National Vector Data – Identification Rules And Change Management

Edition 1.0

2012-10-15

Natural Resources Canada
Mapping Information Branch
Centre for Topographic Information – Sherbrooke
2144, King Street West, Suite 010
Sherbrooke, Quebec Canada
J1J 2E8

Telephone: +01-819-564-4857 / 1-800-661-2638 (Canada and USA)
Fax: +01-819-564-5698
E-mail: supportGeoBase@NRCan.gc.ca
URL: http://www.geobase.ca

Copyright Notice

REVISION HISTORY

Date	Version	Description
2012-09-27	1.0	Original version combining the documents: *National Vector Data – Identification Rules, Edition 3.0* and *National Vector Data – Change Management, Edition 4.0* both dated November 2010. Addition of the National Railway Network (NRWN) theme.

TABLE OF CONTENTS

ABBREVIATIONS

GUID	Globally Unique Identifiers
ID	Identifier
IEEE	Institute of Electrical and Electronics Engineers – USA
MUNI	Municipal Boundaries
NHN	National Hydrographic Network
NID	National Identifier
NRCan	Natural Resources Canada
NRN	National Road Network
NRWN	National Railway Network
NSDI	National Spatial Data Infrastructure – USA
NVD	National Vector Data
UUID	Universal Unique Identifier

TERMS AND DEFINITIONS

IEEE 802 Address – *[Adresse IEEE 802]*

Digital address proposed by Committee 802 of the Institute of Electrical and Electronics Engineers (USA) that is one basic component of UUIDs.

National Identifier (NID) – *[Identifiant national (IDN)]*

Unique national identifier attributed to each NVD object. Each NID is a 32 lower case character string (without blanks) representing a UUID generated at random.

National Vector Data (NVD) – *[Données vectorielles nationales (DVN)]*

The NVD is intended to be the best vector representation of interest phenomena that is broadly available across Canada. Changes occur when a more up-to-date or more accurate representation than the previous one becomes available. The National Road Network (NRN), the National Hydrographic Network (NHN), the Municipal Boundaries (MUNI), and the National Railway Network (NRWN) are examples of NVD.

Object – *[Objet]*

Information technology model of a real world phenomenon.

~~Unique~~ Universally Unique Identifier (UUID) – *[Identifiant universel unique (IDUU)]*

Unique identifier within a universe defined in an application domain. UUIDs are those proposed by the ISO 19118 standard: *Geographic information – Encoding*. They are represented by a 32-character hexadecimal string.

The definition and method used for the generation of a UUID is defined in the current document.

1 BACKGROUND

This documentation is intended for National Vector Data (NVD) users and providers on GeoBase and aims at describing and standardizing data maintenance.

The NVD is intended to be the best vector representation of interest phenomena that is broadly available across Canada. These are vector data layers with common characteristics (such as a unique and permanent identifier for each of its basic constituent) and therefore share the same data specifications.

NVD present the possibility to be updated on a regular basis by various actors. To allow this, we use update mechanisms among NVD partners. These mechanisms are founded on change management principles established on two basic concepts:

- Identification rules for objects composing NVD
- Change management rules for the data.

The **identification rules** for NVD are presented in the first section of the current document while the **change management rules** are presented in the second section.
Or
The **identification rules** for NVD and the **change management** rules used to be presented in two separate documents. For the first time, they have now been grouped in the current document as section 2 and section 3 respectively.

Objects are information technology models of real world phenomena. If the real world phenomenon of interest is considered to be a geographic entity, then the term feature is typically used to describe the model of the entity. In other words, a feature is a type of object.

2 INTRODUCTION TO NVD IDENTIFICATION RULES (SECTION 2.1)

Objects that depict real-world phenomena vary over time, either by their description, by their precision, or by the instruments and methods involved in their initial acquisition or eventual change. It is therefore possible that more than one representation of the same phenomena exist. Our goal is to build and maintain ***only one*** representation of each NVD product (the National Road Network, for instance) while allowing that the representations originate from multiple sources (the best source available). Identifiers play a fundamental role in ensuring proper sharing and exchange of objects that have been modified at the source and that have already been provided to users. The implementation of a standard for the permanent identification of a phenomenon and its application must achieve two primary objectives:

- Facilitate the management and distribution of object changes in an incremental manner;
- Facilitate the conflation process, if necessary.

