

Item 9 of the Provisional Agenda

Activities Relating to the Working Group on Toponymic Data Files and Gazetteers

**Advances in Spatial Data Management in
Canada's National Geographical Names Database**

Submitted by Canada*

* Prepared by Kristina Kwiatkowski, Toponymy Specialist, Geographical Names Board of Canada Secretariat, Natural Resources Canada (Canada)

Progrès réalisés dans la gestion des données spatiales de la

Base de données toponymiques du Canada

RÉSUMÉ

Au cours des deux dernières années, la Base de données toponymiques du Canada (BDTC), base de données nationale de référence pour les toponymes canadiens, a fait l'objet d'améliorations structurelles et techniques importantes. Le présent document porte sur la migration du modèle de données de la BDTC qui, d'un modèle axé sur les attributs est devenu un modèle axé sur les données géospatiales, la gestion des liens toponymiques au sein de la BDTC, le développement d'une nouvelle application Web pour accéder à la base de données et la diffusion d'une série de produits de données ouvertes de la BDTC.

Contexte

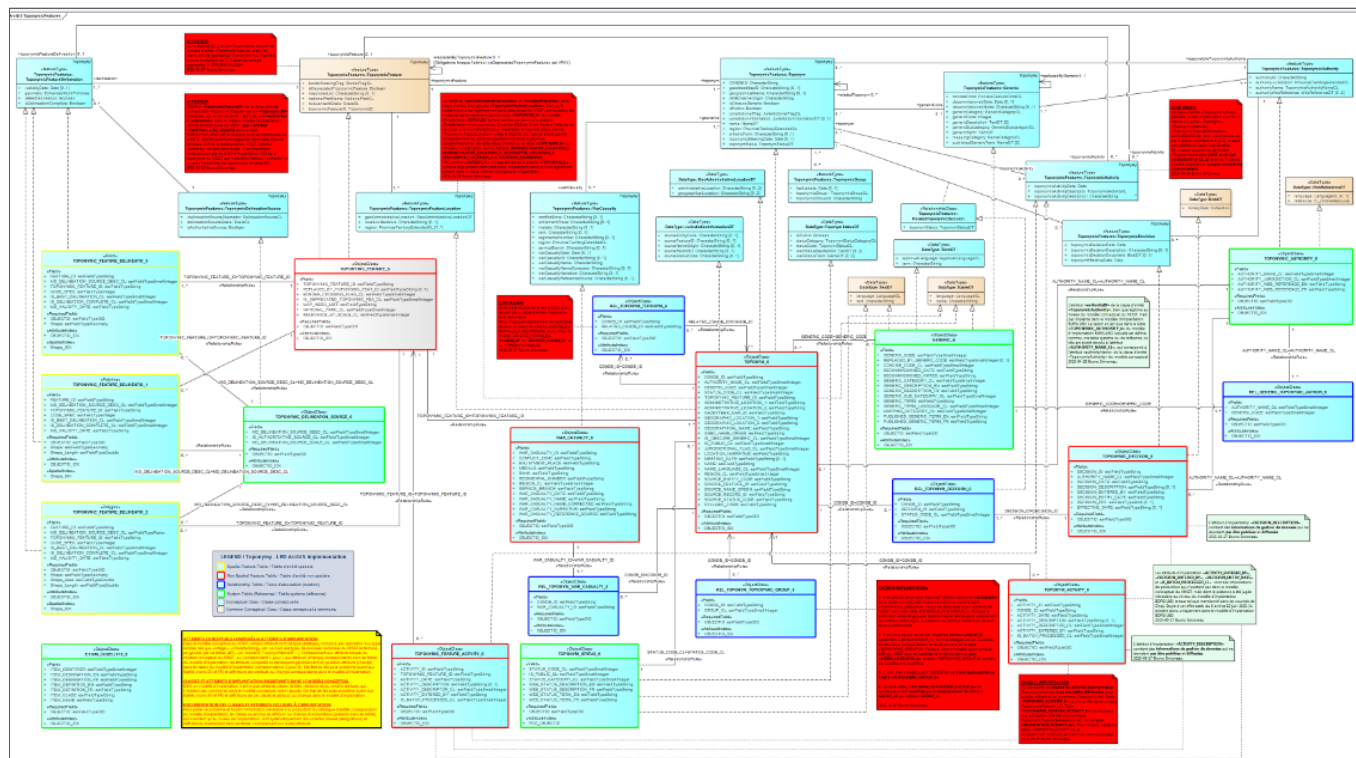
La Commission de toponymie du Canada (CTC) est l'organisme de coordination national chargé des normes et des politiques en matière de toponymes canadiens. La CTC se compose de ministères et d'organismes fédéraux, provinciaux et territoriaux, ayant chacun des pouvoirs et des responsabilités particuliers au sein de leurs territoires de compétence. Collaborant ensemble à titre d'organisme intergouvernemental national, les membres de la Commission de toponymie du Canada coordonnent leurs efforts pour assurer la gestion uniforme des toponymes.

