
- *TGNIDNO unique identification*
- *DATENT date entered*

HIERARCHY FOR DISPLAY
- *DISPHIER display hierarchical position*

MAIN HIERARCHY
- *FACET display hierarchical position*
- *PARENT immediate parent key*
- *PARPREF preferred flag*

SIMPLE HIERARCHY
- *CONTINT continent*
- *NATION nation*
- *FIRSTLEV first level administrative subdivision*
- *SECNDLEV second level administrative subdivision*
- *INHABPL inhabited place*

PREFERRED NAME
- *PREFNAME preferred name*
- *PNAFLAGS language and history flags*
- *PNASORT sort string*
- *PNAID preferred name ID*

  NAME DETAIL
  - *PNCONTR contributor*
  - *PNADSP name display date*
  - *PNASTRT start date*
  - *PNAEND end date*
  - *PNASOURC source*
  - *PNASRCID source ID*
  - *PNAPAGE page*

ALTERNATE NAMES
- *ALTNAME alternate name*
- *ANAFLAGS language and history flags*
- *ANASORT sort string*
- *ANAID alternate name ID*
- *ANASEQN sequence of name*

  NAME DETAIL
  - *ANCONTR contributor*
  - *ANADSP name display date*
  - *ANASTRT start date*
  - *ANAEND end date*
  - *ANASOURCE source*
  - *ANASRCID source ID*
  - *ANAPAGE page*

PREFERRED PLACE TYPE
- *PREFPTP preferred place type*
- *PPTCODE code for placetype*
- *PPTAGE page*

  DATES
  - *PPTDSP placetype display date*
  - *PPTSTRT start date*
  - *PPTEND end date*

OTHER PLACE TYPES
- *OTHERPTP other place types*
- *OPTCODE code for placetype*
- *OPTAGE place type age*
- *OPTSEQN sequence of place type*

  DATES
  - *OPTDSP name display date*
  - *OPTSTRT start date*
  - *OPTEND end date*

COORDINATES
- *LAT latitude*
- *LATDIR direction of latitude*
- *LONG longitude*
- *LONGDIR direction of longitude*
- *LATNUM latitude decimal*
- *LONGNUM longitude decimal*

DESCRIPTIVE NOTE
- *DESCNOTE descriptive note*

THE PLACE TYPE LIST

PLACE TYPE
- *PTYP place type*
- *PTYPCODE place type code*

THE BIBLIOGRAPHY

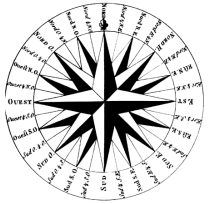
SOURCE
- *BIBKEY unique identification*
- *BIBBRIEF brief citation*
- *BIBBRSRT brief citation sort string*
- *CITE full citation*

REC FORMAT DATA DICTIONARY

In the following chart, the names of sections appear in small caps (**GENERAL RECORD INFORMATION**). The names of field tags are in uppercase (**STATUS**). For each field, the following is indicated: the field tag label, whether or not the field is repeatable, a description of the data, an example of the data, an explanatory note, and the columns to which this field corresponds in the TGN Relational Files format.

EXAMPLE

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO REL. TABLES FORMAT*
The Place Record						
GENERAL RECORD INFORMATION		N				
STATUS	***record status***	N	An alphabetical code indicating if the record is new (**n**), corrected/revised (**c**), or deleted (**d**).	n	For first edition of TGN, all records are New.	N/A
TGNIDNO	***unique identification***	N	A numeric code that uniquely identifies each place represented in TGN.	7011179		geog_main. geog_key
DATENT	***date entered***	N	The date a record was created [yyyymmdd].	20000602		N/A



Getty Thesaurus of Geographic Names
REC format DATA DICTIONARY

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
Place Record						
GENERAL RECORD INFORMATION		N				
LEN	***length of record***	N	Number of characters in the entire record including <carriage return>, <line-feed>, and all delimiters.	3453		
STATUS	***record status***	N	An alphabetical code indicating if the record is new (**n**), corrected/revised (**c**), or deleted (**d**).	n		N/A
TGNIDNO	***unique identification***	N	A numeric code that uniquely identifies each place represented in the TGN.	7011179		geog_main. geog_key
DATENT	***date entered***	N	The date a record was created [yyyymmdd].	20000602		N/A
HIERARCHY FOR DISPLAY		Y				

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
DISPHIER	***display hierarchical position***	N	The full list of parents for the place in display (note) format, listing parents in ascending order, to the level of continent. Names are in the vernacular, where applicable.	Siena (Siena province, Toscana, Italia, Europe)	If a place has multiple parents, only one parent is included in this string.	Display names or preferred names of parents traced through geog_parents. parent_key.
MAIN HIERARCHY		Y				
FACET	***facet***	N	The facet in which the place is located.	Current World	In the first release of TGN:REC, all places are in the Current World facet	N/A
PARENT	***immediate parent key***	N	The unique numeric ID of the place that is the immediate parent for the place represented in this record.	7003168		geog_parents. parent_key
PARPREF	***preferred flag***	N	Flag indicating whether or not this parent is the preferred parent.	P	legal values are P or N.	geog_parents pref_flag
SIMPLE HIERARCHY		Y				
CONTINT	***continent***	Y	Preferred name of the continent on which the place is located.	Europe	Since places may have more than one parent, this field and the following fields may repeat together as a group.	**continent** traced through geog_parents. parent_key.
NATION	***nation***	Y	If the place is located in a nation,the preferred name of that nation.	Italia		**primary political unit** traced through geog_parents. parent_key

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
FIRSTLEV	***first level administrative subdivision***		If the place is located in a first level subdivision in a nation, the name of that subdivision.	Toscana		**first level subdivision** traced through geog_parents. parent_key
SECNDLEV	***second level adminsitrative subdivision***		If the place is located in a second level subdivision in a nation, the name of that subdivision.	Siena province		**second level subdivision** traced through geog_parents. parent_key
INHABPL	***inhabited place***		For neighborhoods and other subdivisions of inhabited places, the name of the inhabited place in which the neighborhood is located.	*[e.g., for Montmartre neighborhood]* Paris		**inhabited place** or **deserted settlement** traced through geog_parents. parent_key
PREFERRED NAME		N				
PREFNAME	***preferred name***	N	The preferred name for the place in the TGN, usually the vernacular name.	Siena		geog_name. place_name

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
PNAFLAGS	***language and history flags***	N	Flags indicating information about the current/historical status (C or H) of the name, whether or not it is a vernacular name (V or O). Other flags are those indicating if name is ADJ = adjectival name form, ADJE = preferred English version of adjectival name form, ENG = preferred English name, FIPS = code indicated for the place by the Federal Information Processing Standards, USP = U.S. Postal abbreviation for a state, ISO = 3-letter ISO code for nation, ISO2 = 2-letter ISO code for a nation, ISON = numeric ISO code for a nation, DIS = special display name to be used when place is displayed as a broader context in a horizontal string, and whether or not it is the preferred English name, a special Display name, an ISO code, or a US Postal abbreviation for a US state.	C,V	C = Current H = Historical *V = Vernacular* *O = Other language* ADJ = adjectival ADJE = adjectival English DIS = display ENG = preferred English FIPS = FIPS code ISO = ISO 3-letter code ISO2 = ISO 2-letter code ISON = ISO numeric code USP = US Postal Code	geog_name. age_flag, geog_name.lang, geog_name. other_flags
PNASORT	***sort string***	N	Normalized name; all caps with punctuation, diacritic codes, numbers, and spaces removed.	SIENA		
PNAID	***preferred name ID***	N	The unique numeric key identifying this name.	47413		geog_name. placename_id
(NAME DETAIL)		Y				
PNCONTR	***contributor***	N	The initials or acronym of the contributor of the preferred name.	BHA		name_detail. contrib
PNADSP	***name display date***	N	Note describing the time period when the name was used, and other information about the name.	used since 13th cen.	Note that in merged TGN record, there may be only one date per name.	name_detail. display_date

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
PNASTRT	***start date***	N	Date indicating the earliest actual or estimated year delimiting the span of time indicated in display date above.	1200	Dates BCE are preceded by minus sign.	name_detail. start_date
PNAEND	***end date***	N	Date indicating the latest actual or estimated year delimiting the span of time indicated in display date.	9999	Dates BCE are preceded by minus sign. If the name is still used for the place, this date is **9999.**	name_detail. end_date
PNASOURC	***source***	N	Brief citation for the source where the preferred name was found.	Hook, Siena (1979)		biblio. bib_brief (linked to name_detail through bib_key)
PNASRCID	***source ID***	N	Unique numeric key identifying this source.	9000009		name_detail. bib_key
PNAPAGE	***page***	N	Volume (if applicable) and page number, plate number, etc., in source where name was found.	6 ff.		name_detail. bib_page
ALTERNATE NAMES		Y				
ALTNAME	***alternate name***	N	Any alternate name(s) associated with the place, PLUS the same added fields as for Pref name.	Saena Julia		geog_name. place_name & geog_name. pref_flag; plus geog_name. seqn
ANAFLAGS	***language and history flags***	N	See PNAFLAGS above	H,O	See PNAFLAGS above	geog_name. age_flag, geog_name.lang, geog_name. other_flags

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
ANASORT	***sort string***	N	Normalized name; all caps with punctuation, diacritic codes, numbers, and spaces removed.	SAENAJULIA		
ANAID	***alternate name ID***	N	The unique numeric key identifying this name.	141380		geog_name. placename_id
ANASEQN	***sequence of name***	N	Sequence position in which the name will be ordered.	4	Values = 1-x.	name_detail. seqn
(NAME DETAIL)		Y				
ANCONTR	***contributor***	N	The initials or acronym of the contributor of the alternate name.	FDA		name_detail. contrib
ANADSP	***name display date***	N	Note describing the time period when the name was used, and other information about the name.	Roman		name_detail. display_date
ANASTRT	***start date***	N	Date indicating the earliest actual or estimated year delimiting the span of time indicated in display date above.	-100	Dates BCE are preceded by minus sign.	name_detail. start_date
ANAEND	***end date***	N	Date indicating the latest actual or estimated year delimiting the span of time indicated in display date.	300	Dates BCE are preceded by minus sign. If the name is still used for the place, this date is **9999.**	name_detail. end_date
ANASOURC	***source***	N	Brief citation for the source where the alternate name was found.	Webster's Geographical Dictionary (1984)		biblio. bib_brief (linked to name_detail through bib_key)
ANASRCID	***source ID***	N	Unique numeric key identifying this source.	9006449		name_detail. bib_key
ANAPAGE	***page***	N	Volume (if applicable) and page number, plate number, etc., in source where name was found.	1114		name_detail. bib_page

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
PREFERRED PLACE TYPE		N				
PREFPTP	***preferred place type***	N	Word or phrase that characterizes a significant aspect of the place, including its size, function, role, political anatomy, or physical characteristics.	inhabited place		placetype_list. placetype holds value (linked to record by geog_ placetype. placetype _code)
PPTCODE	***code for placetype***	N	Numeric code for the placetype.	83002		geog_ placetype. placetype _code
PPTAGE	***place type age***	N	Flag indicating if the preferred place type is is Current (C), Historical (H), or Both current and historical (B).	C		geog_placetype. age_flag
(DATES)		Y				
PPTPDSP	***placetype display date***	N	Note describing the time period when the place type applied, and other information about the place type.	settled by Etruscans	Place type dates are repeatable if the placetype applied in separate periods of time (e.g., capital during Roman times, again in Middle Ages).	placetype_ dates.display_date
PPTPSTRT	***start date***	N	Date indicating the earliest actual or estimated year delimiting the span of time indicated in display date above.	-800	For inhabited place and deserted settlement, indicates span of habitation. Dates BCE are preceded by minus sign.	placetype_ dates. start_date

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
PPTPEND	***end date***	N	Date indicating the latest actual or estimated year delimiting the span of time indicated in display date.	9999	Dates BCE, are preceded by minus sign. If the name is still used for the place, this date is **9999.**	placetype_ dates. end_date
OTHER PLACE TYPES		Y				
OTHERPTP	***other place type***	N	Word or phrase that characterizes an additional significant aspect of the place, including its size, function, role, political anatomy, or physical characteristics.	commune (administrative)		placetype_list. placetype holds value, (linked to record by geog_placetype. placetype _code)
OPTCODE	***code for placetype***	N	Numeric code for the placetype.	81112		geog_ placetype. placetype _code
OPTAGE	***place type age***	N	Flag indicating if this place type is is Current (C), Historical (H), or Both current and historical (B).	C		geog_placetype. age_flag
OPTSEQN	***sequence of place type***	N	Sequence position in which the place type will be ordered.	2	Values = 1-x.	geog_placetype. seqn
(DATES)		Y				
OPTPDSP	***placetype display date***	N	Note describing the time period when the place type applied, and other information about the place type.	since 1125	Place type dates are repeatable if the placetype applied in separate periods of time (e.g., capital during Roman times, again in Middle Ages).	placetype_ dates.display_date
OPTPSTRT	***start date***	N	Date indicating the earliest actual or estimated year delimiting the span of time indicated in display date above.	1125	Dates BCE are preceded by minus sign.	placetype_ dates. start_date

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
OPTPEND	***end date***	N	Date indicating the latest actual or estimated year delimiting the span of time indicated in display date.	9999	Dates BCE are preceded by minus sign. Where placetype is still applicable, value is **9999**.	placetype_ dates. end_date
COORDINATES		N	The point of latitude and longitude on the globe representing the location of the place.			
LAT	***latititude***	N	The degrees and minutes of the latitude, which is the angular distance north or south of the equator, measured in degrees and minutes along a meridian.	43 19	Format is two digits, one space, two digits.	geog_main. lat
LATDIR	***direction of latitude***	N	Letter indicating if latitude is North or South of equator.	N	Legal values = N or S.	geog_main. lat_dir
LONG	***longitude***	N	The degrees and minutes of the longitude, which is the angular distance east or west of the Prime Meridian at Greenwich, England.	011 19	Format is three digits, one space, two digits.	geog_main. long
LONDIR	***direction of longitude***	N	Letter indicating if latitude is East or West of the Prime Meridian.	E	Values = E or W.	geog_main. long_dir
LATNUM	***latitude decimal***	N	Latitude expressed as decimal fractions of degrees.	43.316	Whole degrees are represented by a two-digit decimal number ranging from 0 through 90, with any decimal fraction of a degree separated by a decimal point; latitudes south of the equator are designated by a minus sign.	geog_main. lat_num

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO RELATIONAL FILES FORMAT*
LONGNUM	***longitude decimal***	N	Longitude expressed as decimal fractions of degrees.	11.316	Whole degrees are represented by a three-digit decimal number ranging from 0 through 180, with any decimal fraction of a degree separated by a decimal point; longitudes west of the Prime Meridian are designated by a minus sign.	geog_main. long_num
DESCRIPTIVE NOTE		N				
DESCNOTE	***descriptive note***	N	Note describing the history, physical characteristics and significance of the place.	Founded by Etruscans; ruled by Romans 3rd cen. BC & Lombards 6th cen.; constitution 1179; medieval seat of Ghibelline faction; …		geog_main. descr_note

Place Type List

PLACE TYPE						
PTYP	***place type***	N	Place type term.	inhabited place		placetype_list. placetype
PTYPCODE	***place type code***	N	A numeric code that identifies each place type.	83002		placetype_list. placetype_code

Bibliography

SOURCE						
BIBKEY	***unique identification***	N	A numeric code that uniquely identifies each source represented in the TGN.	9000009		biblio.bib_key
BIBBRIEF	***brief citation***	N	Brief citation for the source; to be used in displays.	Hook, Siena (1979)		biblio.bib_brief
BIBBRSRT	***brief citation sort string***	N	Normalized brief citation; all caps with punctuation, diacritic codes, numbers, and spaces removed	HOOKSIENA		biblio. bib_brief_sort
CITE	***full citation***	N	Full citation for the source.	Hook, Judith. Siena: A City and Its History. London: Hamish Hamilton, 1979.		biblio.citation

CHAPTER 4
USMARC FORMAT

OVERVIEW

The TGN MARC data release uses an 8-bit, code-extended ASCII character set. Diacritics for the TGN MARC release are not the same as listed in *Appendix A;* instead, the TGN MARC release uses the ANSEL character set (ANSI Z39.47).[5] The database is at least 500 megabytes in size. The TGN was compiled using a custom-built editorial system that allowed the Vocabulary Program 1) to merge records from separate contributors that represent the same place, 2) to build and rearrange hierarchies, and 3) to accept, edit, and manage contributions to the database.[6]

The TGN USMARC release format includes the heading for the place record, its hierarchical context, latitude and longitude, its place types, bibliography and contributors, and notes. Although a brief introduction to the structure of a USMARC authority record is given here, please see http://lcweb.loc.gov/marc for a complete description of the *USMARC Format for Authorities Data*.

The MARC portion of a TGN:MARC record consists of four sections arranged in the following order:

- **Record leader**
- **Record directory**
- **Control fields**
- **Variable fields**

Each TGN:USMARC record ends with a record terminator (Hex 1D). In addition, a <carriage return> and a <line feed> (Hex 0D0A) follow the end of each record. The file is thus not stream file, but an ASCII text file with 3 non-printing characters (Hex 1F, Hex 1E (field terminator), and Hex 1D) used as delimiters.

Record Leader: The first 24 character positions (0-23) of a record are the record leader and are used for computer processing of the record. Leader elements are not repeatable.

Record Directory: The record directory is an index to the location of each variable control field and data field within a record. It begins at character position 24. Every field (control or variable) in a record has a directory entry. Each entry gives the field's tag, length, and starting character position. At the end of the whole series of directory entries is a field terminator (Hex 1E).

Control Fields: Control fields provide information useful for processing a record such as identification numbers and dates. Each control field is identified by a three-character numeric tag. Control fields do not contain indicator positions or subfield codes.

Variable Fields: Variable fields contain the TGN preferred name, variant names, and related data. Data within the variable fields is formatted using tags, indicators, and subfield codes. These conventions are

[5] For the ANSEL character set, see "MARC 21: Specifications for Record Structure, Character Sets, and Exchange Media" at http://lcweb.loc.gov/marc/specifications .

[6] Contributions will be accepted from selected institutions in an automated format prescribed by the Vocabulary Program. The contribution format and a new Vocabulary Coordination System are scheduled to be in place by Spring, 2001. The original TGN editorial system operated under the Sybase Relational Database Management System and Sybase SQL Server release 4.9.2; it was developed using APT Edit release 4.0.2. TGN ran on Sun SPARC workstations running Sun OS release 4.3.4.

referred to in USMARC as *content designators*, since they provide information about the field and identify the type of data that follows them. Each field ends with an end-of-field character (Hex 1E).

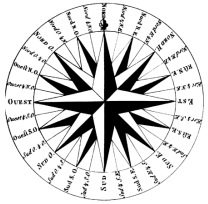
Tags: A tag is a three-character numeric label for a field, (e.g., 150). Tags appear only in the directory portion of an TGN:MARC record, not in the variable field portion. Indicators follow the tag and further define the information in the field. Indicators consist of two characters: numeric (e.g., 01), blanks, or a combination of a blank and a number. Subfield codes separate and distinguish information within afield. The format of subfield content designators consists of a delimiter, "‡", and a lower-case alphabetical character or number, (e.g., ‡w, ‡3). *A record may or may not contain all of these variable fields.* Some fields, including 151, 451, 670, and 902, may be repeated in separate occurrences within a record. *Each field ends with a field terminator (Hex 1E).*

EXAMPLE

```
Len 00914
Sta n
Typ z
Lev
Bas 00157
Enc n
001 7016740¶
008 000602 n anznnbabn           b ana     d¶
151   ‡aMercia (England, United Kingdom, Europe) historic
      region‡c(7016740,scheme=TGN)‡dH,V,N‡eMERCIA‡f150562¶
551   ‡wg‡aEngland (7002445,scheme=TGN)¶
670   ‡aColumbia Lippincott Gazetteer (1961) [BHA]‡bMercia¶
670   ‡aCanby, Historic Places (1984) [VP]‡bMercia¶
670   ‡aEncyclopædia Britannica  (1988) [VP]‡bMercia¶
678   ‡iArea comprising modern Staffordshire, Derbyshire,
      Nottinghamshire & parts of West Midlands & Warwickshire; was
      Anglo-Saxon kingdom; gained greatest power under King Offa (ruled
      757-96); declined after rise of kings of Wessex & divided by Danes
      in 9th cen.¶
901   ‡ahistoric region‡b81501‡cC¶
902   ‡akingdom‡b82141‡cH‡d1‡efrom ca. 500 until dissolved by Danish
      attack in 874‡f500‡g874¶
905   ‡aEurope‡bUnited Kingdom‡cEngland¶
```

TGN MARC FORMAT DATA DICTIONARY

In the following chart, the names of field tags are in left column, followed by the name of the field, an indication whether or not the field is repeatable, a description of the data, an example of the data, an explanatory note, and the fields to which this field corresponds in the TGN REC format.



