

Procédure d'assignation d'un Identificateur d'entité (IDE)

--- Version approuvée par les membres du COCG et de la CTC, les 12 mai 2009 et 6 août 2009 respectivement. ---

Introduction

La base de données toponymique du Canada (BDTC) décrite au site <<http://toponymes.rncan.gc.ca>> est un dépôt national de noms de lieux et d'entités géographiques officiellement adoptés ou approuvés par la Commission de toponymie du Canada (CTC). Dans la BDTC, une entité chevauchant une frontière provinciale ou territoriale possède plus d'un enregistrement dans la BDTC. Il s'agit d'un toponyme ayant deux enregistrements ou plus, un pour chaque province ou territoire dans lequel l'entité se situe, car ces toponymes sont approuvés par l'autorité provinciale ou territoriale, et les métadonnées (telles que les dates d'approbation) peuvent différer. Dans le même ordre d'idées, une entité ayant deux toponymes officiels, tel qu'un toponyme d'intérêt pancanadien, possède deux ou plus de deux enregistrements dans la BDTC, soit un dans chaque langue officielle, et pour chaque province ou territoire dans lequel elle se situe. De plus, chaque fois qu'une entité change de toponyme ou de code générique, de nouveaux enregistrements sont créés, un pour chaque changement. Ainsi la BDTC contient souvent plusieurs enregistrements pour la même entité. Par conséquent, dans la BDTC, une entité peut être associée à plusieurs toponymes (parfois répétés) et conséquemment, plus d'une clé de la BDTC lui sont associées.

Afin d'identifier de façon unique chaque entité, il fût proposé de créer un identificateur unique d'entité (IDE) et de l'insérer dans la BDTC. Cet identificateur est également utilisé par le Service de recherche de toponymes (voir <<http://srt.rncan.gc.ca/gnss-srt>>) du Service canadien de toponymie (voir <<http://sct.rncan.gc.ca>>). Le IDE demeure associé à son entité même si son toponyme change. L'IDE est unique pour toute entité ayant un toponyme au Canada. Quand le toponyme d'une entité est remplacé par un nouveau toponyme, l'ancien et le nouveau toponyme ont le même identificateur d'entité. Ainsi, il est possible de consulter l'historique des toponymes d'une entité.

Ce document définit l'identificateur d'entité et décrit la procédure d'assignation d'un identificateur d'entité.

Contexte cartographique

Dans le contexte de données cartographiques, une entité peut être représentée à différentes échelles de cartes ou à différentes projections cartographiques. Les différentes représentations peuvent ne pas être identiques (visuellement ou numériquement), bien qu'il s'agisse de la même entité. À la même échelle, deux cartes provenant de différentes sources ou différentes éditions d'une même carte, peuvent représenter une même entité différemment (si ce n'est pas visuellement, cela peut-être de façon numérique). Le but de l'IDE est de lier ces sources diverses ensemble. Par conséquent, l'IDE établit les délimitations d'une entité spécifique provenant de diverses sources de données cartographiques comme étant une seule et même entité.

En résumé, une entité cartographique pourrait être représentée différemment selon les sources, les échelles, les projections ou les éditions et chaque représentation aura, notamment, un Identificateur de géométrie tout en gardant le même IDE (voir Tableau 1). En d'autres mots, le but de l'IDE consiste à définir l'étendue et la limite géométrique d'une entité. Par conséquent, l'IDE aurait pu être nommé Identificateur d'étendue. De plus, lorsque plusieurs identificateurs de géométrie sont liés au même IDE, toutes les géométries représentent alors la même entité mais à des niveaux de détail (échelles) différents, avec différentes formes (projections) et différentes qualités (précisions).

Même si une entité est parfois désignée par plusieurs toponymes et par plusieurs géométries ils sont tous raccordés si et seulement si l'IDE est utilisé. Par exemple, au Tableau 1, une entité a plusieurs toponymes connexes (changements successifs de toponymes), par conséquent, elle est associée à plusieurs clés de la BDTC; cependant, l'IDE demeure le même pour chacun de ces toponymes.