La CTC est appuyée par un secrétariat assuré par Ressources naturelles Canada (RNCAN), ministère fédéral, qui fournit l'infrastructure et le soutien à la BDTC à titre de base de données nationale de référence pour les toponymes canadiens et élément clé de l'infrastructure des données géospatiales du Canada. RNCAN regroupe les données toponymiques, les délimitations spatiales des entités et les décisions relatives aux nouvelles dénominations communiquées par les autorités de dénomination de la Commission de toponymie du Canada (CTC) dans la base de données nationale. En février 2016, plus de 550 000 toponymes, dont 390 000 noms officiels, étaient gérés dans la BDTC.

Modèle de données

En 2015, le modèle de données de la BDTC a fait l'objet d'une transformation fondamentale, ayant évolué d'un modèle axé sur les attributs vers un modèle de données géospatiales interexploitable et plus efficient. Reposant sur un modèle relationnel basé sur des normes ISO, le nouveau modèle de données améliore considérablement la fonctionnalité et l'interopérabilité de la base de données nationale. L'ancien modèle de données contenait un tableau unique comportant plus de 90 champs. Le nouveau modèle de données géospatiales contient 45 tableaux et peut traiter des données spatiales et relationnelles. La transformation du modèle de données permet des liens entre les toponymes et les délimitations spatiales des entités. La BDTC contient plus de 133 000 délimitations spatiales, principalement pour les phénomènes hydrographiques, tels que les lacs et les rivières. Le nouveau modèle peut gérer des documents de divers formats, tels que les documents PDF de décisions, les délimitations de fichiers Shapefile et les fichiers sonores. Il permet également d'appuyer plus efficacement la validation des données, la surveillance des bases de données et les rapports statistiques.

FIGURE 1 : MODÈLE DE DONNÉES DE LA BDTC DE 2015



Identificateurs d'entité de la BDTC

L'élément essentiel de la transformation de la BDTC qui, d'une base axée sur les attributs est devenue une base de données géospatiales, a été l'inclusion d'identificateurs d'entité (IDE). Les IDE ont été conçus pour identifier de façon unique chaque entité nommée; mis en œuvre à titre d'identificateur unique universel, ils comportent 32 caractères alphanumériques. Les IDE demeurent associés à une entité, quels que soient les changements apportés aux toponymes de l'entité. Ils permettent le retrait de tous les toponymes ayant la même étendue spatiale. Cependant, en cas de modification importante de l'étendue spatiale de l'entité, un nouvel IDE sera attribué. Par exemple, si l'eau d'un lac s'évaporerait et que le lac se transformerait en deux lacs distincts plus petits, deux nouveaux IDE seraient attribués à chacun des deux nouveaux lacs.

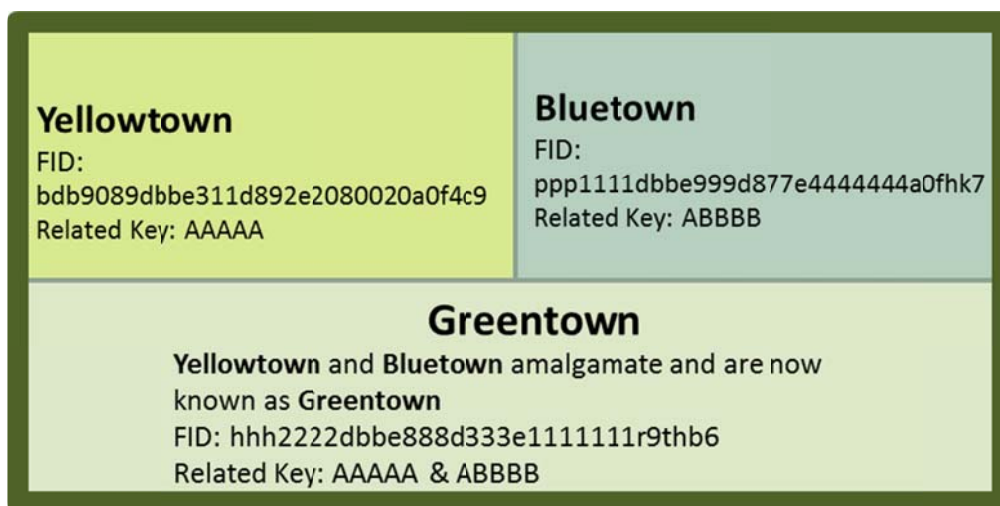
En 2015-2016, RNCAN a poursuivi ses travaux en ajoutant des IDE à tous les toponymes figurant dans la BDTC. Ce processus d'ajout d'IDE était semi-automatisé. Des exigences d'attribution automatique ont été créées pour attribuer des IDE aux liens de parenté non complexes. Par exemple, on aurait automatiquement attribué le même IDE à un nom actuellement officiel et un nom auparavant officiel apparentés ayant la même définition générale.