Getty Thesaurus of Geographic Names
USMARC format DATA DICTIONARY

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO REC FORMAT*
Place Record *(USMARC format)*						
RECORD LEADER						
Len	*length*	N	Character position 0-4. Length of the record.	03069	Total number of characters in the record; including the record terminator character, but not the <carriage return><line feed>.	LEN
Sta	*record status*	N	Character position 5. An alphabetical code indicating if the record is new (**n**), corrected/revised (**c**), or deleted (**d**).	n		STATUS
Typ	*record type*	N	Character position 6. Type of record (**z** indicates "authority")	z		
Lev	*undefined*	N	Character position 7-9.	[spaces]		
	indicator count	N	Character position 10. Number of character positions used for indicators at the beginning of variable fields.	2		

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO REC FORMAT*
	subfield code count	N	Character position 11. Number of the characer position used for each subfield code in a variable data field.	2		
Bas	*base address*	N	Character position 12-16. Base address of the data.	00457	The computer-generated, five-character numberic string that indicates the first character positon of the first variable control field in a record.	
Enc	*encoding level*	N	A one-character code that indicates whether the record is complete. Values are **n** or **o**.	n	n = Compele authority record; the record meets national level record requirements. o = Incomplete authority record; TGN records do not always contain full content designation.	
	undefined	N	Character position 18-19.	[spaces]		
	length of field length	N	Length of the length-of-field portion of the record.	4		
	starting character length	N	Character position 21. Length of the starting character positoin portion of each record directory entry.	5		
	implementation position	N	Character position 22. Length of the implementation portion of each record directory entry.	0		
	undefined	N	Character position 23.	0		

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO REC FORMAT*
RECORD DIRECTORY						
	tag	N	Character position 0-2. Three-character numeric symbol that identifies a control or variable field	`001`	Directory begins at position 24, and subsequent positions are counted from there.	
	field length	N	Character position 3-6. Number of characters in the field.	`0008`		
	starting character position	N	Character position 7-11. The character position of the first character of the field relative to the first character of the field control field that follows the record directory.	`00000`		
CONTROL FIELDS						
001	*TGN unique identification*	N	A numeric code that uniquely identifies each place represented in TGN.	`7011179`		TGNIDNO
008	*date entered*	N	The date a record was created [yymmdd], and other information.	`000602 n anznnbabn b ana d`		DATEENT
	date entered on file	N	Character positions 0-5. The year, month, day that a TGN record was created.	`000602`		
	geographic place	N	Character position 6. Direct or indirect geographic subdivision.	`[space]`		
	Romanization scheme	N	Character position 7. Type of Romanization scheme used for transliterations.	`n`	**n** = not applicable	
	language of catalog	N	Character position 8. Language in which the record is written.	`[space]`	[space] = no information provided.	
	kind of record	N	Character position 9. Indicates the kind of term in the 150 field.	`a`	**a** = preferred name.	
	descriptive cataloging rules	N	Character position 10. Cataloging rules used for the record.	`n`	**n** = not applicable	

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO REC FORMAT*
	subject heading system / thesaurus	N	Character position 10. Indicates the subject heading system or thesaurus building conventions used to formulate the contents of the 1xx field.	z		
	series type	N	Character position 12. Type of series.	n	**n** = not applicable	
	numbered series	N	Character position 13. Indicates if the record is a numbered or unnumbered series.	n	**n** = not applicable	
	heading use—main or added entry	N	Character position 14. Indicates if the preferred name, when used as a heading, is appropriate as a main or added entry.	b	**b** = not appropriate	
	heading use—subject added entry	N	Character position 15. Indicates if the preferred name, when used as a heading, is appropriate as a subject added entry.	a	**a** = appropriate	
	heading use—series added entry	N	Character position 16. Indicates if the preferred name, when used as a heading, is appropriate as a series added entry.	b	**b** = not appropriate	
	subject subdivision	N	Character position 17. Type of subject subdivision.	n	**n** = not applicable	
	undefined	N	Character positions 18-27. Undefined character positions.	[spaces]		
	government agency	N	Character position 28. Type of government agency.	[space]	[space] = not a government agency	
	reference evaluation	N	Character position 29. Reference evaluation.	b		
	undefined	N	Character position 30. Undefined character position.	[space]		
	record update	N	Character position 31. Record update in process.	a	**a** = record can be used	
	personal name	N	Character position 32. Undifferentiated personal name.	n	**n** = not applicable	

LABEL	FIELD NAME	REP	DESCRIPTION	EXAMPLE	NOTE	MAPPING TO REC FORMAT
	level of establishment	N	Character position 33. Indicates the extent to which the 1xx heading conforms to the subject heading system or thesaurus convention indicated in 008/11.	a	**a** = fully established (descriptors) **n** = not applicable (guide terms)	
	undefined	N	Character position 34-37. Undefined character positions.	[spaces]		
	modified record	N	Character position 38. Indicates whether record is modified from its intended content (e.g., missing characters due to character set limitations).	[space]	[space] = not modified	
	cataloging source	N	Character position 39. Indicates the creator of the record.	d	**d** = organization other than a national bibliographic agency or participant in a cooperative cataloging program	
VARIABLE FIELDS						
151	*preferred name of the place, display hierarchical position and other information*	N	The preferred name for the entry, the full list of parents for the place in display (note) format, listing parents in ascending order, to the level of continent. Names are in the vernacular, where applicable. It also includes the preferred place type, latitude value, latitude direction, longitude value, longitude direction, TGN ID number, flag indicating if the name is vernacular or historical, the normalized sort name, the sequence number of the name, and the numeric ID of the name.	‡aSiena (Siena province, Toscana, Italia, Europe) inhabited place‡c43 19 N 011 19 E (7011179, scheme=TGN)‡dC, V,N‡eSIENA ‡f47413¶	If a place has multiple parents, only one parent is included in this string.	DISPHIER, PREFPTP, LAT, LATDIR, LONG, LONGDIR, TGNIDNO, PNAFLAGS, PNASORT, PNAID

LABEL	*FIELD NAME*	*REP*	*DESCRIPTION*	*EXAMPLE*	*NOTE*	*MAPPING TO REC FORMAT*
451	*alternate name, and additional information*	Y	It includes the alternate name, history and vernacular flags for the name, normalized sort name for the name, the ID for the name, and the sequence number for the name	‡aSaena Julia‡dH,V,N ‡eSAENAJULIA ‡f141380‡g4¶		ALTNAME, ANAFLAGS, ANASORT, ANAID, ANASEQN
551	*immediate parent*	Y	The name and unique numeric ID of the place that is the immediate parent for the place represented in this record.	‡wg‡aSiena province (7003168,scheme=TGN)¶		DISPHIER, PARENT
670	*sources for the record*	N	Published source for the record, contributor of the name, and the name found in the cited source	‡aAnnuario Generale (1980) [BHA]‡bSiena¶		PNASOURC, ANASOURC, PREFNAME, ALTNAME, PNCONTR, ANCONTR
678	*descriptive note*	N	Descriptive note for the place.	‡iFounded as Etruscan hill town; later was Roman city of Sena Julia (3rd cen. BC); flourished under Lombard kings (6th cen. AD); was medieval self-governing commune (constitution in 1179); was medieval seat of pro-imperial, Ghibelline power;…		DESCNOTE

LABEL	FIELD NAME	REP	DESCRIPTION	EXAMPLE	NOTE	MAPPING TO REC FORMAT
901	*preferred place type*	N	The preferred place type, its numeric code, the history flag, the display date note, the start and end dates for the display date.	‡ainhabited place‡b83002‡cC ‡esettled by Etruscans (flourished 6th cen. BCE)‡f-800 ‡g9999¶		PREFPTP, PPTCODE, PPTAGE, PPTPDSP, PPTPSTRT, PPTPEND
902	*other place types*	Y	Another place type, its numeric code, the history flag, the display date note, the start and end dates for the display date.	‡acommune (administrative) ‡b81112‡cC‡d2 ‡esince 1125 ‡f1125‡g9999¶		OTHERPTP, OPTCODE, OPTAGE, OPTSEQN, OPTPDSP, OPTPSTRT, OPTPEND
903	*coordinates*	N	Latitude value, latititude direction, longitude value, longitude direction, decimal latitude value, decimal longitude value	‡a43 19‡bN‡c011 19‡dE‡e43.317 ‡f11.317¶		LAT, LATDIR, LONG, LONGDIR, LATNUM, LONGNUM
905	*simple hierarchy*	N	The primary administrative subdivision levels: continent, nation, first level subdivision, second level subdivision	‡aEurope‡bItalia ‡cToscana ‡dSiena province¶		CONTINT, NATION, FIRSTLEV, SECNDLEV, INHABPL

Place Type List *(simple file of full place type list; not USMARC format)*						
PTYP	*place type*	N	Place type term.	inhabited place		
PTYPCODE	*place type code*	N	A numeric code that identifies each place type.	83002		
Bibliography *(simple file of full bibliography; not USMARC format)*						
BIBKEY	*unique identification*	N	A numeric code that uniquely identifies each source represented in TGN.	9000009		
BIBBRIEF	*brief citation*	N	Brief citation for the source; to be used in displays.	Hook, Siena (1979)		
BIBBRSRT	*brief citation sort string*	N	Normalized brief citation; all caps with punctuation, diacritic codes, numbers, and spaces removed	HOOKSIENA		
CITE	*full citation*	N	Full citation for the source.	Hook, Judith. Siena: A City and Its History. London: Hamish Hamilton, 1979.		

CHAPTER 5
CONTENTS OF THE TGN

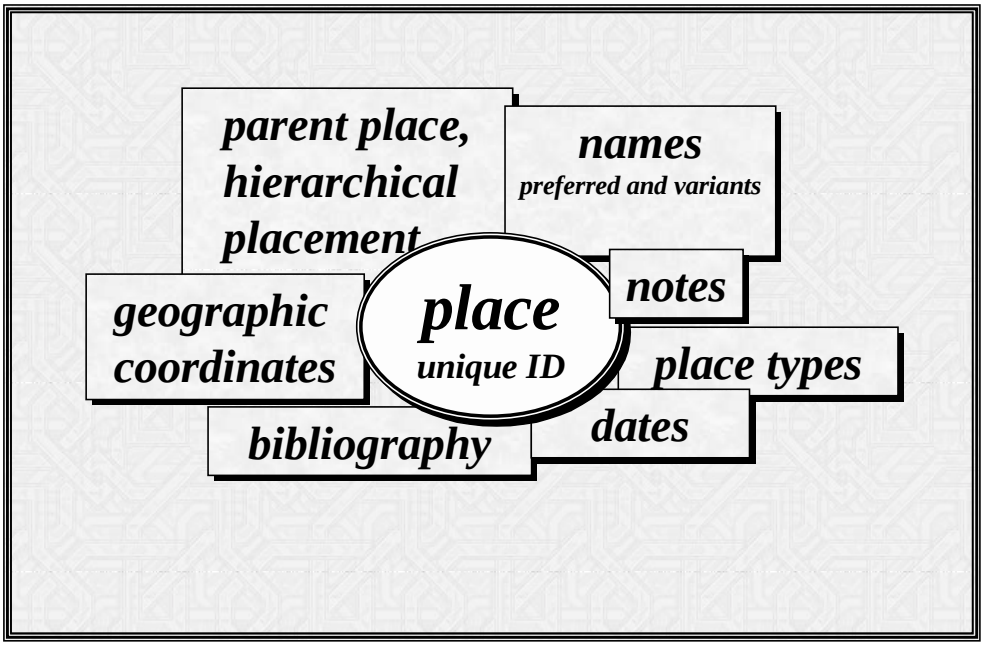
OVERVIEW

The *Getty Thesaurus of Geographic Names* (TGN) is a hierarchically arranged source of geographic data, global in scope. It is intended to be a source of geographic names for cataloging and retrieving records about art and cultural heritage, but it may have broader application in other disciplines. In the world of art, architecture, and material culture, geographic names are used to record the location of the art object, building, or artifact, its place of origin, the loci of activity of the creator, and the sites of the creator's birth and death.

The TGN provides users with vernacular and English names of places, as well as variant names in other languages and historical names for some places. Other information also may be included to identify and describe the place. Places in the TGN are arranged in hierarchies to show broader and narrower contexts, ranging from continents to nations to states to inhabited places.

The logical focus of a record[7] in the TGN is a ***place***. Attributes of the place include hierarchical position, names, "place types," coordinates, dates, and sources (contributors and bibliographic citations).

Logical focus of the TGN "record" is a place



[7] Note that for the purposes of this chapter, "record" refers to an intellectual record – not a database record. The intellectual record comprises data linked through the unique numeric ID for the place, the *geog_key*.

The record is identified uniquely in the database by a numeric key. For most records, the minimum information needed by a user to identify the place and intellectually distinguish it from other places include a name, place type, and position in the hierarchy. However, since there are so many homographs in the TGN, in some cases coordinates or a distinguishing note are also necessary to disambiguate places.

The TGN is a compiled resource. A TGN record may contain information from multiple contributors. Contributors include various Getty projects and qualified outside institutions who collect information about geographic places. The Vocabulary Program also contributes original information to the TGN. All information about a single place is merged to allow display as a single, uniform, homogenous set of information. At the same time, each contributor's individual contribution is flagged, so a display may alternatively illustrate the diversity and conflicting nature of information about a single place.

The TGN incorporates "flexible standards" in order to allow contributions from a wide variety of institutions with established, diverse practice. While critical standards for technical details, structure, content, and editorial guidelines are required of contributors, other content and editorial guidelines are considered non-critical and are therefore recommended, but not required. For example, it is critical that core fields are included and that they are indexed or formatted in a way that will allow retrieval. However, it is not critical that the "preferred name" be in the vernacular language (although this is highly recommended) or that the "descriptive note" be phrased according to a uniform style (although this is recommended too). Employing flexible standards means that the TGN database as a whole is not entirely consistent or totally uniform.

Relationships

The TGN is a thesaurus that includes equivalence, associative, and hierarchical relationships.

Equivalence Relationship. All relationships between names within the same TGN record are equivalence relationships. In the example below, all names refer to the same city, Lisbon, Portugal.

Lisboa (preferred)
Lisbon (preferred English)
Lisbonne
Lissabon
Lisabon
Olissibona
Ulixbone
Luzbona
Lixbuna
Felicitas Julia
Olisipo

Among all the names that refer to the place, one is indicated as the "preferred name," comparable to the "descriptor" in the AAT. This is the "vernacular" or local-language name most often found in scholarly or authoritative published sources (e.g., the name in bold in the example above). The English name is also indicated. Institutions who wish to use the TGN as an authority may use one of these two names to refer to the place consistently.

Variant and alternate names in the record include names in other languages, names transliterated into the Roman alphabet by various methods, names in natural or inverted form (particularly for physical features, e.g., "Etna, Mount"), nicknames, official names, and historical names. Misspellings may be included if they are found in published sources.

Hierarchical Relationship. The “hierarchy” in the TGN refers to the method of structuring and displaying the places within their broader contexts. Hierarchical relationships in the TGN represent part/whole relationships and are typically indicated with indention, as in the example below.

.....Europe.....(continent)
.....Italia.....(nation)
.....Toscana.....(region)
.....Siena..... (province)
.....Abbadia San Salvatore.....(inhabited place)
.....Argiano..... (inhabited place)
.....Asciano.... (inhabited place)
.....Badia a Isola.... (inhabited place)
.....Bagno Vignoni.....(inhabited place)
.....Buonconvento.... (inhabited place)
.....Campiglia dei Fosci.... (inhabited place)
.....Casciano...(inhabited place)
[...]
.....Scrofiano.....(inhabited place)
.....Serre di Rapolano.....(inhabited place)
.....Siena.....(inhabited place)
.....Sinalunga.....(inhabited place)
.....Sovicille.....(inhabited place)
.....Staggia..... (inhabited place)
.....Torrenieri.....(inhabited place)

The TGN is polyhierarchical, meaning that a place may have multiple "parents" or broader contexts. For example, the US state of Hawaii is administratively part of the United States in North America, but it is physically located in Oceania.

The TGN hierarchy includes both physical features and political/administrative entities. Major subdivisions of the hierarchy typically include the continent, nation, first level subdivision, second level subdivision, inhabited place, and possibly neighborhood. Most nations have one level of administrative subdivision above inhabited place, and many have two levels. Generally, the hierarchy in the TGN goes only to the level of the inhabited place. However, the level of neighborhood has been included for some of the world's largest cities. Releases of the TGN will generally not include levels for streets or buildings within the boundaries of inhabited places.

If users integrate the TGN’s hierarchical structure into their database of objects, they can link their object record to a specific inhabited place (e.g., *Campo Largo* in Argentina in the example below). However, they can still retrieve the record by searching for a broader context (e.g., *Chaco province, Argentina*), without having had to do redundant data entry of the inhabited place and all its broader contexts in each relevant object record. In a less sophisticated application, the TGN is still a source of the vernacular and English name (if applicable) of the place and its broader contexts, which users could implement in almost any automated or manual application.

Partial hierarchy showing places in Chaco province, Argentina

World
South America (continent)
Argentina (nation)
Chaco (province)
Barranqueras (inhabited place)
Basail (inhabited place)
Campo Largo (inhabited place)
Charadai (inhabited place)

Associative Relationship.[8] Associative relationships may exist between the records for places in the TGN. For example, if an inhabited place has been physically moved (as when the location has been deemed unsafe due to flood or earthquake), there should be an associative relationship between the original settlement and the new settlement. For example, Ocotepeque, Honduras was originally located to the NE of the current city, but was moved after the Marchala river flooded in 1935; *Ocotepeque* and *Nueva Ocotepeque* could be linked by the associative relationship.

SCOPE

The TGN contains more than one million names and other information for around 900,000 places. The places are arranged in hierarchies. The scope is global; all continents and all nations are represented. Physical features (including mountains and rivers) and administrative entities (including towns and nations) are included. Historical places may also be included. The target audience of the TGN is the art, architecture, and material culture communities.

The focus of the TGN record, the "place," is represented by a name and a "place type" in the hierarchy above. Each place record includes a name, a place type, and a position in the hierarchy. For most users, this is the minimum information needed to identify the place and intellectually distinguish it from other places.[9] The name alone does not identify it because there are many homographs among place names: For example, there are over 70 inhabited places called "Springfield" in the TGN. Many of them are located in various states in the USA; however, recording the state and nation alone still does not identify the place. Since one state can have multiple Springfields, it is necessary to record the county as well. Among the outstanding issues regarding implementation of the TGN (or any geographic information) in a retrieval tool or cataloging system is how to incorporate the parents' names (and their variant names and abbreviations) in a meaningful way.

[8] In the current edition of TGN, the associative relationships are not included; they weill be included in future relaeases.
[9] In an automated system, the place should be identified by the unique key (in the relational files, the *geog_key*).

Labels containing minimum information for two places with names that are homographs

Springfield (Delaware county, Pennsylvania, United States), inhabited place

Springfield (Montgomery county, Pennsylvania, United States), inhabited place

In addition to names, place types, and hierarchical position, a TGN record may include other information, including coordinates, dates, notes, and bibliography.

Geographic places in the TGN can be either physical or political entities. They include physical features such as continents, rivers, and mountains. Geographic places also include entities defined by political boundaries, such as empires, nations, states, districts, townships, cities, and neighborhoods. Although the Vocabulary Program focused editing on political entities, physical features were included in the database because they were considered useful by potential users of the TGN.