Every occurrence of NVD basic features must be uniquely and permanently identified. As an example, each geometric object in the NRN: *Road Elemen*t, *Ferry Connection*, and *Junctio*n that describes specific characteristics of the linear network must also be uniquely identified.

The following section presents the definition and description of NVD identifiers and the section after that presents how NVD identifiers are generated.

3 DEFINITION AND DESCRIPTION OF IDENTIFIERS (Section 2.2)

According to [*Bédard et al.*, 2000], the identifiers must be permanently assigned or persistent. To ensure their stability, the assigned IDs must be insignificant (inconsequential) in their expression [1] (free translation). In other words, the IDs must not contain any information relative to the data. Past experience has demonstrated that encapsulating information within the ID can cause ID modification without any real change ever having occurred in the data.

Within the ISO 19118: *Geographic Information – Encoding* [2] standard, there is a definition for UUIDs that corresponds exactly to the fundamental requirements sought after for NVD:

> “An application domain defines a universe and an identification scheme called universal unique identifiers (UUIDs). A UUID is assigned to an object when it is created and is stable over the object's entire life span. The UUID of a deleted object cannot be used again. UUIDs are required for long-term distributed data management and for implementing update mechanisms. These identifiers are also called persistent identifiers. A special name server may be used to resolve persistent identifiers. The identifiers are unique within a well-defined limited universe defined by an application domain.”
>
> [New definition in ISO 19118 : 2011 instead of 2005]
>
> “An application domain defines a universe and an identification convention called domain unique identifiers (DUIDs). A DUID is assigned to an object when it is created and is stable over the entire life span of the object. A DUID of a deleted object shall not be used again. A DUID is mostly used for implementation reasons. DUIDs are required for long term distributed data management and for realizing update mechanisms. These identifiers are also called persistent identifiers. A special name server may be used to resolve persistent identifiers. The identifiers shall be unique within a well-defined, limited universe defined by an application domain.”
>
> +
>
> NOTE The term “universal unique identifier” has the same meaning as “domain unique identifier”. This annex uses it and “UUID” as its abbreviation to keep the compatibility for the communities in use. The application of UUIDs in this International Standard is not limited to a particular mathematical basis, like ISO/IEC 11578 or ISO/IEC 9834-8. The “universal unique identifier” is not actually qualified by application domains, in contrast to the “domain unique identifier”, which is qualified by the application domain.
>
> EXAMPLE An application domain can want to use a two-component identifier. The first component is the domain name and the next component is an integer instance number. The components can be separated by a colon “:”. The instance numbers can be encoded in hex. There are no restrictions on the size of the instance number. Examples of two UUIDs in a domain called “example” are: “example:F23C30” and “example:FFFFFF12345A”.

The UUID consists of a 16-byte record and must be void of padding between fields. The hexadecimal values “a” to “f” must be lower case. The total size is 128 bits. For use as human-readable text, a UUID string representation (32 characters) is specified as a sequence of fields. The following string is a UUID example:

- 378a3917e824422cb25f268b8295da51

For more information on UUIDs: http://www.opengroup.org/onlinepubs/9629399/apdxa.htm#tagcjh_20

This ISO definition for identifiers has been adopted for “National Identifier (NID)” of the NVD. In other words, a NID is a 32 character lower case string representing a UUID generated randomly and void of padding between fields. The NID (and therefore UUID) generation mechanism is presented in the following section.

4 NVD IDENTIFIER GENERATION (Section 2.3)

ID uniqueness is one of the fundamental characteristics of NVD. The method retained for unique ID generation consists of using a unique ID generation algorithm[i] that can be used by data producers with no particular management of range and domain.

NID (and therefore UUID) generation does not require a registration authority for each single identifier. Instead, it requires a unique time value for each UUID generator (computer). This unique time value is specified as an IEEE 802 address, which is usually already applied to network-connected systems. This 48-bit address can be assigned based on an address block obtained through the IEEE registration authority (incorporated into the equipment). This UUID specification assumes the availability of an IEEE 802 address within all equipments that generate a UUID.

