

Feature ID Assigning Procedure

--- Version approved by CCOG and GNBC members, May 12, 2009 and August 6, 2009 respectively. ---

Introduction

The Canadian Geographical Names Data Base (CGNDB) described at <http://geonames.nrcan.gc.ca> is a national repository of names of places and geographical features officially adopted or recognized by the Geographical Names Board of Canada (GNBC). In the CGNDB, a feature that crosses provincial or territorial boundaries has more than one CGNDB record. The name is repeated in two or more records, one for each province or territory that it lies in, because names are approved by each provincial and territorial names authority, and the metadata (such as approval dates) may differ. Similarly, a feature that has two official names, such as a pan-Canadian name, has two or more CGNDB records, one in each official language, for each province or territory that it lies in. In addition, a feature that changes its name or generic code gets additional records, one for each change. A CGNDB Key is associated with each record. Hence, in the CGNDB, a feature may be associated with many names (sometimes repeated) and more than one CGNDB Key.

To easily identify each feature, it was proposed that a Feature Identifier (Feature ID or FID) be created and stored in the CGNDB and provided through the Canadian Geographical Names Service (see <http://cgns.nrcan.gc.ca>). That FID would remain associated with an entity regardless of name changes for the entity. The FID is to be unique for any named feature in Canada. When a feature name is replaced by a new name both the previous name and the new name have the same FID. Thus, it is possible to refer to previous names of the feature.

This document defines the FID and describes the procedure to assign an FID.

Cartographic context

In the context of cartographic data, a feature may be represented at various map scales or various map projections. The various representations may not be identical (visually or digitally), even though they are the same feature. At the same scale, two maps from different sources, or different editions of a map, may display the same feature but represent it differently, if not visually, then digitally. The intent of the FID is to link these various sources together also. The FID identifies delineations obtained from various sources of cartographic data as being the same feature.

In summary, a map feature could be represented differently depending on sources, scales, projections or editions, and each representation would have a different Geometry ID, but keep the same Feature ID (see Table 1). In other words, the intent of the FID is to be able to identify the global extent and limits of a feature. The Feature ID could have been called the Extent ID. When many Geometry IDs are linked to the same FID, the implication is that all these geometries define the same feature, though at various degrees of detail (scales), various shapes (projections), and various qualities (accuracies).

Even though a feature may be described by multiple names and multiple geometries, they all get linked together if the FID is used. For example, in Table 1, a feature has several related names (successively changed names) and, so, is assigned many CGNDB keys. However, the FID is the same for each of these names.

Again in Table 1, a feature has been digitized by several agencies, at various scales, on various map projections and on several map editions, each version having a Geometry ID but the same FID. Table 1 demonstrates that a FID links “many versions of a toponym” to “many versions of its digital delineation”.

Table 1. Demonstrating the use of the Feature ID (FID).

CGNDB key	Feature ID	Geometry ID	Feature ID
A (lake)	A	P (1:1M)	A
B (new name)	A	Q (1:7.5M)	A
C (French name)	A	R (1:50K)	A
D (English name)	A	S (1:50K) v.2	A
Etc.	A	T (1:20K)	A
F (another lake)	B	U (1:50K)	B

Implemented as a universal unique identifier (UUID)

The Feature ID was implemented as a universal unique identifier (UUID) also called a global unique identifier (GUID). The Feature ID is 32 characters long (such as 6978bbe2ba3611d892e2080020a0f4c9). The algorithm to generate a UUID is freely available on the Internet and is well known by cartographic agencies.

Operational assumptions:

- All groups assigning/maintaining feature identifiers (FIDs) must be capable of generating FIDs, which are universally unique (UUIDs).
- FIDs may be assigned to both named and unnamed features.
- There is a one to many relationship between one FID and many CGNDB Keys.
- Name changes will not affect FID assignments.
- Geometric delineations of a given feature at various scales (e.g., atlas maps, topographic maps) will share the same FID.

Within the project called the National Hydrographic Network (NHN), the members of Canadian Council on Geomatics (CCOG) have adopted the UUID of 32 characters to identify hydrographic features. Here, it is suggested that a 32-character UUID be also the standard to identify all named features adopted by the GNBC.

International standard

The Feature ID was chosen as a UUID because it is an ISO standard and it is available on many computers by simply typing "uuidgen". If not available on your computer, your IT people can install it. It is freely available on the web. Most cartographic agencies use the UUID as part of their digital cartographic database. For example, typing "uuidgen", displays say, 6978bbe2-ba36-11d8-92e2-080020a0f4c9. This should be copied and pasted into the database field but without the hyphens. In a production environment, the UUID is generated and inserted in the database automatically.