What is a geographic place?

Geographic places in the TGN include physical features such as continents, rivers, and mountains.

PARTIAL PLACE TYPE LIST FOR PHYSICAL FEATURES

ocean	pond	river	creek
lakes	lake	intermittent lake	bay
strait	glacier	river mouth	waterfall
desert	oasis	wetland	marsh
continent	subcontinent	fault	volcano
dunes	arroyo	canyon	fluvial island
mountain	mountain range	mountain system	hill
hills	plateau	valley	basin
field	plain	cape	island group
island	peninsula	reef	shore
cave	crater	depression	pampa
prairie	savanna	jungle	forest

Geographic places also include entities defined by political and administrative boundaries, such as empires, nations, states, districts, townships, and cities

PARTIAL PLACE TYPE LIST FOR ADMINISTRATIVE AND POLITICAL ENTITIES

primary political unit	independent sovereign nation	empire
federation	nation	kingdom
dependent state	first level subdivision	country
bailiwick	autonomous municipality	canton
city-state	special city	colony
commissary	commonwealth	metropolitan area
independent city	county	department
dependency	parish (political)	province

region	shire	state
territory	external territory	unincorporated territory
overseas territory	possession	occupied territory
second level subdivision	diocese	archdiocese
parish (ecclesiastical)	third level subdivision	arrondissement
borough	precinct	township
inhabited place	locale	rural community
hamlet	village	town
city	American Indian reservation	Aboriginal reserve

Administrative entities include those defined by boundaries set up by standard, independent sovereign states. In addition, administrative entities in the TGN may include those with government and boundaries defined by other authorities, such as ecclesiastical or tribal authorities.

Administrative subdivisions and neighborhoods of the largest cities are within the scope of the TGN. Included are subdivisions such as urban districts, quarters, neighborhoods, or suburbs. Also within the scope of the TGN are large physical or man-made features within the boundaries of a city, including streams, hills, woods, or large parks. Smaller features typically found within the boundaries of a city are *outside* the scope of the TGN. Therefore, generally excluded from the TGN are buildings, landmarks, and streets within cities. Note that buildings are occasionally considered within the scope of the TGN; these are limited to structures that are located in the countryside (for example, abbeys, villas, and shopping centers), where they serve as a place name in the absence of a larger populated place.

Likewise, some other large, major man-made features are within the scope of the TGN, for example the Great Wall of China and the Via Appia.

Cultural and political groups are *outside* the scope of the TGN; however, the political state of a cultural or political group, and the territory within its boundaries, are *within* the TGN's scope. For example, the Ottoman Turks are outside the scope of the TGN, though the *Ottoman Empire* could be included in TGN.

The scope of the TGN may include places that are no longer extant. This includes deserted settlements, such as ghost towns and archaeological sites. Lost physical feature could include features such as submerged islands.

Within the scope of the TGN are included "general regions," which are recognized, named areas with undefined, controversial, or ambiguous borders. Examples include the Middle East, which refers to an area in SW Asia and NE Africa, but which has no defined borders and may be variously interpreted to mean different sets of nations. General regions are not organized political entities, though they may have their roots in historical entities such as kingdoms, duchies, or traditional tribal areas.

Historical entities may also be included; the number of historical entities will grow through contributions. Historical names and information regarding ***extant*** places may also be included in the TGN. For example, the record for the extant Turkish city of Pamukkale includes the historical name *Hierapolis*, as well as notes and dates describing the history of the city.

THE RECORD FOR EACH PLACE

The minimum record[10] for each place represented in the TGN includes a unique numeric identifier, a name for the place, at least one "place type," and a hierarchical position (indicated through a link to the immediate "parent"). Altogether, some 25 or so items of information may be associated with a place, and some of these data (including names, place types, and parents) are repeating. Some of the data may not be appropriate to display for end-users (see *Chapter 6: Recommendations for Using the TGN*).

Sample TGN "Record" Display for the End-User

(7000457)

- **Firenze** (Firenze province, Toscana Italia) (inhabited place)

Coordinates: Lat: 43 47 N Long: 011 15 E
(represented in degrees minutes direction)
Lat: 43.783 Long: 11.250
(represented in decimal degrees and fractions of degrees)

Note: Was Roman military center at head of navigation on Arno river & on Casian Way; escaped capture by Goths 5th cen.; was thriving center by 12th cen.; torn by medieval Guelph/Ghibelline civil strife; ruled by Medici family from 1434.

Hierarchical Position:

- Europe .....(continent)
- Italia .....(nation)
- Toscana .....(region)
- Firenze .....(province)
- Firenze .....(inhabited place)

Names:

Firenze (C,V, Pref)
Florence (C,O, prefEng)
Florencia (C,O)
Florenz (C,O)
Fiorenza (H,V).....medieval
Florentia (H,V).....name of Roman colony on N bank of Arno

Place Types:

inhabited place (C) .....site of ancient settlement, later founded as colony by Romans in 1st cen. BC, at foot of Etruscan hill town Fiesole
city (C)
regional capital (C)
provincial capital (C)
commune (administrative) (C)
river settlement (C) .....developed on both sides of the Arno river, is subject to periodic flooding; most bridges were destroyed in WWII
tourist center (C)
archiepiscopal see (C) .....bishops were established here early; today is famed for huge cathedral & baptistry & for numerous other churches
industrial center (C).....factories located in suburbs produce precision instruments & other items
cultural center (C) .....noted as great center of art & literature since Middle Ages, especially flourished 14th-16th cen.
transportation center (C)..... for road & river traffic since Roman times, now is also

[10] For the purposes of this chapter, "record" refers to an intellectual record — not necessarily a database record.

a major hub for rail traffic
craftsman center (C) .....famed for traditional products, including textiles, glass, ceramics, metal wares, leatherwork, art reproductions & furniture
educational center (C)
financial center (C)..... Florentines were paramount bankers in Europe by 15th cen.
capital (H) .....of duchy of Tuscany
municipium (H)

Sources:

Fiorenza ..... Companion Guide: Florence (1979), 14 [VP]
Firenze ..... Columbia Lippincott Gazetteer (1961) [BHA]
Webster's Geographical Dictionary (1984) [GRLPA]
Times Atlas of the World (1992), 66 [VP]
Companion Guide: Florence (1979), 62 ff. [FDA]
Florence..... Webster's Geographical Dictionary (1984) [FDA, GRLPA]
Canby, Historic Places (1984), I, 296 [VP]
Encyclopaedia Britannica (1988), IV, 838 [VP]
Webster's Geographical Dictionary (1988), 400 [VP]
Florencia .....Rand McNally Atlas (1994), I-56 [VP]
Cassell's Spanish Dictionary (1978), 317 [VP]
Florentia..... Princeton Encyclopedia (1979), 331 [GRLPA,VP]

Names

Place names (in relational files, *geog_name.place_name*) in the TGN are generally limited to established proper names; however, additional names may include documented nicknames, abbreviations, or codes (e.g., an ISO code). The TGN contains vernacular and English names of places, as well as variant names in other languages and historical names for some places. The adjectival form of names for those entities that are often used to represent the nationality of artists is also included (e.g., *Italian*, *Flemish*, or *Florentine*). The inclusion of multiple names that refer to the same place has always been one of the primary goals of the TGN, because this allows the vocabulary to be used as a retrieval tool that can provide access across databases that may have been indexed using any of the variant names. This is the critical difference between the TGN as a "vocabulary" and other sources designed to be used solely as "authorities."

Names for London, England

London (C,V)(preferred)
Londres (C,O)
Londra (C,O)
Lundin (H,V)
Lundenne (H,V)
Lundene (H,V)
Lundenburg (H,V)
Lundonia (H,V)
Londinium (H,V)
Lundinium (H,V)[11]

[11] For names, "C" indicates a current name, "H" a historical name, "V" that the language is vernacular (i.e., local, but transliterated where the language is written in a non-Roman alphabet), and "O" that the language is other than vernacular. End-users should have access to a key that explains what these flags mean.

Preferred Name

Although the TGN uses multiple names to refer to the same place, it was necessary to flag one name to identify to a computer system which name to use in the hierarchy, in a label for the place, or in alphabetical lists. Therefore, although all variant names provide access to the place record, one name is flagged as the "preferred name," also known as the "entry form" or "descriptor" (in relational files, *geog_name.place_name where geog_name.pref_flag = P*). Given that the audience for the TGN is international, the name flagged "preferred" is the name currently commonly used in the place (the "vernacular" or local name).[12] The preferred name in the example above is "London." Note that the vernacular name is not necessarily technically in the official language spoken in the nation; for example, *Los Angeles* is a Spanish phrase, but it is the vernacular name of a city in the United States where the common language (and the language used in official government transactions) is English.

Note that a place can have multiple vernacular names, as when multiple languages are spoken in a place (e.g., Switzerland). A place can have multiple vernacular names even when there is a single vernacular language (e.g., *Abydel* and *Abbey Dell* are both vernacular names for the same town in Orange county, Indiana, where the local language is English). There can be multiple vernacular names for a place when the vernacular language is not written in the Roman alphabet and variant transliterations exist (as in *Ashkhabad* and *Aşhabad*, Turkmenistan). Where it is difficult or inappropriate to indicate a vernacular name (as with the names of continents or oceans), the preferred name is English by default.

Because the preferred name is meant to be used in displays, length is a practical consideration. Therefore, the preferred name is not necessarily the "official name," but rather the commonly used vernacular name (e.g., for Jordan, *al-Urdunn* is preferred, not the official *al-Mamlakah al-Urdunnīyah al-Hāshimīyah*, Hashemite Kingdom of Jordan).

The Vocabulary Program has attempted to indicate the correct current vernacular name for places; however, owing to the vast number of names in the TGN database, editors could not research each name. Editors have relied upon names published in atlases and encyclopedias whose editorial policy preferred the vernacular name; given the rapid pace of geopolitical changes and the lag-time involved in publications listing them, it is possible that in some instances the newest, current vernacular name may not be correctly indicated.

The official name and other variants are included in the record for the place and can be used for access, even though they are not the "preferred" name. The place record may contain alternate names that include all variations in spelling, names in different languages, nicknames, and historical names. If there is a commonly used English name for the place, this is flagged (in the relational files, *where geog_name.other_flags = ENG*) and can be used for displays instead of the so-called "preferred name." See an example of an alternate English display in the section on *The Hierarchy*.

Variant Names

Variant (or alternate) names may include all variations in spelling (including differences in diacritical marks, punctuation, or capitalization),[13] names in different languages, nicknames, and historical names. Certain types of names are flagged in the data, and the flags should be available to end-users. Where the flag or some other abbreviation is displayed to end-users, they should have access to a key explaining what

[12] For names in languages that do not use the Roman alphabet, the names have been transliterated.

[13] If a name appears in a well-known, published source, it is a valid alternate name for that place in TGN, even if it is a misspelling.

the flags mean. In the examples below, current names are flagged "C" and historical names are flagged "H" (in the relational files, *where geog_name.age_flag = C or H*). Vernacular names are flagged "V" (in relational files, *where geog_name.lang = V*).[14] Names in languages other than the vernacular are flagged "O" (for "other language," in relational files, *where geog_name.lang = O*).[15]

Names may also have other flags. The name flagged "ENG" is the commonly used English name for the place (in relational files, *where geog_name.other_flags = ENG*, indicated as ***preferredEng*** for the end-user in the sample display below). "ISO" indicates the three-letter ISO code for nations (in relational files, *where geog_name.other_flags = ISO*, indicated as ***ISO*** in sample display below).[16] "USP" indicates the two-letter U.S. Postal Service code for states in the USA (in relational files, *where geog_name.other_flags = USP*). "ADJE" is the preferred English adjectival form of the name.

Names and flags for nation of Jordan

al-Urdunn *(C, V, Preferred)*
Al-Urdunn *(C,V)*
Al Urdunn *(C,V)*
Urdunn *(C,V)*
al-Mamlakah al-Urdunnīah al-Hāshimīyah *(C,V)*
Al Mamlakah al Urdunniya al Hashemiya *(C,V)*
Jordan *(C, O, preferred Eng)*
Hashemite Kingdom of Jordan *(C,O)*
Jordanie *(C,O)*
Royaume hachémite de Jordanie *(C,O)*
Jordanië *(C,O)*
Jordania *(C,O)*
Jordânia *(C,O)*
Giordano *(C,O)*
JOR *(C, O, ISO)*
Transjordan *(H,O)*
Jordanian *(C,O, English adjectival)*

"DIS" indicates the display form of the name (in relational files, *where geog_name.other_flags = DIS*), used in strings of names when the place is presented as a broader context, but when using its preferred name would result in ambiguity or confusion. This usually occurs when an administrative subdivision has the same name as an inhabited place. For example, Tokushima the inhabited place and Tokushima the prefecture (in Japan) have the same preferred name. Therefore, in horizontal displays of the hierarchical string, the prefecture should be represented by its display name (***Tokushima prefecture***) in order to avoid confusing it with the inhabited place of the same name (see example below).

[14] Each place has only one name tagged as preferred (in relational files, *where geog_name.pref_flag = P*), and that name generally is a vernacular name. However, each place may have more than one vernacular name, so every name flagged as "vernacular" is not necessarily the preferred name.

[15] In the future, TGN will have added structure to allow an indication of the specific language of the name.

[16] International Organization for Standardization. *Codes for the Representation of Names of Countries and Their Subdivisions*Genève: International Organization for Standardization, 1999. The Vocabulary Program also used various other ISO publications and newsletters.

Label with "display name" for Tokushima prefecture

Tokushima (*Tokushima prefecture, Shikoku region, Nihon*) (inhabited place)

Alphabet

Names and all other information in the TGN records are expressed in the Latin alphabet. The sources used in the TGN were generally written in Western European languages; therefore, names in languages that use other alphabets were already transliterated before they were incorporated into the database. In those records edited by the Vocabulary Program, variant names derived by various transliteration standards were included, where possible.

Diacritics in the Relational Files and REC formats are represented by numeric codes in the TGN data; a translation of the codes and diacritical marks is located in *Appendix A* of this manual. The Vocabulary Program has used these codes rather than diacritical marks because existing international standards for diacritics pose technical obstacles for many of our contributors and users.[17] The TGN USMARC format uses ANSEL for diacritics (the standard ANSI Z39.47).

Names may be alphabetized using the normalized version of the name (in the relational files, *geog_name.sort_name*), which is stripped of all diacritical codes, punctuation, spaces, and translated to all uppercase letters.

Sequence of Names

In a full record for a place, names are arranged in a logical order. The sort order (in relational tables, *geog_name.seqn*) should allow names to be arranged with the preferred name at the top of the record, and variant names arranged with current names before historical names. Note that historical names are arranged in reverse chronological order, from the most recent to most ancient names. Within a set of names used during the same period, the vernacular names precede the names in other languages.

In some cases, the TGN describes the time frame when a given place name was used. This is described in a "display date" (in relational tables, *name_detail.display_date*), which is a short note describing when the name was used; other information pertinent to the name may be included. The database also includes associated years that represent the beginning and end of the span indicated in the display date. These "*start*" and "*end*" dates (in relational tables, *name_detail.start_date* and *name_detail.end_date*) can be used in retrieval, but should be hidden from end-users.

[17] It is possible that TGN will adopt the Unicode Standard (ISO 10646) in the future.

Sequence, history/language flags, and dates for names, Siena, Tuscany, Italy

0 Siena (C,V) used since 13th cen. ***(start: 1200, end: 9999[18])***
1 Sienna (C,O)
2 Senae (H,V) medieval ***(start: 800, end: 1500)***
3 Sanna (H,V)
4 Saena Julia (H,V) Roman (*start: -100, end: 300)*
5 Sena Julia (H,V)
6 Sena (H,V) Etruscan ***(start: -800, end: 100)***

Dates for place names may be known with different levels of specificity and varying shades of certainty. In the display date, editors express exact and imprecise dates based on what information is available. In the start and end date, the editors can record a known or estimated time span in years, based on what is known. Since conventions used to describe approximate dates (e.g., "circa") may vary depending upon context, this gives editors flexibility in establishing appropriate date spans for retrieval.

Display dates for names, with associated start and end dates

(*for town in New York state, USA, exact day of name change is known*)
Sleepy Hollow **new name certified by village trustees on Dec. 11, 1996** ***(start: 1996, end: 9999)***

(*for US state, estimated searching dates for "circa")*
Indiana **named by land developers ca. 1765, roughly means "land of the [American] Indian"** ***(start: 1750, end: 9999)***

(*for variant name for Thebes, Egypt, estimated searching dates for "century"*)
Nowe **used from 22nd cen. BC** ***(start: -2100, end: -29)***

(*for variant for Siena, Italy, searching dates informed by life dates of Julius Caesar)*
Saena Julia **Roman** ***(start: -100, end: 300)***

(*for variant for Vienna, Austria, estimated searching dates express broad span of time*)
Vindobona **ancient** ***(start: -400, end: 1500)***

[18] TGN records an end date "9999" to indicate that the name is still used. Dates BCE (before current era) are indicated with negative numbers (i.e., using a minus sign).

Contributors

Contributors of names to the TGN are noted by their initials or an abbreviated form of their name (in relational files, *name_detail.contrib*). End-users should have access to a key that lists the full name of the Getty project and the initials (see ***Appendix B*** for description of these contributors).

Bibliography of the History of Art	**BHA**
Foundation for Documents of Architecture	**FDA**
Getty Research Library Photo Archive	**GRLPA**
Vocabulary Program	**VP**

Note that where a contributor's data was received in hard-copy and entered by hand by the Vocabulary Program, the contributor will be listed as "VP," and the project or institution that supplied the data will be listed as a Source (see *Bibliography* below).

Bibliography

In most cases, contributors and the Vocabulary Program editors have listed the sources used for names (in relational files, table = *BIBLIO*). Most names in the TGN were derived from standard general reference sources, including atlases, loose maps, gazetteers, guidebooks, geographic dictionaries, and encyclopedias. Web sites, including NIMA,[19] provided many names. Many names in the USA are from USGS.[20] Other sources include books on the history of art and architecture, journal articles, newspaper articles, newsletters from ISO and the United Nations, letters and telephone calls to embassies, inscriptions on art objects, and catalog records of repositories of art objects.

[19] United States National Imagery and Mapping Agency. GEOnet Names Server [online database]. Edited by the U.S. Board on Geographhic Names. Washington, DC: National Imagery and Mapping Agency, 2000.

[20] United States Department of the Interior and U.S. Geological Survey. *Geographic Names Information System Digital Gazetteer*. Reston, Virginia: U.S. Geological Survey, 1994.

Contributors and bibliographic sources for names, Beijing, China

0 Beijing
BHA,[21] Columbia Encyclopedia (1975)
Encyclopædia Britannica (1976)
GRLPA, Webster's Geographical Dictionary (1984) 936
VP, Times Atlas of the World (1992) 24

1 Peking
BHA, Webster's Geographical Dictionary (1984) 936
VP, US National Imagery and Mapping Agency, GEOnet Names Server (1996)

2 Pei-Ching
VP, US National Imagery and Mapping Agency, GEOnet Names Server (1996)

3 Pékin
BHA
FDA, Cassell's French Dictionary (1973) 550
VP, US National Imagery and Mapping Agency, GEOnet Names Server (1996)

4 Pequín
VP, Cassell's Spanish Dictionary (1978) 917

5 Pechino
VP, Cassell's Italian Dictionary (1983) 365

6 Peip'ing
VP, US National Imagery and Mapping Agency, GEOnet Names Server (1996)

7 Ch'i
VP, Webster's Geographical Dictionary (1984) 936

8 Khanbalik
VP, Times Atlas of World History (1994) 347

9 Cambaluc
VP, Canby, Historic Places (1984) II, 721-722

The sources for each name are generally cited for the end user by the brief citation (in relational files, *biblio.bib_brief*). A key listing the full citation for these brief references should be available to end-users (in relational files, *biblio.bib_brief* and *biblio.citation* for each *bib_key*).

[21] Contributors in this example: BHA, Bibliography of the History of Art; FDA, Foundation for Documents of Architecture; GRLPA, Getty Research Library Photo Archive; VP, Vocabulary Program. Note that users will also need access to a key that lists the contributor initials with the corresponding full name for the contributing project (see *Appendix B*).