The algorithm described above provides producers the needed flexibility while working within a network of partners. The algorithm can be used by all closest to source data producers to modify the data and add a new NID when needed. ***NIDs should only be generated and assigned by authorized organizations*** (these include those who provide the data or the changes at the source). Specific care must be given to the management of NIDs. These NIDs will eventually allow for data synchronization between organizations. Data users must ensure that they make ***no alterations whatsoever to these NIDs*** value in order to ensure synchronization**.** Modifications to NIDs would render them useless for data synchronization.

The assignation and persistence rules of the NID are further explained in the following section of the current document.

[i] Readers wishing to use a standards-body definition of UUIDs/GUIDs should refer to: ISO/IEC 11578:1996 Information technology -- Open Systems Interconnection -- Remote Procedure Call *http://www.iso.org/iso/catalogue_detail.htm?csnumber=2229* *(not free)* or DCE 1.1: Remote Procedure Call Open Group Technical Standard Document Number C706, August 1997, 737 pages. (Supersedes C309 DCE: Remote Procedure Call 8/94, which was the basis for the ISO specification) http://www.opengroup.org/publications/catalog/c706.htm

5 INTRODUCTION FOR CHANGE MANAGEMENT (Section 3.1)

Many projects (or the literature) deal with update management and time modelling [see References 3, 4, 5]. The NVD change management model herein was developed in cooperation with the Centre for Research in Geomatics (CRG) at Laval University [see Reference 6] and has been defined to be consistent with the terminology used by the Open Geospatial Consortium.

Change management rules allow to characterize the evolution of features in order to identify any changes that may have occurred between two versions of a feature, whether successive or not. The discrepancies observed between two versions are referred to as the differential[ii]. In the context of the NVD, the purpose of change management is to facilitate synchronization of databases from various producing partners and customers based on current national views (see Figure 1: Evolution of the database in time).

The change management process must also make it possible to reconstitute the dataset as it was on a previous date. Depending upon the particular change management strategy used (as discussed below), the process may allow for time travel through the dataset, such that the state of individual features in the dataset can be ascertained at arbitrary times.

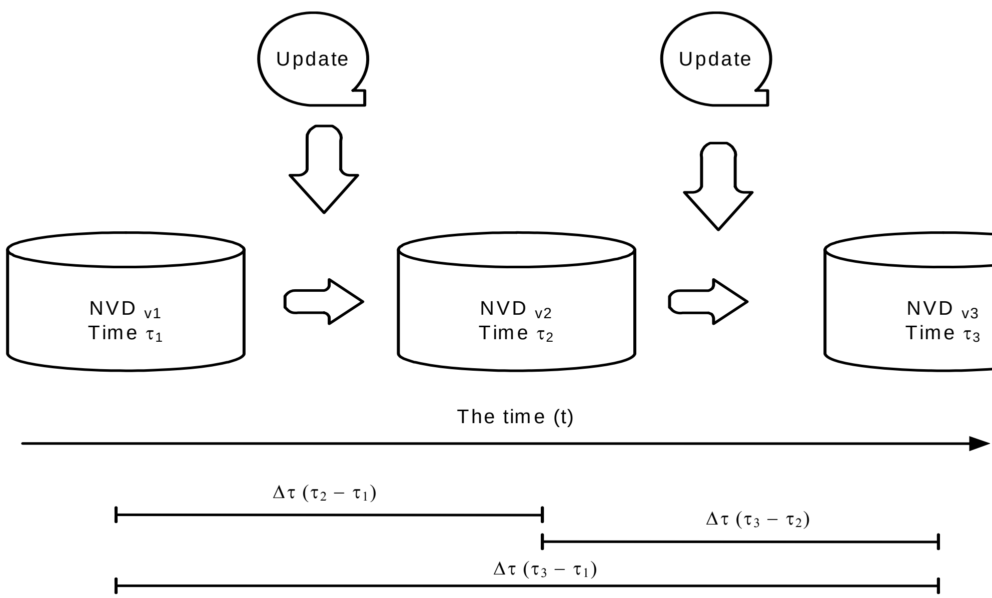


Figure 1: Evolution of the database in time

In the next sections, we will present the life cycle of an object (in other words, what a change is) as well as the types of effect on objects and we will provide examples of changes for each NVD product that can go through updating.

ii The differential corresponds to the set of differences observed between two landmarks of the territory [see reference 6].