Également dans le Tableau 1, lorsqu'une entité fut numérisée par plusieurs agences, avec différentes échelles, avec diverses projections et pour plusieurs éditions de cartes, chaque version a un Identificateur de géométrie, mais conserve le même l'IDE. Le Tableau 1 démontre qu'un IDE relie "plusieurs versions d'un toponyme" à "plusieurs versions de sa délimitation numérique".

Tableau 1. Exemple d'utilisation de l'identificateur d'entité (IDE).

Clé de la BDTC	Identificateur d'entité	Identificateur de géométrie	Identificateur d'entité
A (lac)	A	P (1/1000000)	A
B (nouveau nom)	A	Q (1/750000)	A
C (nom français)	A	R (1/50000)	A
D (nom anglais)	A	S (1/50000) v.2	A
Etc.	A	T (1/20000)	A
F (autre lac)	B	U (1/50000)	B

Le FID est implémenté en tant qu'identificateur universel unique (UUID)

L'IDE fut implémenté en tant qu'identificateur universel unique (UUID), aussi nommé identificateur unique global (GUID). L'IDE est constitué d'une séquence longue de 32 caractères, par exemple 6978bbe2ba3611d892e2080020a0f4c9. L'algorithme qui génère un UUID est disponible gratuitement sur le Web, et il est bien connu du milieu cartographique.

Suppositions d'opérations :

- Tous les groupes qui assignent ou gèrent les IDEs doivent être capables de générer des identificateurs qui sont universellement uniques (UUID).
- Les IDEs peuvent être attribués à toute entité, qu'elle soit nommée ou non.

- Il existe une relation de « un à plusieurs » entre un IDE et plusieurs clés de la BDTC.
- Les changements de noms n'affecteront pas l'attribution d'IDEs.
- Les délimitations géométriques à différentes échelles d'une même entité (par exemple, des cartes de l'Atlas et des cartes topographiques) partageront le même IDE.

Pour le projet du Réseau hydrographique national (RHN), les membres du Conseil canadien de la géomatique (COCG) ont adopté le UUID de 32 caractères pour identifier les entités hydrographiques. On suggère, ici, que le UUID soit aussi la norme pour identifier toutes les entités « nommées » adoptées par la CTC.

Norme internationale

L'IDE a été choisi de type UUID pour sa norme ISO qui est disponible sur la majorité des ordinateurs, simplement en tapant "uuidgen". S'il n'est pas disponible sur votre ordinateur, les gens de votre département d'informatique peuvent vous l'installer. Il est disponible gratuitement sur le Web. La plupart des organisations cartographiques utilisent le UUID dans leurs bases de données cartographiques numériques. Ainsi, en tapant "uuidgen", cela affichera, par exemple, 6978bbe2-ba36-11d8-92e2-080020a0f4c9 à l'écran. On peut alors copier et coller cette chaîne de caractères dans un champ de la base de données tout en s'assurant d'éliminer les traits d'union. Dans un environnement de production, ce UUID est automatiquement généré et inséré dans la base de données.

Durée de vie de l'IDE

Les règles suivantes proviennent ou sont adaptées du document « *Vector Data Identification Rules* », RNCAN, 1983.

- L'identificateur est assigné de façon permanente.
- Un UUID est assigné à une entité lorsqu'il est créé et ne changera pas pour toute la longévité d'une entité. L'IDE d'une entité supprimé ne peut être réassigné.
- Le UUID procure de la flexibilité pour des travaux impliquant plusieurs partenaires. L'algorithme bien connu du UUID peut être utilisé par tout usager qui désire modifier les données et pour ajouter un nouvel IDE. **Les identificateurs d'entité devraient être générés uniquement par des organisations autorisées.** Une attention particulière doit être portée aux IDEs. Ces IDEs vont éventuellement permettre la synchronisation de données entre organismes. Les utilisateurs de données doivent s'assurer qu'ils ne font **aucune modification aux valeurs des IDEs** : des modifications rendraient les IDEs inutiles lors d'une synchronisation de données.

Création initiale des IDEs

Une création initiale d'IDEs et leur association à des entités géographiques « nommées », a été effectuée et rendue disponible en décembre 2004, par la Section des applications toponymiques, Centre d'information topographique, Ressources naturelles Canada.