Latitude and Longitude

Most of the place records in the TGN include geographic coordinates (in relational files, *geog_main.lat, geog_main.lat_dir, geog_main.long*, and *geog_main.long_dir*). Latitude is the angular distance north or south of the equator, measured along a meridian. Longitude is the angular distance east or west of the Prime Meridian at Greenwich, England. Latitude and longitude are expressed in degrees and minutes.[22] Geographic coordinates represent a single point for each place, corresponding to a point in or near the center of the inhabited place, political entity, or physical feature. For linear features such as rivers, the point represents the source of the feature. The Vocabulary Program editors entered coordinates using a variety of sources.[23]

Coordinates for state of Indiana, USA

Latitude: 40° 00' N
Longitude: 086° 15' W

The TGN also includes coordinate data expressed as decimal degrees (in relational files, *geog_main.lat_num* and *geog_main.long_num*), where the minutes of latitude and longitude are converted to decimal fractions of degrees (e.g., *40.000* and *-086.250* are the decimal degrees for the example of Indiana above). Coordinates south of the equator and west of the Prime Meridian are expressed as negative numbers. *See Chapter 5: Recommendations for Using the TGN* for further discussion of decimal degrees.

Descriptive Note

For some records, a note has been included to describe the place (in relational files, *geog_main.descr.note*). This descriptive note may provide additional information about the place, or expand on information in other fields. Topics that may be covered in this note include the political history of the place, a physical description, and its importance relative to other places or to the history of art and architecture.

Descriptive note for state of Indiana, USA

Originally inhabited by Mound Builders; when Europeans arrived, Algonquian-speaking Miami confederation, Potowatomi & Delaware lived here; claimed by French in 17th cen., by British in 1763; taken by USA in Revolution; was strongly pro-Union in Civil War.

[22] When the new system is implemented in 2001, TGN will be able to capture four bounding coordinates.

[23] Major sources include United States Department of the Interior and U.S. Geological Survey. *Geographic Names Information System Digital Gazetteer.* Reston, Virginia: U.S. Geological Survey, 1994; United States National Imagery and Mapping Agency. GEOnet Names Server [online database]. Edited by the U.S. Board on Geographic Names. Washington, DC: National Imagery and Mapping Agency, 1996-2000; *Times Atlas of the World*. 9th ed. New York: Times Books, 1994; Rand McNally and Company. *The New International Atlas*. Chicago: Rand McNally, 1994; and *Cambridge World Gazetteer: A Geographical Dictionary*. Cambridge, England; New York: Cambridge University Press, 1990.

The sources for descriptive notes and all other information are linked indirectly through a name.[24]

Descriptive notes and display dates are written in English. However, in order to make the information understood more easily by an international audience, the Vocabulary Program editors have rarely used abbreviated words. Abbreviations occur only when the full word can be understood easily by virtue of its context in the note (e.g., preceding or following a date); they include the following abbreviated words: "ca." for "circa," "cen." for "century," "mill." for "millennium," and "&" for "and."

Place Types

Geographic places in the TGN can be either physical or political entities. Physical features include continents, rivers, and mountains; political entities include empires, nations, states, districts, townships, cities, and neighborhoods. The "place type"[25] in the TGN is a term that characterizes a significant aspect of the place, including its role, function, political anatomy, size, or physical characteristics.

Place types, history flags, and display dates for Marrakesh, Morocco

inhabited place (C, Pref) ***founded by Yusuf ibn-Tashfin in 1062***
city (C) ***modern European city founded 1913***
provincial capital (C)
commercial center (C)
transportation center (C)
religious center (C) ***important medieval center of Islam***
royal residence (H)
capital (H) ***of Almoravid dynasty, until 1147; of Almohades from 1147; of Morocco, 1550-1660***

Preferred Place Type

One place type in each record is flagged "preferred" (in relational files, *where geog_placetype.pref_flag = P*). As with preferred names, a preferred place type is necessary because one place type must be distinguished to appear with the preferred name in an identification of the place (as in the hierarchy or in lists).

[24] In the current TGN data model, sources are linked only to the name; therefore, to deal with this limitation, editors cited sources for other information indirectly by linking the source to a name (generally the entry form name in the published source). In the future, sources will be linked directly to the data.

[25] In relational files, *geog_placetype.placetype_code* links the place type to the *geog_key;* the code is linked to the term by *placetype_list.placetype_code* and *placetype_list.placetype.*

The preferred place type for cities, towns, and villages is "inhabited place"; for places that are no longer inhabited, the preferred place type is "deserted settlement."[26] The preferred place type for sovereign nations is "nation"; it is "dependent state" for semi-autonomous states.[27] The preferred place types of subdivisions within a nation generally are the terms used by the nation (or English translations of the terms). The preferred place types of physical features are terms that describe the individual feature.

Each record has a preferred place type, and for some records, additional place type terms have been included. The Vocabulary Program editorial policy was to add place type terms that describe the most important qualities and roles of the place.

Sort Order and Dates for Place Types

As is true of names, place types also are ordered logically by a sequence number (in relational files, *geog_placetype.seqn*). The order generally represents order of importance, with the preferred place type first, and others listed with current before historical, and then in order of importance.

For place types, "C" indicates a current place type, "H" a historical place type that no longer applies (in relational files, *where geog_placetype.age_flag = C or H*). Occasionally, when a place type has applied for a period in history, then does not apply for a period, but again currently applies, it may be flagged "B" (for "both historical and current").[28] A key explaining these flags should be accessible to end-users.

As with names, place types may have associated display dates (in relational files, *placetype_dates.display_date*), which are indexed with years representing the start and end of the span (in relational files, *placetype_dates.start_date* and *placetype_dates.end_date*). These "*start*" and "*end*" dates should not be visible to the end-user.

[26] Note that occasionally the precise location of an abandoned settlement is unknown (e.g., Gath, Israel), and is labeled by the place type "lost settlement." Abandoned structures that were not actually inhabited (e.g., Stonehenge, England) are labeled "ruins."

[27] Excludes territories (1) to which Antarctic Treaty is applicable in whole or in part, (2) without permanent civilian population, (3) without internationally recognized civilian government, or (4) representing unadjudicated unilateral or multilateral territorial claims. *2000 Encyclopaedia Britannica Book of the Year. Events of 1999.* Chicago: Encyclopaedia Britannica Inc., 2000.

[28] The date is a repeating set of fields in the TGN database, and thus a place type (or a name) may have multiple ranges of time associated with it; therefore it may be *both* current and historical.

Place types, sort order, flags, display dates, and retrieval dates for Paris, France

0 inhabited place (C, Pref) (settled as Gallic (Parisi tribe) fishing village on Îsle de la Cité) ***(start: -800, end: 9999)***
1 city (C)
2 national capital (C) (since accession of Hugh Capet to French throne in 987) ***(start: 987, end: 9999)***
3 department capital (C)
4 river port (C)
5 financial center (C)
6 commercial center (C)
7 transportation center (C)
8 cultural center (C)
9 industrial center (C)
10 tourist center (C)
11 trade center (C)
12 educational center (C)
13 episcopal see (C) (since 3rd cen.) ***(start: 200, end: 9999)***
14 countship (H) (Carolingian) ***(start: 700, end: 1000)***

THE HIERARCHY

The "hierarchy" in the TGN refers to the method of structuring and displaying the inhabited places within their broader contexts. The name of a city alone does not identify it; the city exists within the larger context of a state, and the state in turn exists within the larger context of the nation. The name "Columbus" attached to the location of an art work does not identify its location. Even adding the nation, "United States of America," does not identify the location, because in the USA there are more than a dozen Columbuses. One would have to add the level of state — and even county — to make the location clear, saying, for example, that Henry Moore's sculpture, "Large Arch," is located in "Columbus, Bartholomew county, Indiana, United States."[29]

Hierarchical display of Columbus, Indiana

World
North and Central America (continent)
United States (nation)
Indiana (state)
Bartholomew (county)
Columbus (inhabited place)

View of the Hierarchy

The 900,000 places in the TGN are arranged in hierarchies to represent the modern world. Historical places are also included. The top of the TGN hierarchy is the World, and all continents and all nations are represented below.

To display hierarchies of the TGN, indentation is generally used to indicate part/whole relationships—that is, broader and narrower contexts. For example, all inhabited places that are indented one level beneath a region are part of that region (sometimes referred to as "children" of a "parent," in relational files, link of *geog_parents.geog_key* to *geog_parent.parent_key*).

Since the TGN hierarchy has too many levels to display on a monitor simultaneously, the display should ideally show only the first level below the target place, and as many levels above it as possible. In the example below, an ellipsis (...) indicates that there are additional places beneath a given level in the hierarchy.

[29] The preferred name of the place with its full string of parents' names and its preferred place type usually uniquely identifies the place. However, this is not true in all cases, because sometimes there is more than one inhabited place with the same preferred name in the same administrative district; such places may be distinguished by their geographic coordinates. In an automated system, the numeric keys (in relational files, *geog_key*) are unique identifiers of places.

Levels of hierarchy in display, with "Bénin" as target

World
Africa (continent)
Bénin (nation)
Alibori (river)
Atakora (province) ...
Atlantique (province) ...
Borgou (province) ...
Kouffo (river)
Mono (province)...
Ouémé (province) ...
Ouémé (river)
Zou (province) ...

Places are generally represented in the hierarchy by their "preferred" or entry-form name (generally the vernacular, in relational files, *geog_name where geog_name.pref_flag = P*) and a "preferred" place type *(*in relational files, *placetype_list.placetype* linked to place by *geog_placetype.placetype_code where geog_placetype.pref_flag = P*). For continents and oceans, the default preferred name is English. See section on *Preferred Names* for further discussion.

Partial hierarchical display for island of Rhodes, Greece, using preferred names (generally vernacular)

World
Europe (continent)
Ellás (nation)
Aiyaíon (region)
Sporádhes (department)
Ródos (island)
Attáviros (mountain)
Kámiros (deserted settlement)
Líndhos (deserted settlement)
Afándou (inhabited place)
Apolakkiá (inhabited place)
Arkhangélos (inhabited place)
Anithi (inhabited place)
Asklepios (inhabited place)
Kalavárdha (inhabited place)
Kattaviá (inhabited place)
Prasonísi, Ákra (cape)
Líndhos (inhabited place)
Plimiri (inhabited place)
Ródhos (inhabited place)...
Triánda (inhabited place)
Vroulia (inhabited place)

English View of the Hierarchy

In addition to the so-called "preferred" name (usually in the vernacular, in relational files, *geog_name.place_name where geog_name.pref_flag = P*), the TGN also flags a preferred English name for the place where appropriate (in relational files, *where geog_name.other_flags = ENG*). The hierarchy could be displayed using the English names instead of the "vernacular" names.

Note, however, that most places do not have a commonly used English name, and the preferred (vernacular) name must be used.[30] In the example below, the nation Greece and the island Rhodes are displayed with their preferred English names, but many places in this section of the hierarchy must still be displayed with their preferred vernacular names.

Partial hierarchical display for island of Rhodes, Greece, using preferred English names

World
Europe (continent)
Greece (nation)
Aegean Islands (region)
Sporades (department)
Rhodes (island)
Attáviros (mountain)
Kámiros (deserted settlement)
Líndhos (deserted settlement)
Afándou (inhabited place)
Apolakkiá (inhabited place)
Arkhangélos (inhabited place)
Arnithi (inhabited place)
Asklepios (inhabited place)
Kalavárdha (inhabited place)
Kattaviá (inhabited place)
Lindos (inhabited place)
Prasonísi, Ákra (cape)
Plimiri (inhabited place)
Rhodes (inhabited place) ...
Trianda (inhabited place)
Vroulia (inhabited place)

[30] In relational files, an algorithm would look for *geog_name.place_name where geog_name.geog_name.other_flags = ENG*, and where there is none, would default to *geog_name.place_name where geog_name.pref_flag = P*.

Islands in the Hierarchy

Inhabited places and physical features that are located on islands usually have been positioned under the island in the TGN hierarchy. That is, if places are located on an island, the island will generally form a level in the hierarchy. This was necessary for island nations, and the rule was extended to include all islands in the database.[31]

Note that, in this edition of the TGN, a similar rule was not applied to other physical features such as peninsulas, nor to general regions that have no defined administrative boundaries.[32] Therefore, although the record for a general region, such as the Middle East, may appear in the TGN, the places that are located inside this area are not linked to it. For some historical entities, "children" may not be linked to the "parent."

Physical and Political Geography

While we are used to seeing the physical and political world coexist on maps, some complexity is added when these unlike entities are merged into a single hierarchy in the TGN. Though it is useful for retrieval and for some displays to have all entities living together in the same hierarchy, it may also be desirable to be able to display the political hierarchy without physical features.

Place types [33] help describe and index places. Place types in the TGN are derived from a controlled list based on terms in the AAT. The TGN took advantage of the AAT's hierarchical structure, which groups similar things together, and enhanced it to allow users to sort the political world separately from the physical world. This feature could also be useful when the TGN is utilized in a retrieval tool.

Section of the TGN place type list

20000 <natural physical world>
21000 <landscapes>
21100 <bodies of water & components>
21101 body of water
21102 ocean
21103 sea
21104 pond
21105 river
21106 tributary (river)
21107 stream
21108 creek
21109 brook
21110 watercourse
21111 aquifer
21112 lakes
21113 lake
21118 intermittent lake

[31] See also section on *Administrative Subdivisions* below.

[32] This was due to constraints of time and the difficult arbitrary decisions required in defining the boundaries of these entities.

[33] Place types are linked to the place record by (in relational files) *geog_placetype.placetype_code*; the term associated with the code is found in *placetype_list.placetype.*

The TGN's list of controlled vocabulary for place types is arranged in an implied hierarchy, meaning that related things are identified by having sequential numeric codes. For example, all physical features have a code within the range 20000-29999, and all flowing bodies of water have a code within the range 21105-21110. Therefore, all places with a place type code within the range 20000-29999 could be excluded from a political hierarchy (programmers would need to analyze how to construct the resulting hierarchical relationships, and account for holes in the hierarchy). Also, all places with a place type code within the range 21105-21110 could be included in a retrieval of rivers and other flowing bodies of water.

While this coded system of place types helps to separate the physical from the political world, additional methods may be employed to pull out the political hierarchy.

Administrative Subdivisions

The main subdivisions in the political world include continents, nations, and main political subdivisions within nations (e.g., states, regions, provinces, etc.). Whenever possible, every nation has at least one level of subdivision, and some nations have two (e.g., the USA has the levels of state and county). However, the level of these subdivisions is not consistent across the hierarchy: That is, you cannot count on finding the first-level internal subdivision directly under the nation because there could be intervening levels representing the physical world. Likewise, inhabited places are not necessarily directly under an administrative subdivision and not always on the same level, even within a single nation (note the inhabited places under Pate Island in the example below).

Hierarchy with physical feature as intervening level between administrative levels and inhabited place, Coast province, Kenya

Africa (continent)
- **Kenya (nation)**
  - **Coast (province)**
    - **Bura (inhabited place)**
    - **Galana (river)**
    - **Gazi (inhabited place)**
    - **Kinango (inhabited place)**
    - **Lamu (inhabited place)**
    - **Lugards Falls (waterfalls)**
    - **Mkunumbi (inhabited place)**
    - **Pate Island (island)**
      - **Pate (inhabited place)**
      - **Rasini (inhabited place)**
    - **Rabai (inhabited place)**
    - **Shimoni (inhabited place)**
    - **Tana (river)**

To further complicate the situation, the first-level subdivisions are not labeled with a consistent preferred place type across the database, because it was inappropriate to use a consistent term. For example, the first-level administrative subdivision is called "state" in the United States, "region" in Italy, "province" in Kenya, etc. In addition, the same term may have different significance in different nations. For example,

the "province" is the first-level subdivision in Kenya, but "province" is the second-level subdivision (below "region") in Italy.

To clarify the levels, the TGN adds place type terms "primary political unit," "first-level subdivision," and "second-level subdivision" that appear as non-preferred place type terms in place records. In the above example, the full record for Kenya contains a place type code "81002" for "primary political unit," and the full record for Coast province contains a place type code "81100" for "first-level subdivision." This would allow implementors to build hierarchies of only the political entities. It also meets a critical need for users who have simple systems that require a set number of levels, generally "primary political unit" (i.e., nations and dependent states), "first-level subdivision," optional "second-level subdivision," and "inhabited place." Note that, in order to construct such a hierarchy, analysis would have to be done to determine how to skip over unwanted hierarchical levels and how to deal with resulting holes in the TGN hierarchies. (The Getty will not aid in this analysis.)

Generally, the hierarchy in the TGN goes only to the level of the inhabited place. However, the level of neighborhood has been included for some of the world's largest cities. Releases of the TGN will generally not include records for streets or buildings within inhabited places.[34]

Partial hierarchy illustrating level below inhabited place, Rome, Italy

World
Europe (continent)
Italia (nation)
Lazio (region)
Roma (province)
Roma (inhabited place)
Borgo (neighborhood)
Castel Romano (neighborhood)
Trastevere (neighborhood)
Acilia (rione)
Borgata Borghesina (rione)
Borgata Fidene (rione)
Capannelle (rione)
Casalotti (rione)
Campo Marzio (field)
Aventino (hill)
Campidoglio (hill)

[34] Although the data structure would allow the addition of buildings, an issue arises with this suggestion, because the attributes of buildings are different from those of places (e.g., architects, construction materials, etc. are appropriate attributes for buildings). Logically, records for architectural structures would be compiled in a separate vocabulary, albeit ideally linked to the geographic vocabulary. Currently there is an occasional building in the release version of the TGN hierarchy, when the building is located in the countryside outside any city and serves as a place name for the area.

Urban Expansion

The TGN attempts to accommodate the expansion and diminishment of urban areas. For example, it is often true that modern cities have incorporated formerly independent towns. When possible, the names of these formerly independent places are included as neighborhoods in the TGN (e.g., Greenwich Village and Harlem, now neighborhoods in New York City).

Hierarchy with formerly independent villages Greenwich Village and Harlem as neighborhoods in modern city

World
North and Central America (continent)
United States (nation)
New York (state)
Kings (county) [35]
New York (inhabited place)
Manhattan Island (borough)
Chinatown (neighborhood)
Greenwich Village (neighborhood)
Harlem (neighborhood)
Washington Heights (neighborhood)
Central Park (park)
Franklin D. Roosevelt Island (island)
Wards Island (island)

Also, a large city may be deserted and subsequently occupied by several smaller towns (e.g., Thebes, Egypt). In the example below, the villages on the site of Thebes would have as preferred parent the governorate by which they are administered. However, in order to allow retrieval down the hierarchical path of Thebes, the villages have the site of the deserted settlement Thebes as an alternate parent (indicated in the example with "N," for "non-preferred parent"; see *Multiple Parents* below).

Hierarchy with villages on site of ancient Thebes, Egypt, where Thebes is non-preferred parent

World
Africa (continent)
Mīsr (nation)
As-Saʿīd (region)
Qinā (governorate)
Thèbes (deserted settlement)
Al-Karnak (inhabited place) [N]
Al-Uqsur (inhabited place) [N]
Qurna (inhabited place) [N]

[35] Note that New York City is linked to five counties (see *Multiple Parents* below). In this display, only one, King's County, is displaying (i.e., Manhattan appears as though it is not displaying with its "correct" grandparent).

Special Cases in Hierarchical Position

The TGN editorial policy is to publish places only when it is possible to determine their correct hierarchical position. However, in certain cases it is considered appropriate to publish the places even when this is not true – for example, when information about the internal subdivisions of a nation is unavailable at the time of release. In such cases, the inhabited places for that nation appear under a place-saver level (marked "lost & found/...") under the nation.[36]

Partial hierarchy with inhabited places under level marked "lost & found"

World
Asia (continent)
Al-Yaman (nation)
North Yemen (region)
Al-Bayda' (governorate) ...
Al-Hudayadah (governorate) ...
lost & found/North Yemen
Al-Fazah (inhabited place)
Al-Luhayyah (inhabited place)
Al-Mukha (inhabited place)

In the TGN, if a physical feature crosses a boundary, it is placed under the next highest level in the hierarchy. In other words, the river or mountain range is placed under the level of the hierarchy that entirely contains it. For example, the Amazon river crosses national boundaries, so it is placed under the next highest level, the continent of South America.