6 OBJECT LIFE CYCLE (Section 3.2)

The NVD is intended to be the best vector representation of interest phenomena that is broadly available across Canada. Changes occur when a more up-to-date or more accurate representation than the previous one becomes available.

The effects on NVD data are established based on the previous representation. Data life cycle is therefore limited by two events. The cycle always begins with an "addition" (assignment of a new NID) and ends with "retirement". Between these two events, geometric or descriptive modification or confirmation of the previous state can occur, while maintaining the same NID. Data with the effects "addition," "geometric or descriptive modification", and "confirmation" are said to be *active* (or current) features. Features with the effect "retirement" are said to be *non active* (historical) data.

7 EFFECT TYPES ON OBJECTS (Section 3.3)

Various kinds of updates are recognized based on the *effect* of the update on the data. (These are analogous to transaction types in the database world.) Effects affect either the existence or the evolution of an object and can be classified as follows:

Addition (Existence)

When a new object has no geometric counterpart in the NVD, a new feature is *added*, which has a new NID.

Retirement (Existence)

When a feature no longer represents an entity, the feature is *retired*. The feature is removed from the current data while maintaining its NID.

Modification (Evolution)

A feature is said to be *modified* if one or more of its descriptive attributes or its geometric representation is changed. In this case the initial NID is preserved, as the new version will have the same NID value as the version it replaces. Two types of modification are possible.

Descriptive Modification

A descriptive modification occurs if one version of a feature is replaced by a second version with the same NID but with one or more differences in attribute values. For example, the surface of a specific road may have changed from “unpaved” to “paved”.

Geometric Modification

A geometric modification occurs if one version of a feature is replaced by a second version with the same NID but with a change in the object’s geometric representation.

Four types of geometric modification are currently defined within the NVD. Each has a certain level of complexity. In comparing two representations (old and new), the following change management methods are recognized:

- **First Method**

Comparing the vertices of an old and a new version of an object. If any vertex is different positionally, or if any vertex has been added or removed, the old representation of the object is retired and a new representation added. The new version has a new NID and no explicit reference to the original version is retained. Using this method means that **there is no tracking of geometric modifications**, making time travel through the data difficult or almost impossible.

- **Second Method**

Comparing the locations of the old and the new junctions. Two junctions always bound a network linear element. Any modification along a linear element (geometric representation) may occur between its junctions. These are treated as geometric modifications where the NID is preserved. If, for whatever reason, one of the old junctions located at one end of the network linear element has changed, then this network linear element is retired and a new one added with a new NID.

- **Third Method**

Comparing topological links of network linear elements. If the representation of the linear element junctions has maintained the same topological links (even if the junctions have moved and the network linear element geometry has been modified), then these changes are treated as a geometric modification and the network linear element and junctions maintain their respective NID.

- **Fourth Method**

The fourth method is the most lenient. The geometric representation of a feature may be redefined in any way, but the feature itself is considered to continue and will retain the same NID. For example, the boundaries of an Aboriginal land or a municipality may be modified or extended. In this case the old version of the feature is replaced by the new version with the same NID.

For any given NVD dataset, **only one method of geometric modification can apply**. It must be specified in the dataset metadata.

Confirmation (Evolution)

Addition, retirement, and modification all pertain to change. However, an entity may have been revisited, with no changes required to the description of its geometric or attribute properties. In such a case, there is *confirmation* of the existing description for the entity.

8 EXAMPLES OF CHANGE (Section 3.4)

8.1 National Road Network (NRN) Data (Section 3.4.1)

The following example is designed to illustrate NRN update management for better comprehension. Figure 2 : Example of Changes Following a NRN Updatebelow demonstrates the comparison between the original data and new data. In terms of geometry, a single Road Element (object 6) was added with respect to the original data. In terms of description, the type of surface of the Road Element (object 2) has changed from “unpaved” to “paved”.