FID lifetime

The following rules were taken or adapted from *National Vector Data Identification Rules*, NRCan, 2003.

- The Identifier is permanently assigned.
- A UUID is assigned to an entity when it is created and is stable over the entity's entire life span. The FID of a deleted record cannot be reassigned.
- The UUID provides flexibility while working with several partners. The well-known UUID algorithm may be used by any data user to modify the data and add a new FID. **FIDs should only be generated by authorized organizations.** Specific care must be given to FIDs. These FIDs will eventually allow data synchronization between organizations. Data users must ensure that they make **no alterations whatsoever to the FID values.** Modifications would render FIDs useless for data synchronization.

Initial creation of FIDs

The initial creation of FIDs and their association to "named" geographical features were performed by the Toponymy Applications Section, Centre for Topographic Information, Natural Resources Canada, and made available in December 2004.

During this initial creation, Aboriginal, French, and English names for the same feature were assigned the same FID. Features that cross provincial or territorial boundaries and have a name adopted by each jurisdiction were assigned the same FID.

FIDs were generated only for official names, stored in the CGNDB and are available on the CGNS. No effort was made to assign a FID to rescinded or deleted/withdrawn names.

GNBC members - Procedure for assigning an FID

When a new name is approved and becomes official, an FID must be assigned. Here are the steps to follow for the assigning of an FID.

1. Verify if the feature is related to an existing named feature, thereafter called “Related Name”. Refer to Table 2 to see the possible cases.
2. If it is not related to an existing named feature, a new FID is generated. Skip steps 3 and 4.
3. If it is related to an existing feature, the new toponym is assigned the FID of the Related Name, and the Synonym Key of the Related Name (see definition of Synonym Key in the Appendix).
4. If it is related to an existing named feature without a Synonym Key, assign the same Synonym Key to both records, the Synonym Key being the alphabetically lower of the two CGNDB keys.

Notes:

- (i) Even if the extent is represented using another map scale, or another map version, the same FID is assigned (when the intent is to identify the same extent on the ground).
- (ii) An FID is usually assigned to an entire feature, e.g., the whole St. Lawrence River (not part of it as the CGNDB key does. For example, each portion of the river in Ontario and Quebec, has a CGNDB key assigned to it). There are a few exceptions such as the Assiniboine River in Manitoba (see Table 2, 6th type).

CCOG members - Procedure for assigning an FID

The procedure to decide on assigning to a geographical feature a new FID or an existing FID is the same as for the GNBC members above, but without the requirement to input the Synonym Key.

When a feature is already named: CCOG members must obtain the existing FIDs from either their local GNBC member, the CGNS or Geobase. This is essential, because all named features in Canada have already a National FID published, and this data is in widespread use. GeoBase publishes FIDs for named features in the NHN layer, as well as for the Names layer, and the Atlas of Canada has already assigned FIDs to their delineations. GNBC members have been required to publish National FIDs for named features, through the CGNS Web site, in order to supply FIDs to the NHN and Atlas projects.

Table 2. Example of features having more than one CGNDB key, but assigned the same FID.

<i>Type</i>	<i>Examples</i>	<i>Notes / Comments</i>
1. Pan-Canadian feature	Lake Winnipeg / Lac Winnipeg (MB)	Pan-Canadian features have two names, one for each official language. A pan-Canadian feature may not have an alternate language version. E.g., Alberta, the name is the same in both languages, and so only one CGNDB record exists.
2. Feature with a name in Canada's two official languages – includes National Parks, DND establishments, etc.	Cassie Cape / Cap-des-Caissie (a feature with two official names) (NB) Kouchibouguac National Park / Kouchibouguac, Parc national (NB)	Not like pan-Canadian names, these names are identified by A7 status, and cross-referenced through the related names table (i.e., the Related Key in the CGNS, see the Appendix).
3. Feature with different names approved by different jurisdictions	Mont d'Iberville (QC) / Mount Caubvick (NL) Kitchen Brook (NB) / Cours d'eau Courchesne (QC)	A few features have different names in adjacent provinces.
4. Feature crossing provincial/territorial boundaries	Kicking Horse Pass (AB/BC) Flin Flon (SK/MB)	There are multiple records in the CGNDB for these features – one for each province.
5. Feature with a former name or other cross-reference names	Bytown – former name for Ottawa (ON) Little Oven – former name for Le Petit Four (QC)	Quebec handles creation and updates of name records differently. This issue will have to be addressed in future uploads when the assigning of FIDs is implemented. (cross-references for B4 and B12 status only)
6. Feature with Manitoba Equivalent names, Ontario Alternate names, or others that may fit this group in the future	Assiniboine River / Rivière Assiniboine	Sometimes only portions of a feature may have a bilingual name. E.g., Assiniboine River is also named Rivière Assiniboine, but <u>only</u> where it flows through a French Language Services Area. The Assiniboine River is assigned many FIDs.
7. Feature with multiple names	PrinceAlbert Sound / Kangiryuaq (NT)	Some features have multiple names chosen among the many official languages of a Territory or Province.
8. Feature in two of the above categories, e.g., a pan-Canadian feature that crosses boundaries	Wood Buffalo National Park / Parc national Wood Buffalo (AB/NT) Gulf of St. Lawrence / Golfe du Saint-Laurent (QC, NB, NS, PE, NL)	An FID is assigned to an entire feature.