Physical feature crosses national borders, Amazon river

World
South America (continent)
Acre (river)
Altiplano (upland)
Amazonas (river)
Argentina (nation) ...

[36] In this example, though the capital of each governorate was known and placed in its correct hierarchical position, a map of Yemen was not available detailing to which governorate the other inhabited places belonged. Therefore, they are placed under "lost & found/North Yemen" and "lost & found/South Yemen" in this release of TGN.

Multiple Parents

The TGN does not resolve the issue of spanning borders in the same way for inhabited places as for physical features, because recording the hierarchical position of inhabited places is a higher priority; unusual or complicated relationships of inhabited places to their "parents" are accommodated by the polyhierarchy. If the area of an inhabited place crosses administrative boundaries (as may happen in the USA with a city belonging to two counties), the inhabited place appears under both administrative subdivisions. Likewise, if jurisdiction over an area is disputed between two nations, that area may appear as part of both nations — it may have "multiple parents."

Places in the TGN may be linked to multiple places as broader contexts. In many thesauri, each item may have only a single parent. However, with geographic places, it is necessary for some places to have multiple parents. For example, a dependent state such as Bermuda may be physically located in North and Central America; however, it is politically associated with the United Kingdom. To allow access to Bermuda via either hierarchical path, it must have two parents. To allow the system to have a default hierarchical view for lists and other displays and reports, one parent is indicated as "preferred."[37] In the TGN, Bermuda's preferred position in the hierarchy represents its physical location. In Bermuda's alternate hierarchical view it displays under the United Kingdom. This alternate view is indicated with an uppercase "N" (indicating "non-preferred") in the hierarchy display below.[38]

Dependent state has multiple parents, Bermuda

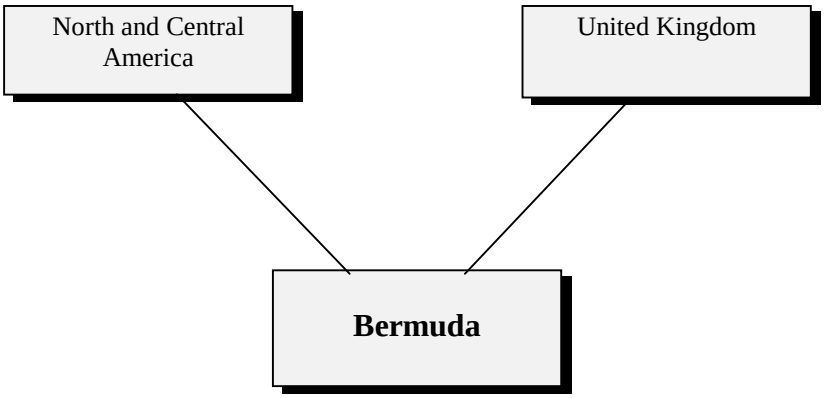
World
North and Central America (continent)
Anguilla (dependent state) ...
Bahamas (nation) ...
Belize (nation) ...
Bermuda (dependent state) ...

World
Europe (continent)
United Kingdom (nation)
Anguilla (dependent state)[N] ...
Bermuda (dependent state)[N] ...
England (country) ...

[37] The main (so-called "preferred") position of dependent states in the hierarchy is with their physical location (generally a continent). The main position of colonies, provinces, and other such entities that have less autonomy is with their political "parent." When jurisdiction over an area is disputed between two nations, the preferred parent is the nation generally accepted by the international community, as reflected in major published atlases and encyclopedias. However, these places should be retrieved equally well regardless of whether the preferred or non-preferred "parent" is used in a search. See also the display problem noted in the foonote for the "New York" example under *Urban Expansion* above.

[38] That is, in relational files, for each *geog_key*, there is only one *geog_parents.parent_key where pref_flag = P*, but in addition there may be one or more *geog_parents.parent_key where pref_flag = N*.

Another way of displaying the multiple parents of Bermuda



Inhabited places on islands also occasionally required the assignment of multiple parents. This happened when the island does not correspond to (or was not contained within) borders of the administrative region, for example when the island is divided between two administrative regions that extend onto the mainland.

CHAPTER 6
RECOMMENDATIONS FOR USING THE TGN

OVERVIEW

The Getty Vocabulary Program application of the TGN may be viewed as a "browser" at http://www.getty.edu/research/tools/vocabulary/. Views of the hierarchy and of the full records will be particularly useful as examples to implementors.

The TGN may be used as a source of vocabulary at the point of data entry and as an aid to retrieval. When used to inform the use of terminology for data entry, the TGN may be used as an authority if the cataloger or indexer consistently uses the same form of the TGN term to refer to a given place; the vernacular name and the English name are flagged and could be used to control terminology. That is, the TGN may be used as an authority by 1) always choosing the name flagged "preferred", which is generally a vernacular name, or 2) choosing the "preferred English" name (in relational files, *geog_name.place_name where other_flags = ENG*) where there is one and, where there is no "preferred English" name (which is true of the majority of the places in the world), default to the "preferred" name. The preferred name may also be used for implementors who need to always have one name to represent the place in hierarchies and other displays. However, even though one name is flagged "preferred" (in relational files, *geog_name.place_name where pref_flag = P*) for technical reasons, it is not expected that all users will want to use that name for their local purposes.

One of the most valuable uses of the TGN can be as a filter or search assistant for querying large, disparate data sets. The hierarchical structure and variant names can provide enhanced access to disparate databases where various names may have been used to represent the same place. Use of the TGN in such an application requires a license; see http://www.getty.edu/research/tools/vocabulary/, and follow the links for "Licenses and Sample Data."

The logical focus of the TGN is a place (not a "name"). Places have attributes, including names. Information about a place is linked across various tables by the unique numeric identifier for the place, in relational files, the *geog_key*. The data associated with the *geog_key* makes up the intellectual "record" for a place.

PROVIDING ACCESS TO THE DATA

It is advised that the TGN place records be accessible through any of the names associated with the place, preferred or variant. Furthermore, note that any name may be accessible in three forms, as illustrated in the example below: The exact name string (in relational files, *geog_name.place_name*), the normalized "sort name" (with case differentiation, spaces, diacritics, and punctuation removed, in relational files, *geog_name.sort_name*), and through a keyword table (where individual words of names have been parsed and normalized, in relational files, *name_keywords.keyword*). These various forms of the name are linked by the *placename_id* (in relational files) for that name, and to the place record by the *geog_key* Generally, the normalized name and normalized keywords are hidden from the end-user.

EXAMPLE:

***exact name string*: Rancho San Diego**

***sort name*: RANCHOSANDIEGO**

***keywords*: RANCHO**
SAN
DIEGO

Where commas appear in TGN names (i.e., with inverted names), it is useful to use the comma as a pivot to create additional access points. For example, if a TGN name is "Etna, Mount", a useful access point will be "Mount Etna" (normalized "MOUNTETNA," and created by using the comma as a cue to recombine words in natural order). Creating such strings should be done in addition to allowing Boolean operators on key words, as described below.

Querying

The TGN data should be accessible by both simple and advanced searches. For more advanced users, querying by name will be enhanced if users may utilize right-hand truncation, when searching for both the full name string or for keywords. For example, querying for "BODA*" (where the asterisk is a wildcard) would bring back dozens of place records in the TGN, including the following:

Boda... (Africa, République centrafri, Lobaye) (inhabited place)
Boda... (Europe, Sverige, Dalarna)[..(inhabited place)
Böda.. (Europe, Sverige, Kalmar, Öland)..(inhabited place)
Bodafors... (Europe, Sverige, Jönköping). (inhabited place)
Boda Glasbruk.. (Europe, Sverige, Kalmar)...(inhabited place)
Bodai... (Asia, Bharat, West Bengal)..(inhabited place)
Bodalangi... (Africa, Congo, Équateur).(inhabited place)
Bodalla... (Oceania, Australia, New South Wales)..(inhabited place)
Bodallin.. (Oceania, Australia, Western Australia..(inhabited place)
Bodan Creek.... (N & C Am., USA, Texas, Angelina)...(creek)
Bodan Draw.. (N & C Am., USA, Wyoming, Johnson)...(valley)
Bodane Ridge... (N & C Am., USA, Tennessee, Johnson)...(ridge)

Searching with Boolean operators on keywords can be very valuable, particularly when the user cannot anticipate whether a multiple-word name is listed in reverse or natural order, when names are composed of more than two or three words, or are in languages with which the user is not familiar. For example, searching for "HAWWARAT AND MAQTA" will retrieve the records with the names "Hawwarat al-Maqta'" and "Maqta, Hawwarat al-" and "Al-Maqta, Hawwarat".

In addition to names, other useful qualifiers for retrieval are broader contexts, place type, and coordinates. For broader contexts, it is useful 1) to allow retrieval on parents, grandparents, and all other ancestors (in relational files, linked to a given *geog_key* up the tree through *geog_parents.parent_key*), or 2) by processing the data, to allow retrieval of places under nations and continents.[39] For place types (in

[39] This would require finding nations and continents through their place type, and noting all children linked to these nations and continents. See *Chapter 5: Contents of TGN: The Hierarchy: Administrative Subdivisions.*

relational files, *geog_placetype.placetype_code*, which is linked to the term in *placetype_list.placetype*), it is useful to allow retrieval on multiple place types, perhaps employing the place type codes to allow users to retrieve similar things. For coordinates (in relational files, *geog_main.lat, geog_main.lat_dir, geog_main.long, geog_main.long_dir*, or *geog_main.lat_num* and *geog_main.long_num*), note that around 10 percent of the places in the TGN do not have coordinates in this release, so searching on coordinates will not always retrieve a full set of relevant records.

Results List

The results list should contain enough information to disambiguate places with the same name. Generally the preferred name, preferred place type, and hierarchical position are enough to distinguish places in the list. Geographic coordinates can also be useful in a results list display. Since the result may have been retrieved based on querying a name other than the preferred name, it is also useful to display the name that met the criterion.

Results lists should be sorted in a way to make it easy for end-users to find the place they are seeking. In the example below, the results are sorted first by name, then by the parent string. In this example, the parent string is listed in descending order so that continents, nations, and subdivisions (e.g., counties) will sort together.

EXAMPLE:

Partial results list for query on exact string "Saint Louis":

Galveston.....	inhabited place.....	N & C Am., USA, Texas, Galveston county Saint Louis
Louisburg.....	inhabited place.....	N & C Am., USA, Kansas, Miami county Saint Louis
Old Saint Louis.....	inhabited place.....	N & C Am., USA, Indiana, Bartholomew county Saint Louis
Saint Louis.....	inhabited place.....	Africa, Réunion, Réunion island
Saint Louis.....	region.....	Africa, Sénégal
Saint Louis.....	inhabited place.....	Africa, Sénégal, Saint-Louis region
Saint Louis.....	inhabited place.....	Europe, France, Alsace, Haut-Rhin
Saint-Louis.....	river.....	N & C Am., Canada, Québec
Saint-Louis.....	inhabited place.....	N & C Am., Canada, Saskatchewan
Saint-Louis.....	inhabited place.....	N & C Am., Guadeloupe, Pointe-à-Pitre, Marie ...
Saint-Louis.....	inhabited place.....	N & C Am., USA, California, Sierra county
Saint-Louis.....	inhabited place.....	N & C Am., USA, Georgia, Macon county
Saint-Louis.....	inhabited place.....	N & C Am., USA, Louisiana, Iberville county
Saint-Louis.....	inhabited place.....	N & C Am., USA, Louisiana, Saint Landry county
Saint-Louis.....	inhabited place.....	N & C Am., USA, Michigan, Gratiot county
Saint-Louis.....	county.....	N & C Am., USA, Minnesota
Saint-Louis.....	county.....	N & C Am., USA, Missouri,
Saint-Louis.....	inhabited place.....	N & C Am., USA, Missouri, Saint Louis City
Saint-Louis.....	deserted settlement..	N & C Am., USA, Ohio, Putnam county
Saint-Louis.....	inhabited place.....	N & C Am., USA, Oklahoma, Pottawatomie county
Saint Louis River.....	river.....	N & C Am., USA Saint Louis

DIACRITICS

Diacritics in the TGN Relational Files and REC format are represented by a code preceding the letter over which the diacritical mark appears.[40] The code should be translated into the correct diacritical mark for display, if possible. For example, the code "$04" is an umlaut. The name represented as "M$04unchen" in the TGN data should be translated to "München" for display to the end-user. Diacritics may appear in various fields, including names, display dates descriptive note, and citations (in relational files, *geog_name.place_name, name_detail.display_date, placetype_dates.display_date, geog_main.descr_note, biblio.bib_brief, biblio.citation*). See *Appendix A, Codes for Diacritics* used in the TGN.

EXAMPLES:

$00Aqua = Áqua	**Radam$02a = Radamà**
$04Orebro = Örebro	**T$01ky$01o = Tōkyō**

If your system cannot display diacritics outside the Latin 1 set (e.g., "ō"), depending upon the needs of your end-user, you may consider suppressing the diacritic for the end-user rather than displaying the TGN codes (e.g., displaying "o" instead of "$01o" or "ō").

DISPLAYING A FULL RECORD[41]

Places in the TGN generally should be available in results displays, and in two other displays for the end-user: The full record view and the hierarchical view. The former is a grouping of the data that makes up the intellectual "record" for the place, linked across tables, by the *geog_key* (in relational files).

Record Key

It may be useful to display the record key (in relational files, *geog_key*) to the end-user. Since there are so many homographs in the TGN, this unique identifier for a place is helpful when users wish to refer to a specific record. The unique identifier is requested by the Vocabulary Program when users correspond about updates to records.

EXAMPLE:

PLACE KEY: 7014444

Label

It may be useful to construct a label to refer to the place, for example at the top of a full record or in lists. The minimum information an end-user generally requires to identify a place is the name, parent string, and place type.

---

[40] In the future, TGN may employ Unicode. The TGN USMARC format uses ANSEL (see standard ANSI Z39.47).

[41] For the purpose of this section, "record" refers to an intellectual record — not a database record.

Note that the TGN has flagged a special Display Name (in relational files, *where geog_name.other_flags = DIS*) to be used in strings of names when the place is presented as a broader context, but when using its preferred name would result in ambiguity or confusion.[42] This usually occurs when an administrative subdivision has the same name as an inhabited place. For example, Siena the inhabited place and Siena the province have the same preferred name, "Siena." Therefore, in horizontal displays of the hierarchical string, the province should be represented by its display name, "Siena province," in order to avoid confusing it with the inhabited place of the same name (e.g., **Siena (Siena province, Toscana, Italia**).

In the example below, the parent string is sorted in ascending order, because this is "natural" order and probably easier for end-users to read (see example in *Results List* section above for parent string in descending order). The name and parent string are combined with preferred place type to identify the place.

EXAMPLE:

Saint Louis (*Saint Louis City, Missouri, United States*), inhabited place

Displaying Names

It is recommended to display the names together with the flags indicating whether the name is Vernacular/Other language (in relational files, *where geog_name.lang = V or O*), and Current/Historical (*where geog_name.age_flag = C or H*). Users may also be interested in other flags as well. Many users wish to know if the name is the "preferred" (*where geog_name.pref_flag* = P) or the "preferred English" name (in relational files, *where geog_name.other_flags = ENG*). It may be useful to translate some or all of the flags into forms more easily understood by the end-user (in the example below, P is translated to "Pref," ENG to "prefEng,"). A key for the flags should be available to the end-user.

EXAMPLE:

C = Current	**Pref = Preferred name**
H = Historical	**prefEng = preferred English name**
V = Vernacular	**ISO = 3-letter ISO code**
O = Other language	**USPS = US Postal Service abbreviation**
	Dis = Display Name

The names should also display with their appropriate "display date" (in relational files, *name_detail.display_date*) describing when the name was used. The start and end dates (in relational files, *name_detail.start_date* and *name_detail.end_date*) that index this display date should ***not*** be displayed to the end-user.[43]

EXAMPLE:

Saint Louis (C,V,Pref) ***named for Louis IX of France in 1764***

[42] Note: Always use the preferred name (or the preferred English name) for the place that is the *focus* of the label, but use the display name for the *parents* in the string (when parents have a display name, otherwise default the parents' preferred name).

[43] The start and end dates may be used in retrieval; however, as yet, most names do not have these data. The structure is included to allow for the possibility that TGN could in the future be employed to display historical views based on dates on names, place types, and parent links.

Sorting Names

Names should be arranged in logical order by sorting them according to sequence numbers (in relational files, *geog_name.seqn*). It is not necessary that the end-user actually see the sequence numbers.

EXAMPLE:

0 **Saint Louis (C,V,Pref)** ***named for Louis IX of France in 1764***
1 **St. Louis (C,V)**
2 **Mound City (C,V)** ***nickname used by boatmen, referring to Native American mounds on the site***
3 **Pain Court (H,V)**
4 **Pancore (H,V)**

Sources and Contributors

Sources and Contributors should be displayed to end-users; the data includes contributors (in relational files, *name_detail.contrib*), brief citations (in relational files, *biblio.bib_brief*, linked to name by *name_detail.bib_key*), and page numbers (in relational files, *name_detail.bib_page*) associated with the names (in relational files, linked by *placename_id*).

EXAMPLE:

Saint Louis:	**Encyclopaedia Britannica (1988) X, 327 / [VP]**
	Getty Research Library Photo Archive, Authority file / [GRLPA]
	Times Atlas of the World (1992) 17 / [VP]
	USGS, GNIS Digital Gazetteer (1994) / [VP]
	Webster's Geographical Dictionary (1984) / [BHA, FDA]

Users should have access to a key explaining the abbreviations of the contributors. See *Appendix B, Contributor Initials List for the TGN*. Future development of the TGN will involve additional contributors; therefore this list will be updated with future releases.

EXAMPLE:

BHA	**Bibliography of the History of Art**
FDA	**Foundation for Documents of Architecture**
GRLPA	**Getty Research Library Photo Archive**
VP	**Vocabulary Program**

It is intended that the citation displayed with name will be the "brief citation" (in relational files, *biblio.bib_brief*). The full citation (in relational files, *biblio.citation*) is quite long and generally inconvenient for such displays. However, users should have access to a key citing the brief citation and the corresponding full citation.

EXAMPLE:

USGS, GNIS Digital Gazetteer (1994)
United States Department of the Interior and U.S. Geological Survey. *Geographic Names Information System Digital Gazetteer.* Reston, Virginia: U.S. Geological Survey, 1994.

Displaying Place Types

Place types should be displayed with the appropriate flags and display dates (in relational files, place types are *placetype_list.placetype* linked to place record by *geog_placetype.placetype_code*; flags are *geog_placetype.age_flag* and where *geog_placetype.pref_flag = P*; display datres are *placetype_dates.display_date*). As was true for names, the start and end dates that index the display date should not be visible to the end-user (in relational files, *placetype_dates.start_date*, *plactype_dates.end_date*).

EXAMPLE:

inhabited place (C, Pref) ***Native American ceremonial site; founded in 1764 as fur station by French pioneer Pierre Lacléde Liguest***

Also, end-users should have access to a key explaining the meaning of the flags.

EXAMPLE:

C = Current **Pref = Preferred place type**
H = Historical

Sorting Place Types

Place types should be arranged in logical order by sorting them according to the sequence numbers (in relational files, *geog_placetype.seqn*). As with the sort order for names, it is not necessary that the end-users actually see the sequence numbers.