Au moment de sa création initiale, on a attribué le même IDE à des toponymes autochtones, français et anglais qui étaient raccordés à la même entité. Pour les entités qui chevauchent les frontières provinciales ou territoriales et qui ont un toponyme approuvé par chaque juridiction, nous leurs avons attribué des IDEs.

Les IDEs furent créés seulement pour les toponymes officiels de la BDTC et sont disponibles sur le Service canadien de toponymie (SCT). Aucun effort n'a été fait pour attribuer un IDE aux toponymes abrogés ou retirés.

Membres de la CTC - procédure d'attribution des IDEs

Lorsqu'un nouveau toponyme est approuvé et devient officiel, alors un IDE doit lui être attribué. Voici les étapes à suivre pour l'attribution d'un IDE :

1. Vérifier si l'entité est reliée à une entité nommée existante (toponyme connexe). Se référer au Tableau 2 pour voir les cas possibles.
2. Si elle ne correspond pas à une entité nommée existante, un nouvel IDE est généré. Passer les étapes 3 et 4.
3. Si elle correspond à une entité nommée existante, il faut attribuer au nouveau toponyme (i) l'IDE du toponyme connexe, ainsi que (ii) la clé de synonyme du toponyme connexe (voir la définition de la clé de synonyme à l'Appendice.)
4. Si elle correspond à une entité nommée existante qui n'a pas de clé de synonyme, il faut attribuer la même clé de synonyme aux deux enregistrements, la clé de synonyme étant la clé de la BDTC dont l'ordre alphabétique est le plus bas.

Notes :

- (i) Même si l'étendue est définie en se basant sur une autre échelle cartographique, ou une autre version de carte, le même IDE doit être attribué (si et lorsque le but est de représenter la même étendue au sol).
- (ii) Les IDEs sont habituellement attribués à une entité entière; par exemple, à tout le fleuve Saint-Laurent (et non pas partiellement comme le fait une clé de la BDTC. Par exemple, la portion du fleuve en Ontario et celle au Québec ont chacune leur clé de la BDTC.) Il existe quelques exceptions comme la Rivière Assiniboine au Manitoba (Voir le Tableau 2, 6^e type.)

Tableau 2. Exemple d'entités ayant plus d'une clé de la BDTC, mais assignés le même IDE.

<i>Type</i>	<i>Exemples</i>	<i>Notes / Commentaires</i>
1. Entité d'intérêt pancanadien	Lake Winnipeg / Lac Winnipeg (MB)	Les entités d'intérêt pancanadien doivent avoir deux toponymes, un dans chaque langue officielle. Une entité d'intérêt pancanadien peut ne pas avoir un deuxième toponyme dans l'une ou l'autre des deux langues officielles, par exemple Alberta : parce que le toponyme est identique dans les deux langues, une seule clé de la BDTC est retenue.
2. Entité ayant un toponyme dans les deux langues officielles du Canada – incluant les parcs nationaux, les établissements de la Défense nationale, etc.	Cassie Cape / Cap-des-Caissie (entité ayant deux toponymes officiels) (NB) Kouchibouguac National Park / Kouchibouguac, Parc national (NB)	À la différence des toponymes d'intérêt pancanadien, ces toponymes ont des statuts A7, et renvoient à la table des toponymes connexes (c'est la Clé de relation dans le SCT, voir l'Appendice.)
3. Entité ayant plusieurs noms, approuvés par des juridictions ou des compétences différentes.	Mont d'Iberville (QC) / Mount Caubvick (NL) Kitchen Brook (NB) / Cours d'eau Courchesne (QC)	Quelques entités chevauchant des provinces ou territoires contiguës possèdent plus d'un toponyme.
4. Entité chevauchant une frontière provinciale ou territoriale	Kicking Horse Pass (AB/CB) Flin Flon (SK/MB)	Plusieurs enregistrements de la BDTC existent pour ces entités; un pour chaque province.
5. Entité ayant des toponymes antérieurs, anciens ou reliés.	Bytown – ancien nom pour Ottawa (ON) Little Oven – ancien nom pour Le Petit Four (QC)	Le Québec gère différemment la création et la mise à jour des enregistrements de ses toponymes. Cet enjeu devra être pris en considération lors des futurs téléchargements au moment où l'attribution des IDES sera implémentée. (noms de renvois pour les statuts B4 et B12 seulement)
6. Entité ayant un toponyme de type Équivalent manitobain, Parallèle ontarien, ou d'autres qui peuvent s'ajouter à ce genre dans le futur.	Assiniboine River / Rivière Assiniboine (MB)	Parfois, seulement une ou des parties d'une entité est ou sont bilingue(s). Par exemple, Rivière Assiniboine est aussi nommée Assiniboine River, mais <u>seulement</u> là où elle coule dans une région désignée en langue française. La rivière Assiniboine a plusieurs IDEs d'attribués.