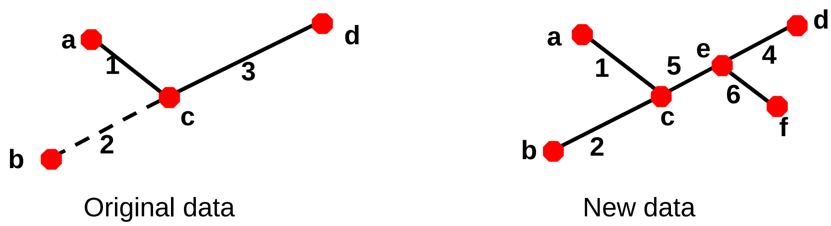


Figure 2 : Example of Changes Following a NRN Update

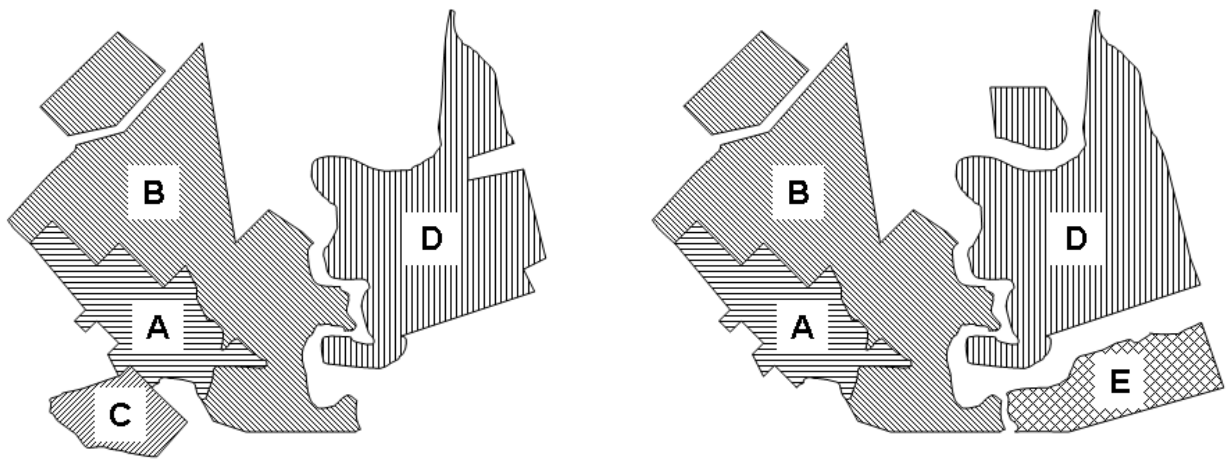
Table 1: NRN Updating Effectsshows the effects observed on objects after the NRN update.

Object	Explanation	Effect
3	No correspondence with a new object.	Retirement
4	No correspondence with an object in the original data; the arrival of object 6 changed the topological structure of the objects (and therefore the geometry).	Addition
5	No correspondence with an object in the original data; the arrival of object 6 changed the topological structure of the objects (and therefore the geometry).	Addition
6	No correspondence with an object in the original data; the feature was not represented.	Addition
e	No correspondence with an object in the original data.	Addition
f	No correspondence with an object in the original data.	Addition
2	Attribute value changed.	Descriptive modification
1	Geometry and attributes did not change.	Confirmation
a	Geometry and attributes did not change.	Confirmation
b	Geometry and attributes did not change.	Confirmation
c	Geometry and attributes did not change.	Confirmation
d	Geometry and attributes did not change.	Confirmation

Table 1: NRN Updating Effects

8.2 Aboriginal Lands (AL) Data (Section 3.4.2)

The following example describes change management for AL data. The AL data always uses the fourth method to manage all changes. The example illustrates the changes that occurred between the original AL dataset and the new AL dataset. The geometry of Indian Reserve A was not altered, but its name was changed. Indian Reserve B was confirmed to exist with the same geometry and description. Indian reserve C was abandoned. Lands were added to Indian Reserve D. Indian Reserve E was created. Table 2 explains the updating effects for each of the five AL objects.



Original data New data

Figure 3: Example of Changes After an Update of AL Data

The table 2 shows the effects observed on objects following updating of AL features in this dataset.