EXAMPLE:

0 **inhabited place (C, Pref)** ***Native American ceremonial site; founded in 1764 as fur station by French pioneer Pierre Lacléde Liguest***
1 **city (C)** ***since 1822***
2 **river port (C)**
3 **industrial center (C)** ***for aircraft, iron, steel & automobiles***
4 **agricultural center (C)**

Coordinates

Latitude and longitude (in relational files, *geog_main.lat, geog_main.lat_dir, geog_main.long, geog_main.long_dir*) in the TGN are expressed in degrees and minutes. It is recommended to clearly separate degrees and minutes in displays, as in any of the examples below (with space, degree symbol, or hyphen). Note that if a period is used to separate degrees and minutes (as is the format in some major atlases), the user may mistakenly assume that the format represents decimal degrees rather than degrees and minutes.[44]

EXAMPLES:

38 37 N, 090 11 W **38° 37’ N, 090° 11’ W** **38-37 N, 090-11 W**

Latitude and longitude expressed in degrees and minutes will probably be most familiar to most end-users. However, the TGN also includes coordinate data expressed as decimal degrees (in relational files, *geog_main.lat_num* and *geog_main.long_num*), which are numbers that can hold decimal points and can be used for calculations and for retrieval (and could be comprehensible to some end-users). For decimal degrees, the minutes of latitude and longitude are converted to decimal fractions of degrees (e.g., *38.616* and *-090.183* are the decimal degrees for the example above). Whole degrees of latitude are represented by a two-digit decimal number ranging from 0 through 90, with any decimal fraction of a degree separated by a decimal point; latitudes south of the equator are expressed as negative numbers (designated by a minus sign). Whole degrees of longitude are represented by a three-digit decimal number ranging from 0 through 180, with any decimal fraction of a degree separated by a decimal point; longitudes west of the Prime Meridian are expressed as negative numbers (designated by a minus sign).

Descriptive Note

It is recommended to display the descriptive note (in relational files, *geog_main.descr_note)* to the end-user.

EXAMPLE:

Located on Mississippi river in E Missouri, adjacent to, but independent of, St. Louis county; largest city in state; given to Spain in 1770, went to France in 1800; after Louisiana purchase, became part of USA; center of Unionist support in Missouri during Civil War.

[44] The expression of coordinates will be updated in future releases of TGN and will include structure for recording seconds and for bounding coordinates.

Other Information

It is also useful to display the hierarchical position of the place in the full record (in addition to allowing navigation of the hierarchy in a hierarchy display). See also section below, *Hierarchy: Non-Preferred Parents.*

Certain data are not appropriate to display to the end-user. This includes the start and end dates intended to index display dates (in relational files, in *name_detail* and *placetype_dates*), normalized sort names (in relational files, *geog_name.sort_name*), and keywords (in relational files, in *name_keywords*).

Depending upon the end-users, they may or may not find the unique keys helpful; the most important keys to end-users are likely to be the key that identifies the "record" (in relational files, *geog_key*) or keys for the individual names (in relational files, *placename_id*). In the example below, keys of the record and of individual names are displayed. Users may also find it useful to see the place type codes (in relational files, *placetype_code*) with the place type term. See *Chapter 4, Contents of the TGN* for another sample record, where only the "record" key is displayed.

SAMPLE RECORD DISPLAY:

PLACE KEY: 7014444
LABEL: **St. Louis *(Saint Louis City, Missouri, United States),* inhabited place**

NAMES:

Saint Louis (C,V,Pref) ***named for Louis IX of France in 1764***
name key: 9867
St. Louis (C,V)
name key: 87326
Mound City (C.V) ***nickname used by boatmen, referring to Native American mounds on the site***
name key: 81124
Pain Court (H,V)
name key: 1254
Pancore (H,V)
name key: 176

COORDINATES: 38° 37' N, 090° 11' W

DESCRIPTIVE NOTE:

Located on Mississippi river in E Missouri, adjacent to but independent of St. Louis county; largest city in state; given to Spain in 1770, went to France in 1800; after Louisiana purchase, became part of USA; center of Unionist support in Missouri during Civil War.

PLACE TYPES:

inhabited place (C, Pref) ***Native American ceremonial site; founded in 1764 as fur station by French pioneer Pierre Lacléde Liguest***
city (C) ***since 1822***
river port (C)
industrial center (C) ***for aircraft, iron, steel & automobiles***
agricultural center (C)

HIERARCHY:

- **North and Central America (continent)**
- **United States (nation)**
- **Missouri (state)**
- **Saint Louis City (independent city)**
- **Saint Louis (inhabited place)**

CITATION/CONTRIBUTOR :

Saint Louis:	**Encyclopaedia Britannica (1988) X, 327 / [VP]**
	Getty Research Library Photo Archive, Authority file / [GRLPA]
	Times Atlas of the World (1992) 17 / [VP]
	USGS, GNIS Digital Gazetteer (1994) / [VP]
	Webster's Geographical Dictionary (1984) / [BHA, FDA]
St. Louis:	**Webster's Geographical Dictionary (1984) / [GRLPA]**
Mound City:	**Faherty, St. Louis History (1989) / [VP]**
Pain Court:	**USGS, GNIS Digital Gazetteer (1994) / [VP]**
Pancore:	**USGS, GNIS Digital Gazetteer (1994) / [VP]**

DISPLAYING THE HIERARCHY

The hierarchy is constructed by following links up the tree through *geog_parents.parent_key* (in relational files). It is recommended that places in the hierarchy be represented by 1) either their "preferred" name (in relational files, *where geog_name.pref_flag = P*) or the "preferred English" name (in relational files, *where geog_name.other_flags = ENG*), and 2) the preferred place type (in parentheses in the example below, in relational files, *placetype_list.placetype*, linked to *geog_key* by *geog_placetype.placetype_code*).

It is recommended that the hierarchy generally be displayed with indentation indicating part/whole relationships. As the user maneuvers through the hierarchical displays, it is typically necessary to display all levels above the target and one level beneath. ***Aisén (region)*** is the target in the example below.

EXAMPLE:

World
South America (continent)
Chile (nation)
Aisén (region)
Anna Pink, Bahía (bay)
Baker (river)
Baker, Canal (channel)
Balmaceda (inhabited place)
Campana, Isla (island)
Chile Chico (inhabited place)

Indicating Levels

When users navigate through levels of the hierarchy, it is generally recommended to show only one level beneath the target. Since the hierarchy is so long at any one level, showing the full depth and breadth beneath the target could be confusing.

EXAMPLE:

World
Africa (continent)
Kenya (nation)
Central (province) ...
Choichuff, Laga (river)
Coast (province) ...

However, some method should be employed to indicate that further levels are present. An ellipsis (...) is used in the example above, though a hierarchy symbol, plus/minus icon, or other icon would also work (as illustrated in example below).

EXAMPLE:

World
Africa (continent)
Kenya (nation)
Central (province)
Choichuff, Laga (river)
Coast (province)

Accommodating Siblings

There can be thousands of siblings at any level of the TGN hierarchy. This poses challenges in creating a coherent hierarchical display. One solution is to simply allow users to scroll through all places at any level; however, this may mean they must scroll through dozens of pages. If this method is employed, it will aid the user if nations and major political subdivisions are highlighted.

EXAMPLE:

World
Europe (continent)
Achen Pass (pass)
Adriatic Sea (sea)
Agepsta, gora (mountain)
Águeda (river)
Ajdar (river)
Andorra (nation)

Another strategy for organizing a long list of siblings is to separate the physical world from the political world, using place types and other features explained in *Chapter 4, Contents of the TGN: Physical and Political Geography*, and *Administrative Subdivisions*. The user could alternate between a view of the physical features and one of the political world, as illustrated in the example below. Note, however, that the complexities of the relationships between the physical and political world must be considered. This is particularly important where a physical feature forms an intervening level between levels of the political world, and where a physical feature, such as an island, can be both a physical and political place.

EXAMPLE:

World

Europe (continent)

[+] **[View PHYSICAL FEATURES]**

- [+] **Andorra (nation)**
- [+] **Bâlgaria (nation)**
- [+] **Belarus (nation)**
- [+] **België (nation)**

Yet another method to accommodate large numbers of siblings is illustrated in the example below. When there are more than 200 places at a given level of the hierarchy, these places could be indexed alphabetically.

EXAMPLE:

World

Europe (continent)

<A> count = 78...
<B> count = 86 ...
<C> count = 109...
<D> count = 67...

Non-Preferred Parents

Places in the TGN may have multiple parents. However, one parent is flagged as preferred. That is, in relational files, for each *geog_key*, there is only one *geog_parents.parent_key where pref_flag = P*, but in addition there may be one or more *geog_parents.parent_key where pref_flag = N*.

The place should display under ***both*** its parents. The example immediately below illustrates the state of Hawaii displaying under its political parent (its preferred parent).

EXAMPLE:

World
North and Central America (continent)
United States (nation)
+ **Alabama (state)**
+ **Alaska (state)**
+ **American Samoa [N] (dependent state)**
+ **Arizona (state)**
+ **Arkansas (state)**
+ **California (state)**
+ **Colorado (state)**
+ **Connecticut (state)**
+ **Delaware (state)**
+ **District of Columbia (state)**
+ **Florida (state)**
+ **Georgia (state)**
+ ***Hawaii (state)***
+ **Idaho (state)**

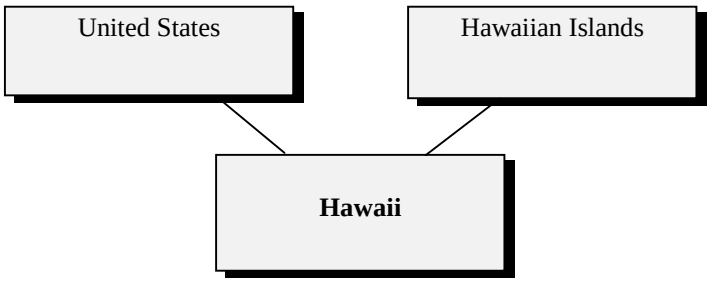
It is useful to alert the end-user when multiple parents exist. One way to do this is to mark the non-preferred link. In the example below, the state of Hawaii is displayed in an alternate hierarchical view with the island group of which it is a part, and the non-preferred relationship is indicated with an "N."

EXAMPLE:

World
Oceania (continent)
Hawaiian Islands (island group)
+ ***Hawaii [N] (state)***
+ **Midway Islands [N] (possession)**

Other graphics may be used to illustrate the relationship of a place to multiple parents, as in the example below.

EXAMPLE:



Multiple parents should also be represented in the full record display for the place.

EXAMPLE:

PLACE KEY:7007249
LABEL: **Hawaii** ***(United States, North and Central America),*** **state**

NAMES:
Hawaii (C,V,Pref) ***name is probably derived from a native word meaning "homeland"***
HI (C,V, USPS)

LAT: 20° 00' N, LONG: 157° 50 W

DESCRIPTIVE NOTE:
One of 2 non-contiguous states; comprises most of the Hawaiian Islands; probably first inhabited by Polynesians who immigrated from Marquesas Islands ca. AD 400; immigrants from Tahiti arrived in 9th cen.; Europeans & Americans arrived in 18th cen.

HIERARCHICAL POSITION:
- **North and Central America (continent)**
- **United States (nation)**
- **Hawaii (state)**

ADDITIONAL PARENTS:
- **Oceania (continent)**
- **Hawaiian Islands (island group)**
- **Hawaii (state)**

PLACE TYPES:
state (C) ***since August 21, 1959***
first level subdivision (C)

CITATION/CONTRIBUTOR :

HI	**Webster's Geographical Dictionary (1988), [VP]**
Hawaii	**National Zip Code Directory (1985) [BHA]**
	Columbia Lippincott Gazetteer (1961) [BHA]
	Random House English Dictionary (1987) [BHA]
	LC Subject Headings, 11th edition (1988) [BHA]
	Webster's Geographical Dictionary (1984) [BHA, FDA]
	Canby, Historic Places (1984), I, 380-381 [VP]
	Times Atlas of the World (1994), Plate 114 [VP]
	Encyclopædia Britannica (1988), XXIX, 447 ff. [VP]
	USGS, GNIS Digital Gazetteer (1994), GNIS88000311 [VP]

APPENDIX A
CODES FOR DIACRITICS

OVERVIEW

The following chart lists the codes used to indicate diacritics in the vocabulary data for AAT,[45] ULAN, and TGN. Each code consists of the dollar sign ($) followed by two numbers. This code is placed before (in front of) the letter to which the diacritical mark applies. The same code can be applied to multiple letters. For example, if an acute accent should be applied to an ***a*** (á), it is recorded as ***$00a***; if an acute accent should be applied to an ***e*** (é), it is recorded as ***$00e***. In some cases, the code means that two diacritics are placed over the same character (e.g., $30). In other isolated cases, the code applies to two adjacent characters (e.g., $57, a digraph).

DIACRITICAL CODES CHART

Getty Vocabularies' Diacritical Codes[46]			
Vocab Code	***Example in Vocab***	***Example to Display or Print***	***Diacritic Name***
$00	Andr$00e $00Aqua	André Áqua	acute accent/ miagkiy znak
$01	T$01oky$01o $01Agra	Tōkyō Āgra	macron
$02	Radam$02a $02Ecole	Radamà Ècole	grave accent
$03	P$03orto $03Etat	Pôrto Êtat	circumflex
$04	M$04unchen $04Orebro	München Örebro	umlaut / dieresis
$05	Fran$05cois $05Sisli	François Şisli	cedilla
$06	Br$06aila $06Uiju	Brăila Ŭiju	breve
$07	Franti$07sek $07Zabari	František Žabari	hacek / wedge
$08	Klaip$08eda $08Ism$08ir	Klaipėda İsmir	superior dot (dot above)
$09	Jap$09ao $09Naupe	Japão Ñaupe	tilde

[45] In the current release, AAT does not use this set of diacritic codes; however, it will use them in 2001.
[46] The asterisk (*) in the Diacritic Name column indicates that the correct diacritic may not display in the Example to Display or Print column.

Getty Vocabularies' Diacritical Codes[46]			
Vocab Code	***Example in Vocab***	***Example to Display or Print***	***Diacritic Name***
$10	Dv$10ur $10Alborg	Dvůr Ålborg	angtrsom / circle above
$12	Gy$12or $12Ogyr	Győr Őgyr	double acute accent / tverdyi znak
$13	W$13lodzim $13Lodz	Włodzim Łodz	Polish l / slashed l
$14	$14Orslev Agers$14o	Ørslev Agersø	Scandinavian o / slashed o
$15	$15D	Ḍ	dot below (sub-dot)
$16	$16L $16l	Ŀ ŀ	dot right side
$17	B$13l$17ed-$00ow	Błędów	right hook
$18	Stra$18sburg	Straßburg	eszett
$19	$19th	þ	small thorn
$20	$20TH	Þ	large thorn
$21	$21th	ð	eth / edh
$22	$22IA	IA *	ligature, first half*
$23	$23ts	ts *	ligature, second half*
$24	$55D$24oc	Đôc *	circumflex and acute accent *
$25	$25u	ụ̉	hook above and dot below (sub-dot)
$26	$26o	ớ	hook above and acute accent
$27	$27A	Ắ	breve and acute accent
$28	$28o	ở	hook above
$29	$29A	Ậ	circumflex and dot below (sub-dot)
$30	$30o	ồ	circumflex and grave accent
$31	$31A	Ằ	breve and grave accent
$32	$32U	Ừ	horn (pseudo question mark) and grave accent
$33	$33E	Ề	horn (pseudo question mark) and circumflex
$34	Th$34o	Thô *	circumflex and hook above *
$35	$35O	Ổ	horn (pseudo question mark) and hook above

Getty Vocabularies' Diacritical Codes[46]			
Vocab Code	***Example in Vocab***	***Example to Display or Print***	***Diacritic Name***
$36	$36O	Õ̉	hook above and tilde
$37	$37A	Ã̆	breve and tilde
$38	$38o	ỗ	circumflex and tilde
$39	$39a	ặ	breve and dot below (sub-dot)
$40	$40a	ả	breve and horn (pseudo question mark)
$41	$41I	Ḯ	dieresis (umlaut) and acute accent
$42	$42O	Ò̉	hook above and grave accent
$46	Na$46zerat Ha$46Zafon	Na̲zerat Ha̲Zafon	underscore / sub-macron
$47	Bra$47sov $47S$01ur	Braşov Şūr	left hook / tail
$48	L$48ebork G$48abin	Lębork Gąbin	right cedilla
$49	$49L	L̦	comma below / sub-comma
$50	$50E	Ė	horn (pseudo question mark)
$55	$55Dakovo Sta$55dur	Đakovo Stađur	d with crossbar / barred d
$56	$56T $56t	Ŧ ŧ	T with crossbar / barred t
$57	$57Aero	Æro	digraph (ligature) AE uppercase
$58	$58Oeuf	Œuf	digraph (ligature) OE uppercase
$59	$59OE	Œ̆	digraph (ligature) OE uppercase and breve
$60	$60oe	œ̆	digraph (ligature) oe lowercase and breve
$65	$65g	ǵ	inverted apostrophe above
$66	$66L	L̓	half-space apostrophe
$67	Nuk$67alofa $67Akko	Nuku‘alofa ‘Akko	ayn
$68	$68	‘	inverted apostrophe
$69	$69	”	double apostrophe
$70	B$70aek	Bæk	digraph (ligature) ae lowercase
$71	Sch$71oelcher	Schœlcher	digraph (ligature) oe lowercase

Getty Vocabularies' Diacritical Codes[46]			
Vocab Code	***Example in Vocab***	***Example to Display or Print***	***Diacritic Name***
$73	K$73ir$73ik-kale	Kırıkkale	Turkish i lowercase / undotted i
$74	$74[SPACE]	·	center dot
$81	G$81hawdex $81Hawdex	Għawdex Ħawdex	h with cross bar / barred h
$85	$85SH $85Sh	**SH** Sh	underscore sh / extended macron sh
$86	$86ZH $85Zh	**ZH** **Zh**	underscore zh / extended macron zh
$91	$91a	ɐ	alif
$92	$92A $92a	ə	schwa
$93	$93N $93n	ŋ ŋ	eng

APPENDIX B
CONTRIBUTOR INITIALS LIST FOR THE TGN

CONTRIBUTORS

Contributors of names to the TGN are noted by their initials or an abbreviated form of their name. End-users should have access to a key that lists the initials and the full name of the Getty project.

Contributors to the TGN

BHA	**Bibliography of the History of Art**
FDA	**Foundation for Documents of Architecture**
GRLPA	**Getty Research Library Photo Archives**
VP	**Vocabulary Program**

BACKGROUND

Contributors include abstracting and indexing projects, collections of images, object-based collections, and vocabulary experts.

The **Bibliography of the History of Art/Bibliographie d'Histoire de l'Art** is jointly sponsored by the J. Paul Getty Trust and the Institut de l'Information Scientifique et Technique (INIST), a division of France's Centre National de la Recherche Scientifique (CNRS). The BHA produces abstracts and indexes of current literature on the history of art, including books, conference proceedings, dissertations, exhibition and dealer's catalogs, and articles from more than 2,500 periodicals. It focuses primarily on European and American art. The Bibliography of the History of Art includes and extends the coverage of its two predecessor art indexes: RAA (Repertoire d'Art et d'Archeologie) from 1973 to 1989 and RILA (International Repertory of the Literature of Art) from 1975 to 1989.

Place names contributed to the TGN by the **Foundation for Documents of Architecture** were derived from the geographic authority in the FDA's experimental database. This database was compiled by FDA in 1988-1989, during a period FDA to cataloged architectural drawings in order to test guidelines proposed by the Architectural Drawings Advisory Group (ADAG). ADAG was a consortium of institutions representing museums, libraries, archives, and art historians world-wide, including the Royal Institute of British Architects; the Centre Canadien d'Architecture/Canadian Centre for Architecture; the National Archives of Canada; the American Architectural Foundation, American Institute of Architects; the Avery Architectural and Fine Arts Library, Columbia University; the Cooper-Hewitt Museum, Smithsonian Institution; the Library of Congress; the National Archives and Records Administration of the United States; the National Gallery of Art, Washington, D.C.; the Victoria and Albert Museum; the Deutsches Architekturmuseum, the École Supérieure des Beaux-Arts, Graphische Sammlung Albertina, and the Royal Library, Windsor. The members of FDA who contributed information were the Canadian Centre for Architecture, the National

Gallery of Art in Washington, the National Archives and Records Administration in Washington, the National Archives of Canada, and the J. Paul Getty Trust. [47]

The **Getty Research Library Photo Archive** provided geographic names used to index their visual collections, which document the history of art from Antiquity to the present. The Photo Archive contains approximately two million photographs, primarily black and white, which document a wide variety of works of art and architecture.