Tableau 2 (suite). Exemple d'entités ayant plus d'une clé de la BDTC, mais assignés le même IDE.

<i>Type</i>	<i>Exemples</i>	<i>Notes / Commentaires</i>
7. Entité ayant plusieurs toponymes	PrinceAlbert Sound / Kangiryuaq (NT)	Quelques entités ont plusieurs toponymes choisis parmi les multiples langues autochtones officielles du territoire ou de la province.
8. Entité faisant partie de deux des catégories précédentes. Par exemple, une entité d'intérêt pancanadien qui chevauche une frontière.	Wood Buffalo National Park / Parc national Wood Buffalo (AB/NT) Gulf of St. Lawrence / Golfe du Saint-Laurent (QC, NB, NS, PE, NL)	Un IDE est attribué pour l'entité entière.

Membres du COCG - procédure d'attribution des IDEs

La procédure pour décider de l'attribution à une entité géographique d'un IDE existant ou bien d'un nouvel IDE, est la même que pour les membres de la CTC, sauf pour l'obligation d'introduire la clé de synonyme.

Quand une entité est déjà nommée: les membres du COCG doivent obtenir les IDEs existants soit de leurs membres de la CTC, du SCT, ou de GeoBase. Ceci est essentiel car toutes les entités nommées au Canada ont leur IDE national déjà publié et ces données sont largement utilisées. GéoBase publie des IDEs pour les entités nommées dans la couche du RHN ainsi que pour la couche des Toponymes, et l'Atlas du Canada a déjà assigné des IDEs à leurs délimitations. Les membres de la CTC ont du publier des IDEs nationaux pour les entités nommées par le biais du site Web du SCT afin de fournir des IDEs aux projets du RHN et de l'Atlas.

Lien avec la délimitation numérique

Les toponymes sont liés avec les délimitations numériques entreposées dans une base de données géospatiales (en se servant de leur IDE). Cependant, comme à l'habitude, les délimitations pour chaque nouvelle décision seront fournies sur des cartes version papier, des fichiers numériques ou des fichiers pour SIG.

L'objectif des membres de la CTC et du Service canadien de toponymie (voir <<http://sct.rncan.gc.ca>>) est de rendre disponibles les délimitations numériques via un service Web normalisé de cartes (SNC) ou un service Web normalisé d'entités (SNE) conforme au « Open Geospatial Consortium » (OGC), voir <<http://www.opengeospatial.org>>).

Élimination d'un identificateur

Comme à l'habitude, un enregistrement et un IDE ne seront pas supprimés même si une entité n'existe plus ou si un toponyme n'est plus officiel; sa géométrie devra exister dans une banque de données géospatiales (pour des fins historiques ou légales.) Lorsque le toponyme n'est plus officiel, l'enregistrement demeure dans la BDTC et dans le SCT mais a le statut « abrogé ». Voir le paragraphe « Durée de vie de l'IDE » ci-haut.

Recherche de toponymes antérieurs ou anciens

L'IDE permet de retrouver les toponymes anciens d'une même étendue au sol.

Étendue ayant des toponymes dans de multiple langues officielles

Les noms d'intérêt pancanadien, les Équivalents manitobains, les noms parallèles ontariens, les parcs nationaux, etc. ont des toponymes français et anglais pour la même étendue. Ainsi, l'IDE fait le lien entre ces toponymes officiels. En plus des toponymes français et anglais, les territoires et quelques provinces reconnaissent plusieurs toponymes autochtones pour quelques entités.