Object	Explanation	Effect
A	Attribute value changed.	Modification
B	Geometry and attributes did not change.	Confirmation
C	No correspondence with an object in the new dataset.	Retirement
D	Geometry was changed.	Modification
E	No correspondence with an object in the original dataset.	Addition

Table 2: AL Data Updating Effects

8.3 Municipal Boundaries (MUNI) Data (Section 3.4.3)

The examples in this section describe change management for MUNI data (municipalities, upper municipalities, and municipal regional areas). The MUNI data always uses the fourth method to manage all changes.

Given the MUNI data is divided into hierarchical levels, change management will be illustrated in two steps: 1) change management for municipalities (lower level of the hierarchy); 2) change management for corresponding upper municipalities (intermediate level of the hierarchy). Change management for municipal regional areas (higher level of the hierarchy) is performed in a similar manner than that for upper municipalities.

8.3.1 Municipalities (Section 3.4.3.1)

Figure 4 illustrates the changes that occurred between the original MUNI dataset and the new MUNI dataset for the municipality hierarchy level.

In Figure 4, objects A through E are municipalities. The geometry of Municipality A was not altered, but its name was changed. Municipality B was confirmed to exist with the same geometry and description. Municipality C was annexed by an adjacent municipality (not shown). Lands were added to Municipality D. Municipality E was created.

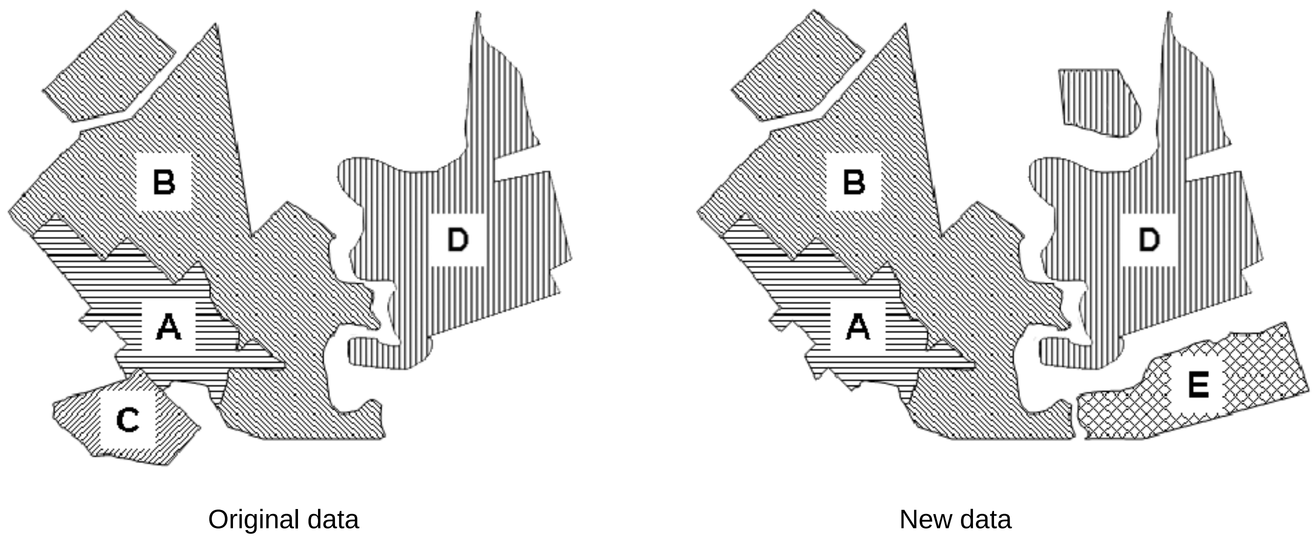


Figure 4: Example of Changes Following an Update of Municipal Boundaries (Municipality Hierarchy Level)

Table 3 shows the effects observed on objects (municipalities) following updating of MUNI in this dataset.