The **Vocabulary Program** (formerly known as the Vocabulary Coordination Group, or VCG) is located in the Getty Research Institute. The Vocabulary Program compiles three vocabularies, the *Art & Architecture Thesaurus* (AAT), the *Union List of Artist Names* (ULAN), and the *Getty Thesaurus of Geographic Names* (TGN).

Other contributors: For some contributors, the Vocabulary Program editors received information in hardcopy and entered it by hand in the editorial system.[48] In such cases, "VP" is listed as the "contributor," and the institution or project that contributed the data is noted as the source in the bibliography.

[47] A standard for cataloging architectural drawings resulted: Porter, Vicki and Robin Thornes. *A Guide to the Description of Architectural Drawings*. New York: G.K. Hall & Co., 1994 (see http://www.getty.edu/gri/standard/fda).

[48] Future contributions will be accepted only in automated form, in the format prescribed by the Vocabulary Program (available in Spring, 2001).

APPENDIX C
SAMPLE TGN DATA

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
GEOG_MAIN	
geog_key	**1114064**
descr_note	Separates Maui & Hawaii islands.
lat	20 26
lat_dir	N
long	156 00
long_dir	W
lat_num	20.433
long_num	-156.000
GEOG_PARENTS[49]	
geog_key	**1114064**
parent_key	7007249
pref_flag	P
GEOG_NAME	
geog_key	**1114064**
placename_id	**115296**
seqn	0
place_name	Alenuihaha Channel
sort_name	ALENUIHAHACHANNEL
pref_flag	P
lang	V
age_flag	C
other_flags	N
geog_key	**1114064**
placename_id	**366362**
seqn	1
place_name	Alennihaha Canal
sort_name	ALENNIHAHACANAL
pref_flag	N

[49] Note that the parent in this table is the state of Hawaii, which is the second sample record.

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
lang	V
age_flag	C
other_flags	N
geog_key	**1114064**
placename_id	**366363**
seqn	2
place_name	Hawaii Channel
sort_name	HAWAIICHANNEL
pref_flag	N
lang	V
age_flag	C
other_flags	N
geog_key	**1114064**
placename_id	**366364**
seqn	3
place_name	Kai o Aleuuihaha
sort_name	KAIOALEUUIHAHA
pref_flag	N
lang	V
age_flag	C
other_flags	N

NAME_DETAIL

geog_key	**1114064**
placename_id	**115296**
bib_key	9006567
bib_page	plate 114
contrib	VP
display_date	
start_date	
end_date	
geog_key	1114064
placename_id	115296
bib_key	9006563
bib_page	GNIS15000186
contrib	VP
display_date	
start_date	
end_date	
geog_key	**1114064**
placename_id	**366362**
bib_key	9006563
bib_page	GNIS15000186

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
contrib	VP
display_date	
start_date	
end_date	
geog_key	**1114064**
placename_id	**366363**
bib_key	9006563
bib_page	GNIS15000186
contrib	VP
display_date	
start_date	
end_date	
geog_key	**1114064**
placename_id	**366364**
bib_key	9006563
bib_page	GNIS15000186
contrib	VP
display_date	
start_date	
end_date	
GEOG_PLACETYPE	
geog_key	**1114064**
placetype_code	21151
seqn	0
pref_flag	P
age_flag	C
PLACETYPE_DATES	
geog_key	
placetype_code	
display_date	
start_date	
end_date	
BIBLIO	
bib_key	**9006567**
bib_brief	Times Atlas of the World (Reprinted 1994)
bib_brief_sort	TIMESATLASOFTHEWORLDREPRINTED
citation	Times Atlas of the World. 9th comprehensive edition. Reprinted with revisions 1994. New York: Times Books, 1994.

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
bib_key	**9006553**
bib_brief	USGS, GNIS Digital Gazetteer (1994)
bib_brief_sort	USGSGNISDIGITALGAZETTEER
citation	United StatesDepartment of the Interior and U.S. Geological Survey. Geographic Names Information System Digital Gazetteer. Reston, Virginia: U.S. Geological Survey, 1994.

PLACETYPE_LIST

placetype_code	**21151**
placetype_search	21151
placetype	channel

NAME_KEYWORDS

geog_key	**1114064**
placename_id	**115296**
keyword	ALENUIHAHA
geog_key	1114064
placename_id	115296
keyword	CHANNEL
geog_key	**1114064**
placename_id	**366362**
keyword	ALENNIHAHA
geog_key	1114064
placename_id	366362
keyword	CANAL
geog_key	**1114064**
placename_id	**366363**
keyword	HAWAII
geog_key	1114064
placename_id	366363
keyword	CHANNEL
geog_key	**1114064**
placename_id	**366364**
keyword	KAI
geog_key	1114064
placename_id	366364
keyword	O
geog_key	1114064
placename_id	366364
keyword	ALEUUIHAHA

BIB_KEYWORDS

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
bib_key	**9006567**
keyword	TIMES
bib_key	9006567
keyword	ATLAS
bib_key	9006567
keyword	OF
bib_key	9006567
keyword	THE
bib_key	9006567
keyword	WORLD
bib_key	9006567
keyword	COMPREHENSIVE
bib_key	9006567
keyword	EDITION
bib_key	9006567
keyword	REPRINTED
bib_key	9006567
keyword	WITH
bib_key	9006567
keyword	REVISIONS
bib_key	9006567
keyword	NEW
bib_key	9006567
keyword	YORK
bib_key	9006567
keyword	TIMES
bib_key	9006567
keyword	BOOKS
bib_key	**9006553**
keyword	UNITED
bib_key	9006553
keyword	STATES
bib_key	9006553
keyword	DEPARTMENT
bib_key	9006553
keyword	OF
bib_key	9006553
keyword	THE
bib_key	9006553
keyword	INTERIOR
bib_key	9006553
keyword	AND
bib_key	9006553
keyword	U

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
bib_key	9006553
keyword	S
bib_key	9006553
keyword	GEOLOGICAL
bib_key	9006553
keyword	SURVEY
bib_key	9006553
keyword	GEOGRAPHIC
bib_key	9006553
keyword	NAMES
bib_key	9006553
keyword	INFORMATION
bib_key	9006553
keyword	SYSTEM
bib_key	9006553
keyword	DIGITAL
bib_key	9006553
keyword	GAZETTEER
bib_key	9006553
keyword	RESTON
bib_key	9006553
keyword	VIRGINIA
bib_key	9006553
keyword	U
bib_key	9006553
keyword	S
bib_key	9006553
keyword	GEOLOGICAL
bib_key	9006553
keyword	SURVEY
FORMER_GEOGKEYS	
new_key	
master_key	
collapse_key	
FORMER_GEOGLINKS	
new_key	
collapse_key	
old_link_key	
FORMER_BIBKEYS	

SAMPLE RECORD 1: Relational Files Format

Alenuihaha Channel (Hawaii, United States, North and Central America) channel

COLUMN	*EXAMPLE*
new_key	
master_key	
collapse_key	
FORMER_BIBLINKS	
new_key	
collapse_key	
old_link_key	

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
GEOG_MAIN	
geog_key	**7007249**
descr_note	One of 2 non-contiguous states; comprises most of the Hawaiian Islands; probably first inhabited by Polynesians who immigrated from Marquesas Islands ca. AD 400; immigrants from Tahiti arrived in 9th cen.; Europeans & Americans arrived in 18th cen.
lat	20 00
lat_dir	N
long	157 50
long_dir	W
lat_num	20.000
long_num	-157.833
GEOG_PARENTS	
geog_key	**7007249**
parent_key	7012149
pref_flag	P
geog_key	**7007249**
parent_key	7006220
pref_flag	N
GEOG_NAME	
geog_key	**7007249**
placename_id	**365**
seqn	0
place_name	Hawaii
sort_name	HAWAII
pref_flag	P
lang	V
age_flag	C
other_flags	N
geog_key	**7007249**
placename_id	**321956**
seqn	1
place_name	HI
sort_name	HI
pref_flag	N
lang	O
age_flag	C

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
other_flags	U
NAME_DETAIL	
geog_key	**7007249**
placename_id	**365**
bib_key	9006303
bib_page	
contrib	BHA
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9005681
bib_page	
contrib	BHA
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9006048
bib_page	
contrib	BHA
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9001601
bib_page	
contrib	BHA
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9006449
bib_page	
contrib	BHA
display_date	
start_date	
end_date	

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
geog_key	7007249
placename_id	365
bib_key	9006449
bib_page	
contrib	FDA
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9006447
bib_page	I, 380-381
contrib	VP
display_date	name is probably derived from a native word meaning “homeland”
start_date	1700
end_date	9999
geog_key	7007249
placename_id	365
bib_key	9005014
bib_page	XXIX, 447 ff.
contrib	VP
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9006549
bib_page	plate 114
contrib	VP
display_date	
start_date	
end_date	
geog_key	7007249
placename_id	365
bib_key	9006553
bib_page	GNIS88000311
contrib	VP
display_date	
start_date	
end_date	
geog_key	**7007249**
placename_id	**321956**
bib_key	9006267
bib_page	1256

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
contrib	VP
display_date	
start_date	
end_date	
GEOG_PLACETYPE	
geog_key	**7007249**
placetype_code	81175
seqn	0
pref_flag	P
age_flag	C
geog_key	**7007249**
placetype_code	81100
seqn	1
pref_flag	N
age_flag	C
PLACETYPE_DATES	
geog_key	**7007249**
placetype_code	81175
display_date	since August 21, 1959
start_date	1959
end_date	9999
BIBLIO	
bib_key	**9001601**
bib_brief	Random House English Dictionary (1987)
bib_brief_sort	RANDOMHOUSEENGLISHDICTIONARY
citation	The Random House Dictionary of the English Language. Stuart Berg Flexner, editor-in-chief. 2nd ed. New York: Random House, 1987.
bib_key	**9005014**
bib_brief	Encyclop$70aedia Britannica (1988)
bib_brief_sort	ENCYCLOPAEDIABRITANNICA
citation	The New Encyclop$70aedia Britannica. 15th ed. Chicago: Encyclop$70aedia Britannica Inc., 1988.
bib_key	**9005681**
bib_brief	LC Subject Headings, 11th edition (1988)
bib_brief_sort	LCSUBJECTHEADINGSTHEDITION
citation	Library of Congress Subject Cataloging Division. Library of Congress Subject Headings. 11th ed. Washington: Cataloging Distribution Service, Library of Congress, 1988.
bib_key	**9006037**

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
bib_brief	Times Atlas of the World (1992)
bib_brief_sort	TIMESATLASOFTHEWORLD
citation	Times Atlas of the World. 9th comprehensive edition. New York: Times Books, 1992.
bib_key	**9006048**
bib_brief	National Zip Code Directory (1985)
bib_brief_sort	NATIONALZIPCODEDIRECTORY
citation	National Five Digit Zip Code and Post Office Directory. Washington: U.S. Postal Service, 1985.
bib_key	**9006267**
bib_brief	Webster's Geographical Dictionary (1988)
bib_brief_sort	WEBSTERSGEOGRAPHICALDICTIONARY
citation	Webster's New Geographical Dictionary. Sprinfield, MA: Merriam-Webster, 1988.
bib_key	**9006303**
bib_brief	Columbia Lippincott Gazetteer (1961)
bib_brief_sort	COLUMBIALIPPINCOTTGAZETTEER
citation	Columbia Lippincott Gazetteer of the World. Edited by Leon E. Seltzer. Morningside Heights, NY: Columbia University Press, 1961.
bib_key	**9006447**
bib_brief	Canby, Historic Places (1984)
bib_brief_sort	CANBYHISTORICPLACES
citation	Canby, Courtlandt. The Encyclopedia of Historic Places. New York: Facts on File Publications, 1984.
bib_key	**9006449**
bib_brief	Webster's Geographical Dictionary (1984)
bib_brief_sort	WEBSTERSGEOGRAPHICALDICTIONARY
citation	Webster's New Geographical Dictionary. Springfield, MA: Merriam-Webster, 1984.
bib_key	**9006553**
bib_brief	USGS, GNIS Digital Gazetteer (1994)
bib_brief_sort	USGSGNISDIGITALGAZETTEER
citation	United States Department of the Interior and U.S. Geological Survey. Geographic Names Information System Digital Gazetteer. Reston, Virginia: U.S. Geological Survey, 1994.
bib_key	**9893654**
bib_brief	Times Atlas of the World (1994)
bib_brief_sort	TIMESATLASOFTHEWORLD
citation	Times Atlas of the World. 9th comprehensive edition. New York: Times Books, 1994.

PLACETYPE_LIST

placetype_code	**81175**
placetype_search	81175

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
placetype	state
placetype_code	**81100**
placetype_search	81100
placetype	first level subdivision
NAME_KEYWORDS	
geog_key	**7007249**
placename_id	**365**
keyword	HAWAII
geog_key	**7007249**
placename_id	**321956**
keyword	HI
BIB_KEYWORDS	
bib_key	**9001601**
keyword	THE
bib_key	9001601
keyword	RANDOM
bib_key	9001601
keyword	HOUSE
bib_key	9001601
keyword	DICTIONARY
bib_key	9001601
keyword	OF
bib_key	9001601
keyword	THE
bib_key	9001601
keyword	ENGLISH
bib_key	9001601
keyword	LANGUAGE
bib_key	9001601
keyword	STUART
bib_key	9001601
keyword	BERG
bib_key	9001601
keyword	FLEXNER
bib_key	9001601
keyword	EDITOR
bib_key	9001601
keyword	IN
bib_key	9001601
keyword	CHIEF

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
bib_key	9001601
keyword	ND
bib_key	9001601
keyword	ED
bib_key	9001601
keyword	NEW
bib_key	9001601
keyword	YORK
bib_key	9001601
keyword	RANDOM
bib_key	9001601
keyword	HOUSE
bib_key	**9005014**
keyword	THE
bib_key	9005014
keyword	NEW
bib_key	9005014
keyword	ENCYCLOPAEDIA
bib_key	9005014
keyword	BRITANNICA
bib_key	9005014
keyword	ED
bib_key	9005014
keyword	CHICAGO
bib_key	9005014
keyword	ENCYCLOPAEDIA
bib_key	9005014
keyword	BRITANNICA
bib_key	9005014
keyword	INC
keywords for other citations are omitted for the sake of brevity	
FORMER_GEOGKEYS	
new_key	
master_key	
collapse_key	
FORMER_GEOGLINKS	
new_key	
collapse_key	
old_link_key	

SAMPLE RECORD 2: Relational Files Format

Hawaii (United States, North and Central America) state

COLUMN	*EXAMPLE*
FORMER_BIBKEYS	
new_key	
master_key	
collapse_key	
FORMER_BIBLINKS	
new_key	
collapse_key	
old_link_key	

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
GEOG_MAIN	
geog_key	**7000457**
descr_note	Was Roman military center at head of navigation on Arno river & on Casian Way; escaped capture by Goths 5th cen.; was thriving center by 12th cen.; torn by medieval Guelph/Ghibelline civil strife; ruled by Medici family from 1434.
lat	43 47
lat_dir	N
long	011 15
long_dir	E
lat_num	43.783
long_num	011.250
GEOG_PARENTS	
geog_key	**7000457**
parent_key	7003163
pref_flag	P
GEOG_NAME	
geog_key	**7000457**
placename_id	45063
seqn	0
place_name	Firenze
sort_name	FIRENZE
pref_flag	P
lang	V
age_flag	C
other_flags	N
geog_key	**7000457**
placename_id	45064
seqn	1
place_name	Florence
sort_name	FLORENCE
pref_flag	N
lang	O
age_flag	C
other_flags	E
geog_key	**7000457**
placename_id	139941

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
seqn	2
place_name	Florencia
sort_name	FLORENCIA
pref_flag	N
lang	O
age_flag	C
other_flags	N
geog_key	**7000457**
placename_id	139942
seqn	3
place_name	Florenz
sort_name	FLORENZ
pref_flag	N
lang	O
age_flag	C
other_flags	N
geog_key	**7000457**
placename_id	165290
seqn	4
place_name	Fiorenza
sort_name	FIORENZA
pref_flag	N
lang	V
age_flag	H
other_flags	N
geog_key	**7000457**
placename_id	164779
seqn	5
place_name	Florentia
sort_name	FLORENTIA
pref_flag	N
lang	V
age_flag	H
other_flags	N
NAME_DETAIL	
geog_key	**7000457**
placename_id	**45063**
bib_key	9006303
bib_page	
contrib	BHA
display_date	
start_date	

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
end_date	
geog_key	7000457
placename_id	45063
bib_key	9004757
bib_page	62 ff.
contrib	FDA
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	45063
bib_key	9006449
bib_page	
contrib	GRLPA
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	45063
bib_key	9006037
bib_page	66
contrib	VP
display_date	
start_date	
end_date	
geog_key	**7000457**
placename_id	**45064**
bib_key	9006449
bib_page	
contrib	FDA
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	45064
bib_key	9006449
bib_page	
contrib	GRLPA
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	45064
bib_key	9006447

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
bib_page	I, 296
contrib	VP
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	45064
bib_key	9005014
bib_page	IV, 838
contrib	VP
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	45064
bib_key	9006267
bib_page	400
contrib	VP
display_date	
start_date	
end_date	
geog_key	**7000457**
placename_id	**139941**
bib_key	9006551
bib_page	317
contrib	VP
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	139941
bib_key	9006541
bib_page	I-56
contrib	VP
display_date	
start_date	
end_date	
geog_key	**7000457**
placename_id	**139942**
bib_key	9006561
bib_page	
contrib	VP
display_date	
start_date	

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
end_date	
geog_key	**7000457**
placename_id	**165290**
bib_key	9004757
bib_page	14
contrib	VP
display_date	medieval
start_date	1100
end_date	1700
geog_key	**7000457**
placename_id	**164779**
bib_key	9006339
bib_page	
contrib	GRLPA
display_date	
start_date	
end_date	
geog_key	7000457
placename_id	164779
bib_key	9006339
bib_page	
contrib	VP
display_date	name of Roman colony on N bank of Arno
start_date	-100
end_date	1500
geog_key	7000457
placename_id	164779
bib_key	9006548
bib_page	343
contrib	VP
display_date	
start_date	
end_date	

GEOG_PLACETYPE

geog_key	**7000457**
placetype_code	83002
seqn	0
pref_flag	P
age_flag	C
geog_key	**7000457**
placetype_code	83040
seqn	1

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83112
seqn	2
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83115
seqn	3
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	81112
seqn	4
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83433
seqn	5
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83351
seqn	6
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83154
seqn	7
pref_flag	N
age_flag	C
geog_key	7000457
placetype_code	83371
seqn	8
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83505
seqn	9
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83131

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
seqn	10
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83506
seqn	11
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83360
seqn	12
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83321
seqn	13
pref_flag	N
age_flag	C
geog_key	**7000457**
placetype_code	83110
seqn	14
pref_flag	N
age_flag	H
geog_key	**7000457**
placetype_code	83045
seqn	15
pref_flag	N
age_flag	H

PLACETYPE_DATES

geog_key	**7000457**
placetype_code	**83002**
display_date	site of ancient settlement, later founded as colony by Romans in 1st cen. BC, at foot of Etruscan hill town Fiesole
start_date	-1000
end_date	9999
geog_key	**7000457**
placetype_code	**83433**
display_date	developed on both sides of the Arno river, is subject to periodic flooding; most bridges were destroyed in WWII
start_date	-100
end_date	9999
geog_key	**7000457**

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
placetype_code	**83154**
display_date	bishops were established here early; today is famed for huge cathedral & baptistry & for numerous other churches
start_date	700
end_date	9999
geog_key	**7000457**
placetype_code	**83371**
display_date	factories located in suburbs produce precision instruments & other items
start_date	1800
end_date	9999
geog_key	**7000457**
placetype_code	**83505**
display_date	noted as great center of art & literature since Middle Ages, especially fourished 14th-16th cen.
start_date	1150
end_date	9999
geog_key	**7000457**
placetype_code	**83131**
display_date	for road & river traffic since Roman times, now is also a major hub for rail traffic
start_date	-100
end_date	9999
geog_key	**7000457**
placetype_code	**83506**
display_date	famed for traditional products, including textiles, glass, ceramics, metal wares, leatherwork, art reproductions & furniture
start_date	-100
end_date	9999
geog_key	**7000457**
placetype_code	**83321**
display_date	Florentines were paramount bankers in Europe by 15th cen.
start_date	1400
end_date	9999
geog_key	**7000457**
placetype_code	**83110**
display_date	of duchy of Tuscany
start_date	1500
end_date	1860
BIBLIO	
bib_key	**9004757**
bib_brief	Companion Guide: Florence (1979)
bib_brief_sort	COMPANIONGUIDEFLORENCE
citation	Borsook, Eve. The Companion Guide to Florence. 4th ed. London: Collins,