Link to digital delineation

Geographical names are linked to the digital delineations stored in geospatial databases (through their FID). Meanwhile, as usual, the delineation for all new decisions will be provided on paper maps, digital files, or GIS files.

Providing the digital delineations through an Open Geospatial Consortium Web Mapping Service and Web Feature Service (OGC WMS and WFS) site (see <<http://www.opengeospatial.org>>) is the objective of the GNBC members and the Canadian Geographical Names Service (see <<http://cgns.nrcan.gc.ca>>).

Deletion of a Feature ID

As usual, a record and an FID will not be deleted when a feature no longer exists, or when a name is no longer official; its delineation shall remain in a spatial database (for historical or legal purposes). When the name is no longer official, the record remains in the CGNDB and CGNS but with a change of status to “rescinded”. See previous paragraph “FID lifetime”.

Searching previous names

The Feature ID enables retrieval of the previous names for the same extent.

Extent with names in multiple official languages

Pan-Canadian names, Manitoba Equivalent names, Ontario Alternate names, national parks, Indian reserves, etc., have French and English names for the same extent, and thus, the FID links these official names. In addition to English and French names, the territories and some provinces recognize multiple aboriginal names for some features. All the names with the same extent are assigned the same FID.

Bibliography

NRCan, 2003. *National Vector Data Identification Rules*, Natural Resources Canada, Geomatics Canada, Centre for Topographic Information, Sherbrooke, Quebec, Canada, 8 pages.

Acronyms

CCOG - Canadian Council on Geomatics (see <<http://www.geobase.ca>>)

CGNDB - Canadian Geographical Names Data Base
(see <http://geonames.nrcan.gc.ca/search/search_e.php>)

CGNS - Canadian Geographical Names Service (see <http://gnss.nrcan.gc.ca/index_e.html>)

GNBC - Geographical Names Board of Canada
(see <http://geonames.nrcan.gc.ca/index_e.php>)

GNSS - Geographical Names Search Service of the CGNS (see <<http://gnss.nrcan.gc.ca/gnss-srt/searchName.jsp?language=en>>)

Appendix

Related database fields

In addition to the Feature ID, five new fields have been added to the CGNS database.

- 1- **Synonym Key** helps manage FIDs. It allows retrieval of all the names with the same extent. When an extent has more than one name, every name is given the same FID and the same Synonym Key. When a record has no Synonym Key, we know there is no other name with the same extent. To simplify its maintenance, the Synonym Key is assigned the alphabetically lowest CGNDB key of all names with the same extent.
- 2- **Related Key** allows retrieval of (i) all the previous names (historical, rescinded), (ii) the bilingual names (official), (iii) the variant names, and (iv) the withdrawn names (for various reasons, including being mistyped), even if the extents differ. The Related Key fulfills the role of the former cross-reference names.
- 3- **History** allows retrieval of either of (i), (ii), (iii), or (iv) above. The only keywords accepted in the “History” field are: 1-Alternate Name, 2-Official Name, 3-Former Name, 4-Variant Name. These numbers were chosen to allow sorting of the toponyms in a meaningful way - from important to less important.
- 4- **List** allows retrieval of specific lists often referred to by GNBC members and users. Such lists include pan-Canadian names, Manitoba Equivalent names, Ontario Alternate names, Parks Canada names, INAC names, DND names, etc. Others may be created. Possible keywords are: PanCan, equiMB, altON, IR, Parks, DND, etc.
- 5- **Language** used together with the Related Key, Synonym Key, FID, History, or List, allows retrieval of or identification of the French names, English names, or Aboriginal names. It also allows an application to automatically display the French name, the English name, or the Aboriginal name, depending on the context. The codes accepted in the “Language” field are the three letter international codes of the ISO 639-3 identifying languages: eng (English), fre (French), crm (Cri, Moose), crk (Cri, Plains), etc. If the field is null, the feature doesn't have another official name in another language. If the field is not null, then the feature has at least one other official name in another language.

Note:

Provinces or territories who contribute to the CGNS, but maintain their own electronic geographical names systems, are not obligated to load or maintain the new fields listed above.