Bibliographie

RNCan, 2003. « *National Vector Data Identification Rules* », Ressources naturelles Canada, Géomatique Canada, Centre de l'information topographique, Sherbrooke, Québec, Canada, 8 pages.

Acronymes

- BDTC - Base de données toponymiques du Canada
(voir <http://geonames.nrcan.gc.ca/search/search_f.php>)
- COCG - Conseil canadien de la géomatique (voir <<http://www.geobase.ca>>)
- CTC - Commission de toponymie du Canada
(voir <http://geonames.nrcan.gc.ca/index_f.php>)
- SCT - Service canadien de toponymie (voir <http://gnss.nrcan.gc.ca/index_f.html>)
- SRT - Service de recherche de toponymie du SCT
(voir <<http://gnss.nrcan.gc.ca/gnss-srt/searchName.jsp?language=fr>>)

Annexe

Champs de la base de données qui sont liés

En plus de l'IDE, cinq nouveaux champs furent ajoutés à la base de données du SCT.

1. **Clé de synonyme** (« Synonym Key ») aide à gérer les identificateurs d'entités. Permet de trouver tous les toponymes ayant la même étendue. Lorsqu'une étendue a plus d'un toponyme, chaque toponyme a le même IDE et la même clé de synonyme. Lorsqu'un enregistrement n'a pas de clé de synonyme, cela signifie qu'il n'y a aucun autre toponyme ayant la même étendue. Afin de simplifier sa maintenance, on assigne comme clé de synonyme la valeur de la clé de la BDTC la plus basse alphabétiquement parmi celles de chaque toponyme ayant la même étendue.
2. **Clé de relation** (« Related ») permet de trouver (i) tous les noms antérieurs (historiques, abrogés), (ii) les noms bilingues (officiels), (iii) les variantes du toponyme et (iv) les noms retirés (pour diverses raisons, incluant les fautes de frappe), même si les étendues diffèrent. La clé de relation remplit le rôle des noms en renvoi.
3. **Historique** (« History ») permet de retrouver chacun des cas (i), (ii), (iii) ou (iv) ci-haut. Les seuls mots clé acceptés dans le champ "History" sont: 1-Alternate Name, 2-Official Name, 3-Former Name, 4-Variant Name. Ces chiffres furent choisis afin d'ordonner les toponymes, de 'important' à 'moins important'.
4. **Liste** (« List ») permet d'extraire des listes spécifiques qui sont souvent utilisées par les usagers et les membres de la CTC. De telles listes incluent les toponymes d'intérêt d'intérêt pancanadien, les toponymes équivalents manitobains, les toponymes parallèles ontariens, les parcs nationaux, les toponymes de la Défense nationale, etc. D'autres peuvent être créées. Les mots clé possibles sont: PanCan, equiMB, altON, IR, Parks, DND, etc.
5. **Langue** (« Language ») lorsqu'elle est utilisée avec les champs « clé de relation », « clé de synonyme », « IDE », « historique », « liste », le champ « langue » permet de retrouver ou d'identifier les toponymes anglais, français ou autochtones. Ce champ permet aussi, dans une application, d'afficher automatiquement le toponyme anglais, français ou autochtone, selon le contexte. Les codes acceptés pour le champ « langue » sont les codes internationaux de trois lettres de l'ISO 639-3 identifiant les langages tel que: eng (anglais), fre (français), crm (Cri, Moose) crk (Cri des plaines), etc. Si le champ est nulle (vide), cela signifie que l'entité n'a pas d'autre toponyme officiel ou pas d'équivalent officiel dans une autre langue. Si le champ n'est pas nulle, alors l'entité possède au moins un autre toponyme officiel dans une autre langue.

Note :

Les provinces ou les territoires qui contribuent au SCT, mais maintiennent leurs propre systèmes électroniques de toponymes, ne sont pas obligés de charger ou de maintenir les nouveaux champs énumérés ci-dessus.