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
	1979.
bib_key	**9005014**
bib_brief	Encyclop$70aedia Britannica (1988)
bib_brief_sort	ENCYCLOPAEDIABRITANNICA
citation	The New Encyclop$70aedia Britannica. 15th ed. Chicago: Encyclop$70aedia Britannica Inc., 1988.
bib_key	**9006037**
bib_brief	Times Atlas of the World (1992)
bib_brief_sort	TIMESATLASOFTHEWORLD
citation	Times Atlas of the World. 9th comprehensive edition. New York: Times Books, 1992.
bib_key	**9006267**
bib_brief	Webster's Geographical Dictionary (1988)
bib_brief_sort	WEBSTERSGEOGRAPHICALDICTIONARY
citation	Webster's New Geographical Dictionary. Sprinfield, MA: Merriam-Webster, 1988.
bib_key	**9006303**
bib_brief	Columbia Lippincott Gazetteer (1961)
bib_brief_sort	COLUMBIALIPPINCOTTGAZETTEER
citation	Columbia Lippincott Gazetteer of the World. Edited by Leon E. Seltzer. Morningside Heights, NY: Columbia University Press, 1961.
bib_key	**9006339**
bib_brief	Princeton Encyclopedia (1979)
bib_brief_sort	PRINCETONENCYCLOPEDIA
citation	The Princeton Encyclopedia of Classical Sites. 2nd ed. Princeton, NJ: Princeton University Press, 1979.
bib_key	**9006447**
bib_brief	Canby, Historic Places (1984)
bib_brief_sort	CANBYHISTORICPLACES
citation	Canby, Courtlandt. The Encyclopedia of Historic Places. New York: Facts on File Publications, 1984.
bib_key	**9006449**
bib_brief	Webster's Geographical Dictionary (1984)
bib_brief_sort	WEBSTERSGEOGRAPHICALDICTIONARY
citation	Webster's New Geographical Dictionary. Springfield, MA: Merriam-Webster, 1984.
bib_key	**9006541**
bib_brief	Rand McNally Atlas (1994)
bib_brief_sort	RANDMCNALLYATLAS
citation	Rand McNally and Company. The New International Atlas. 25th Anniversary Edition. Chicago: Rand McNally, 1994.
bib_key	**9006548**
bib_brief	Times Atlas of World History (1994)
bib_brief_sort	TIMESATLASOFWORLDHISTORY
citation	The Times Atlas of World History. Edited by Geoffrey Barraclough. 4th

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
	edition editied by Geoffrey Parker. Maplewood, NJ: Hammond Incorporated, 1994.
bib_key	**9006551**
bib_brief	Cassell's Spanish Dictionary (1978)
bib_brief_sort	CASSELLSSPANISHDICTIONARY
citation	Cassell's Spanish Dictionary. Completely revised by Anthony Gooch. New York: Macmillan Publishing Co., 1978.
bib_key	**9006561**
bib_brief	NIMA, GEOnet Names Server (1996)
bib_brief_sort	NIMAGEONETNAMESSERVER
citation	United States National Imagery and Mapping Agency. GEO-net Names Server [online database]. Edited by the U.S. Board on Geographic Names. Washington: National Imagery and Mapping Agency, 1996. <http://164.214.2.53/gns/html/index.html>.

PLACETYPE_LIST

placetype_code	**83002**
placetype_search	83002
placetype	inhabited place
placetype_code	**83040**
placetype_search	83040
placetype	city
placetype_code	**83112**
placetype_search	83112
placetype	regional capital
placetype_code	**83115**
placetype_search	83115
placetype	provincial capital
placetype_code	**81112**
placetype_search	81112
placetype	commune (administrative)
placetype_code	**83433**
placetype_search	83433
placetype	river settlement
placetype_code	**83351**
placetype_search	83351
placetype	tourist center
placetype_code	**83154**
placetype_search	83154
placetype	archiepiscopal see
placetype_code	**83371**
placetype_search	83371
placetype	industrial center
placetype_code	**83505**

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
placetype_search	83505
placetype	cultural center
placetype_code	**83131**
placetype_search	83131
placetype	transportation center
placetype_code	**83506**
placetype_search	83506
placetype	craftsman center
placetype_code	**83360**
placetype_search	83360
placetype	educational center
placetype_code	**83321**
placetype_search	83321
placetype	financial center
placetype_code	**83110**
placetype_search	83110
placetype	capital
placetype_code	**83045**
placetype_search	83045
placetype	municipium

NAME_KEYWORDS

geog_key	**7000457**
placename_id	**45063**
keyword	FIRENZE
geog_key	**7000457**
placename_id	**45064**
keyword	FLORENCE
geog_key	**7000457**
placename_id	**139941**
keyword	FLORENCIA
geog_key	**7000457**
placename_id	**139942**
keyword	FLORENZ
geog_key	**7000457**
placename_id	**165290**
keyword	FIORENZA
geog_key	**7000457**
placename_id	**164779**
keyword	FLORENTIA

BIB_KEYWORDS

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
bib_key	**9004757**
keyword	BORSOOK
bib_key	9004757
keyword	EVE
bib_key	9004757
keyword	THE
bib_key	9004757
keyword	COMPANION
bib_key	9004757
keyword	GUIDE
bib_key	9004757
keyword	TO
bib_key	9004757
keyword	FLORENCE
bib_key	9004757
keyword	ED
bib_key	9004757
keyword	LONDON
bib_key	9004757
keyword	COLLINS
bib_key	**9005014**
keyword	THE
bib_key	9005014
keyword	NEW
bib_key	9005014
keyword	ENCYCLOPAEDIA
bib_key	9005014
keyword	BRITANNICA
bib_key	9005014
keyword	ED
bib_key	9005014
keyword	CHICAGO
bib_key	9005014
keyword	ENCYCLOPAEDIA
bib_key	9005014
keyword	BRITANNICA
bib_key	9005014
keyword	INC
keywords for other citations are omitted for the sake of brevity	
FORMER_GEOGKEYS	
new_key	
master_key	

SAMPLE RECORD 3: Relational Files Format

Firenze (Firenze province, Italia, Europe) inhabited place

COLUMN	*EXAMPLE*
collapse_key	
FORMER_GEOGLINKS	
new_key	
collapse_key	
old_link_key	
FORMER_BIBKEYS	
new_key	
master_key	
collapse_key	
FORMER_BIBLINKS	
new_key	
collapse_key	
old_link_key	

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
***General Record Information**	
LEN	3955
STATUS=	n
TGNIDNO=	7000457
DATENT=	19980316
***Hierarchy for Display**	
DISPHIER=	Firenze (Firenze province, Toscana, Italia, Europe)
***Main Hierarchy**	
FACET=	Current World
PARENT=	7003163
PARPREF=	P
***Simple Hierarchy**	
CONTINT=	Europe
NATION=	Italia
FIRSTLEV=	Toscana
SECNDLEV=	Firenze province
***Preferred Name**	
PREFNAME=	Firenze
PNAFLAGS=	C,V
PNASORT=	FIRENZE
PNAID=	45063
***Name Detail**	
PNCONTR=	FDA
PNASOURC=	Companion Guide: Florence (1979)
PNASRCID=	9004757
PNAPAGE=	62 ff.
***Name Detail**	
PNCONTR=	VP
PNASOURC=	Times Atlas of the World (1992)
PNASRCID=	9006037
PNAPAGE=	66
***Name Detail**	
PNCONTR=	BHA
PNASOURC=	Columbia Lippincott Gazetteer (1961)
PNASRCID=	9006303
***Name Detail**	
PNCONTR=	GRLPA
PNASOURC=	Webster's Geographical Dictionary (1984)
PNASRCID=	9006449
***Alternate Names**	
ALTNAME=	Florence
ANAFLAGS=	C,O,ENG
ANASORT	**FLORENCE**

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
ANAID=	45064
ANASEQN=	1
***Name Detail**	
ANCONTR=	VP
ANASOURC=	Encyclop$70aedia Britannica (1988)
ANASRCID=	9005014
ANAPAGE=	IV, 838
***Name Detail**	
ANCONTR=	VP
ANASOURC=	Webster's Geographical Dictionary (1988)
ANASRCID=	9006267
ANAPAGE=	400
***Name Detail**	
ANCONTR=	VP
ANASOURC=	Canby, Historic Places (1984)
ANASRCID=	9006447
ANAPAGE=	I, 296
***Name Detail**	
ANCONTR=	FDA
ANASOURC=	Webster's Geographical Dictionary (1984)
ANASRCID=	9006449
***Name Detail**	
ANCONTR=	GRLPA
ANASOURC=	Webster's Geographical Dictionary (1984)
ANASRCID=	9006449
***Alternate Names**	
ALTNAME=	Florencia
ANAFLAGS=	C,O
ANASORT=	FLORENCIA
ANAID=	139941
ANASEQN=	2
***Name Detail**	
ANCONTR=	VP
ANASOURC=	Rand McNally Atlas (1994)
ANASRCID=	9006541
ANAPAGE=	I-56
ANASOURC=	Cassell's Spanish Dictionary (1978)
ANASRCID=	9006551
ANAPAGE=	317
***Alternate Names**	
ALTNAME=	Florenz
ANAFLAGS=	C,O
ANASORT=	FLORENZ
ANAID=	139942

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
ANASEQN=	3
***Name Detail**	
ANCONTR=	VP
ANASOURC=	NIMA, GEOnet Names Server (1996)
ANASRCID=	9006561
***Alternate Names**	
ALTNAME=	Fiorenza
ANAFLAGS=	H,V,N
ANASORT=	FIORENZA
ANAID=	165290
ANASEQN=	4
***Name Detail**	
ANCONTR=	VP
ANADSP=	medieval
ANASTRT=	1100
ANAEND=	1700
ANASOURC=	Companion Guide: Florence (1979)
ANASRCID=	9004757
ANAPAGE=	14
***Alternate Names**	
ALTNAME=	Florentia
ANAFLAGS=	H,V,N
ANASORT=	FLORENTIA
ANAID=	164779
ANASEQN=	5
***Name Detail**	
ANCONTR=	VP
ANADSP=	name of Roman colony on N bank of Arno
ANASTRT=	-100
ANAEND=	1500
***Name Detail**	
ANCONTR=	VP
ANASOURC=	Princeton Encyclopedia (1979)
ANASRCID=	9006339
ANAPAGE=	331
***Name Detail**	
ANCONTR=	GRLPA
ANASOURC=	Princeton Encyclopedia (1979)
ANASRCID=	9006339
***Name Detail**	
ANCONTR=	VP
ANASOURC=	Times Atlas of World History(1994)
ANASRCID=	9006548
ANAPAGE=	343

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
***Preferred Place Type**	
PREFPTP=	inhabited place
PPTCODE=	83002
PPTAGE=	C
***Dates**	
PPTPDSP=	site of ancient settlement, later founded as colony by Romans in 1st cen. BC, at foot of Etruscan hill town Fiesole
PPTPSTRT=	-1000
PPTPEND=	9999
***Other Place Types**	
OTHERPTP=	city
OPTCODE=	83040
OPTAGE=	C
OPTSEQN=	1
***Other Place Types**	
OTHERPTP=	regional capital
OPTCODE=	83112
OPTAGE=	C
OPTSEQN=	2
***Other Place Types**	
OTHERPTP=	provincial capital
OPTCODE =	83115
OPTAGE=	C
OPTSEQN=	3
***Other Place Types**	
OTHERPTP=	commune (administrative)
OPTCODE=	81112
OPTAGE=	C
OPTSEQN=	4
***Other Place Types**	
OTHERPTP=	river settlement
OPTCODE=	83433
OPTAGE=	C
OPTSEQN=	5
***Dates**	
OPTPDSP=	developed on both sides of the Arno river, is subject to periodic flooding; most bridges were destroyed in WW II
OPTPSTRT=	-100
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	tourist center
OPTCODE=	83351
OPTAGE=	C
OPTSEQN=	6

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
***Other Place Types**	
OTHERPTP=	archiepiscopal see
OPTCODE=	83154
OPTAGE=	C
OPTSEQN=	7
***Dates**	
OPTPDSP=	bishops were established here early; today is famed for huge cathedral & baptistry & for numerous other churches
OPTPSTRT=	700
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	industrial center
OPTCODE=	83371
OPTAGE=	C
OPTSEQN=	8
***Dates**	
OPTPDSP=	factories located in suburbs produce precision instruments & other items
OPTPSTRT=	1800
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	cultural center
OPTCODE=	83505
OPTAGE=	C
OPTSEQN=	9
***Dates**	
OPTPDSP=	noted as great center of art & literature since Middle Ages, especially flourished 14th-16th cen.
OPTPSTRT=	1150
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	transportation center
OPTCODE=	83131
OPTAGE=	C
OPTSEQN=	10
***Dates**	
OPTPDSP=	for road & river traffic since Roman times, now is also a major hub for rail traffic
OPTPSTRT=	-100
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	craftsman center
OPTCODE=	83506
OPTAGE=	C
OPTSEQN=	11

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
***Dates**	
OPTPDSP=	famed for traditional products, including textiles, glass, ceramics, metal wares, leatherwork, art reproductions & furniture
OPTPSTRT=	-1100
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	educational center
OPTCODE=	83360
OPTAGE=	C
OPTSEQN=	12
***Other Place Types**	
OTHERPTP=	financial center
OPTCODE=	83321
OPTAGE=	C
OPTSEQN=	13
***Dates**	
OPTPDSP=	Florentines were paramount bankers in Europe by 15th cen.
OPTPSTRT=	1400
OPTPEND=	9999
***Other Place Types**	
OTHERPTP=	capital
OPTCODE=	83110
OPTAGE=	C
OPTSEQN=	14
***Dates**	
OPTPDSP=	of duchy of Tuscany
OPTPSTRT=	1500
OPTPEND=	1860
***Other Place Types**	
OTHERPTP=	municipium
OPTCODE=	83045
OPTAGE=	C
OPTSEQN=	15
***Coordinates**	
LAT=	43 47
LATDIR=	N
LONG=	011 15
LONGDIR=	E
LATNUM=	43.783
LONGNUM=	11.250

SAMPLE RECORD 4: REC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
***Descriptive Note**	
DESCNOTE=	Was Roman military center at head of navigation on Arno river & on Cassian Way; escaped capture by Goths 5th cen.; was thriving center by 12th cen.; torn by medieval Guelph/Ghibelline civil strife; was an early republic; ruled by Medici family from 1434.
$$	

SAMPLE RECORD 5: USMARC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*

```
Len 03786
Sta n
Typ z
Lev
Bas 00577
Enc n
001 7000457¶
008 000602 n anznnbabn           b ana     d¶
151   ‡aFirenze (Firenze province, Toscana, Italia, Europe) inhabited
      place‡c43 47 N 011 15 E
      (7000457,scheme=TGN)‡dC,V,N‡eFIRENZE‡f45063¶
451   ‡aFlorence‡dC,O,ENG‡eFLORENCE‡f45064‡g1¶
451   ‡aFlorencia‡dC,O,N‡eFLORENCIA‡f139941‡g2¶
451   ‡aFlorenz‡dC,O,N‡eFLORENZ‡f139942‡g3¶
451   ‡aFiorenza‡dH,V,N‡eFIORENZA‡f165290‡g4¶
451   ‡aFlorentia‡dH,V,N‡eFLORENTIA‡f164779‡g5¶
451   ‡aFlorentine‡dC,O,ADJE‡eFLORENTINE‡f1159547‡g6¶
551   ‡wg‡aFirenze province (7003163,scheme=TGN)¶
670   ‡aColumbia Lippincott Gazetteer (1961) [BHA]‡bFirenze¶
670   ‡aCompanion Guide: Florence (1979) [FDA]‡bFirenze¶
670   ‡aWebster's Geographical Dictionary (1984) [GRLPA]‡bFirenze;
      Florence¶
670   ‡aTimes Atlas of the World (1992) [VP]‡bFirenze¶
670   ‡aWebster's Geographical Dictionary (1984) [FDA]‡bFlorence¶
670   ‡aCanby, Historic Places (1984) [VP]‡bFlorence¶
670   ‡aEncyclopædia Britannica (1988) [VP]‡bFlorence¶
670   ‡aWebster's Geographical Dictionary (1988) [VP-ENG]‡bFlorence¶
670   ‡aCassell's Spanish Dictionary (1978) [VP]‡bFlorencia¶
670   ‡aRand McNally Atlas (1994) [VP]‡bFlorencia¶
670   ‡aNIMA, GEOnet Names Server (1996) [VP]‡bFlorenz¶
670   ‡aCompanion Guide: Florence (1979) [VP]‡bFiorenza¶
670   ‡aPrinceton Encyclopedia (1979) [GRLPA]‡bFlorentia¶
670   ‡aTimes Atlas of World History (1994) [VP]‡bFlorentia¶
670   ‡aGRI Photo Study, Authority File (1998) [VP]‡bFlorentia¶
670   ‡aPrinceton Encyclopedia (1979) [VP]‡bFlorentia¶
670   ‡aOED2 on CD-ROM (1994) [VP-ADJE]‡bFlorentine¶
678   ‡iWas Roman military center at head of navigation on Arno river &
      on Cassian Way; escaped capture by Goths 5th cen.; was thriving
      center by 12th cen.; torn by medieval Guelph/Ghibelline civil
      strife; was an early republic; ruled by Medici family from 1434.¶
901   ‡ainhabited place‡b83002‡cC‡esite of ancient settlement, later
      founded as colony by Romans in 1st cen. BC, at foot of Etruscan
      hill town Fiesole‡f-1000‡g9999¶
902   ‡acity‡b83040‡cC‡d1¶
902   ‡aregional capital‡b83112‡cC‡d2¶
902   ‡aprovincial capital‡b83115‡cC‡d3¶
902   ‡acommune (administrative)‡b81112‡cC‡d4¶
902   ‡ariver settlement‡b83433‡cC‡d5‡edeveloped on both sides of the
      Arno river, is subject to periodic flooding; most bridges were
      destroyed in WW II‡f-100‡g9999¶
902   ‡atourist center‡b83351‡cC‡d6¶
902   ‡aarchiepiscopal see‡b83154‡cC‡d7‡ebishops were established here
      early; today is famed for huge cathedral & baptistry & for numerous
```

SAMPLE RECORD 5: USMARC Format

Firenze (Firenze province, Italia, Europe) inhabited place

FIELD LABEL	*EXAMPLE*
	other churches‡f700‡g9999¶
902	‡aindustrial center‡b83371‡cC‡d8‡efactories located in suburbs produce precision instruments & other items‡f1800‡g9999¶
902	‡acultural center‡b83505‡cC‡d9‡enoted as great center of art & literature since Middle Ages, especially flourished 14th-16th cen.‡f1150‡g9999¶
902	‡atransportation center‡b83131‡cC‡d10‡efor road & river traffic since Roman times, now is also a major hub for rail traffic‡f-100‡g9999¶
902	‡acraftsman center‡b83506‡cC‡d11‡efamed for traditional products, including textiles, glass, ceramics, metal wares, leatherwork, art reproductions & furniture‡f-1100‡g9999¶
902	‡aeducational center‡b83360‡cC‡d12¶
902	‡afinancial center‡b83321‡cC‡d13‡eFlorentines were leading bankers in Europe by 15th cen.‡f1400‡g9999¶
902	‡acapital‡b83110‡cH‡d14‡eof duchy of Tuscany‡f1500‡g1860¶
902	‡amunicipium‡b83045‡cH‡d15¶
903	‡a43 47‡bN‡c011 15‡dE‡e43.783‡f11.250¶
905	‡aEurope‡bItalia‡cToscana‡dFirenze province¶