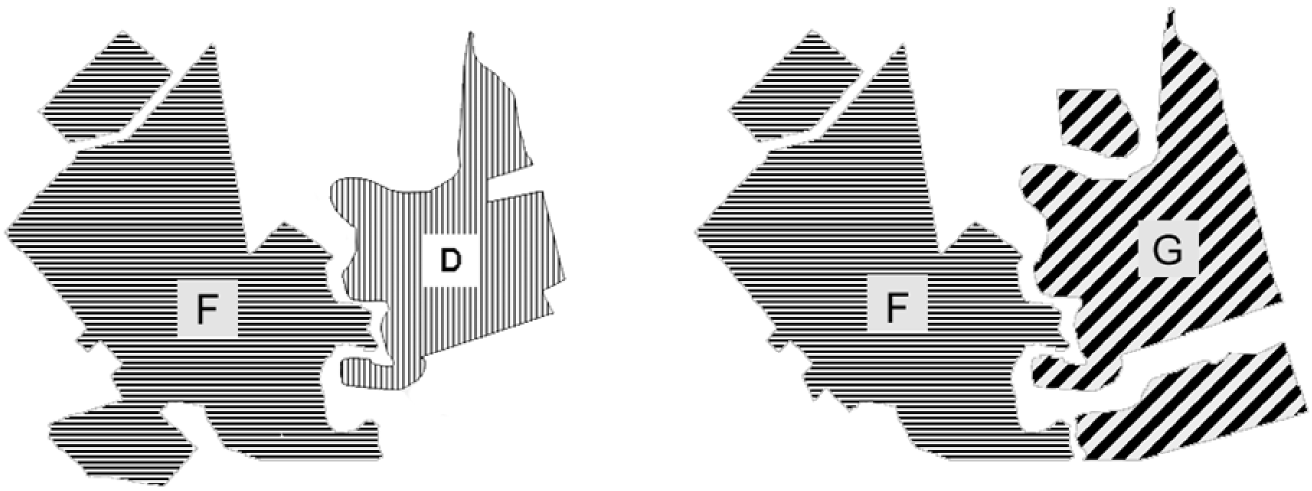
Object	Explanation	Effect
A	Attribute value changed.	Modification
B	Geometry and attributes did not change.	Confirmation
C	No correspondence with an object in the new dataset.	Retirement
D	Geometry was changed.	Modification
E	A new municipality is formed. It has no correspondence with an object in the original dataset.	Addition

Table 3: Updating Effects on Data for Municipalities

8.3.2 Upper Municipalities (Section 3.4.3.2)

Figure 5 illustrates the changes that occurred between the original MUNI dataset and the new MUNI dataset for the upper municipality hierarchy level.

In Figure 5, municipal feature F is the Upper Municipality that is comprised of Municipalities A, B, and C from Figure 4. In the original data, the Municipality D is not included in an upper municipality. After Municipality C is amalgamated to another upper municipality (not shown), the resulting Upper Municipality F has a smaller area than the previous one and contains only two Municipalities (A and B). Municipal feature G, an Upper Municipality, is defined for the first time. It consists of the redefined Municipality D and a new Municipality, municipal feature E.



Original data New data

Figure 5: Example of Change Management for Corresponding Upper Municipalities

The following table shows the effects observed on objects (upper municipalities) following updating of Municipal Boundaries in this dataset.

Object	Explanation	Effect
F	Municipality C is removed from the Upper Municipality F, and thus the geometry of Upper Municipality F changes.	Modification
G	A new upper municipality is defined, consisting of Municipality D and Municipality E.	Addition

Table 4: Updating Effects on Data for Upper Municipalities

REFERENCES

1 Bédard, Y., Larrivée, S. et Proulx, M.-J. "*Travaux de modélisation pour la mise en place de la base de données géospatiales ISIS*", Laval University, March 2000.

2 ISO Technical Committee 211, Working Group 4, "*Geographic Information – Encoding*," ISO/CD 19118, July 15, 2005.

3 Langran, Gail. *Time in Geographic Information Systems*, Éd.Taylor & Francis, 1993, 187 p.

4 PEUQUET, Donna J. *It´s About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems*, Annals of the Association of American Geographers, vol. 84, no 3, 1994, p. 441-461.

5 Worboys, Michael F. *A Unified Model for Spatial and Temporal Information*, The Computer Journal, Vol 37, No. 1, pp. 26-34

6 Pouliot, J., Larrivée, S., and Bédard, Y. *Typologie des mises à jour*, 2000, 11 p.