Pour traiter les dossiers qui ne répondaient pas aux exigences d'attribution automatique, un groupe de travail spécial composé de membres de la CTC a examiné divers scénarios et créé des règles d'attribution d'IDE. Un examen approfondi des liens historiques et de l'étendue spatiale des dossiers a été réalisé par des techniciens en géomatique de RNCAN en collaboration avec les autorités de dénomination provinciales et territoriales pour s'assurer qu'un identificateur d'entité pertinent avait été attribué.

Conservation des liens historiques

Avant l'utilisation d'identificateurs d'entité, les liens entre les toponymes étaient traités à l'aide d'une clé connexe à cinq caractères. La clé connexe suit « l'historique du nom », c'est-à-dire qu'elle suit les divers changements apportés au nom au fil du temps, mais pas nécessairement l'étendue spatiale exacte du nom. Par exemple, lorsque deux municipalités fusionnent et que la nouvelle municipalité prend le nom de l'une des deux anciennes municipalités, on aura trois IDE distincts et la clé connexe reliera historiquement les trois noms.

FIGURE 2 : SCENARIO D'ATTRIBUTION DE L'IDE DE VILLES FUSIONNEES



Application Web de toponymes

Afin d'appuyer l'élaboration d'un nouveau modèle de données pour la BDTC et de faciliter les recherches, les révisions et les mises à jour effectuées par des membres de la CTC, RNCAN a également mis à jour l'outil d'application Web appelé GNApp-II. La nouvelle application lancée en février 2016 est destinée à améliorer l'interaction directe avec la base de données nationale par les membres de la CTC. La nouvelle application a été élaborée à l'aide des intrants des membres par le recensement des exigences et des essais de convivialité. Outre l'apport d'un visualisateur de cartes, l'application a permis d'améliorer les fonctions d'affichage et de recherche. Grâce à l'application GNApp-II, les autorités de dénomination de la CTC peuvent maintenant présenter des documents sur les décisions de dénomination à la base de données et télécharger des délimitations d'entités nommées dans la BDTC en format Shapefile.

Diffusion des données de la BDTC

Le public peut avoir accès à la base de données toponymiques nationale du Canada grâce à un outil de recherche sur le site Web de RNCAN et télécharger les données en divers formats.

Les recherches de noms de lieux en ligne par l'entremise du site Web de RNCAN peuvent se faire à partir du toponyme, du type d'entité, de la province ou du territoire, des coordonnées (latitude/longitude), de la région rectangulaire ou en utilisant une clé unique. Les utilisateurs peuvent également effectuer des recherches par nom contenant des caractères de langues autochtones (écriture syllabique inuktitut ou alphabet latin étendu). Une recherche renvoie le type d'entité, la région, les identificateurs uniques du

nom et de l'entité, la latitude et la longitude, la date d'approbation ou de changement de statut du toponyme et la délimitation spatiale du lieu du toponyme (si ces données sont disponibles) figurant sur une carte de la base de données. Les noms officiels et historiques d'entités géographiques canadiennes, y compris les lieux habités et les entités sous-marines, se trouvent sur le site Web de RNCan à : <http://www.rncan.gc.ca/sciences-terre/geographie/noms-lieux/recherche/9171>

La série de produits toponymiques numérisés provenant de la BDTC a été récemment réévaluée, reformatée, mise à jour et normalisée afin de répondre aux besoins actuels en matière de données. Les dossiers de produits toponymiques portant sur les provinces et les territoires canadiens peuvent être téléchargés dans les formats Shapefile, KML, GML et CSV du portail GéoGratis (<http://geogratias.gc.ca/>) ainsi que du portail de données ouvertes du gouvernement du Canada (<http://ouvert.canada.ca/fr/donnees-ouvertes>).

En outre, RNCan offre une interface de programmation d'applications (API) comme moyen d'accès direct à la BDTC pour les recherches personnalisées.

Prochaines étapes

En 2016-2017, RNCan créera une fonctionnalité supplémentaire dans l'application GNApp-II, permettant notamment aux membres de la CTC d'ajouter des fichiers sonores contenant la prononciation des noms, et ajoutera des fonctions de téléchargement par lot par l'entremise de l'interface de l'application Web.

RNCan travaillera en collaboration avec les autorités de dénomination de la CTC pour réaliser un examen approfondi des diverses listes de codes au sein de la BDTC, qui définissent le type d'entité et le statut du toponyme. Cet examen vise à déterminer si d'autres économies et progrès peuvent être réalisés en renforçant les règles de validation pour la saisie des données dans l'application Web GNApp-II.

RNCan continuera de collaborer avec les autorités de dénomination de la Commission de toponymie du Canada pour définir les formats de diffusion optimaux pour les données toponymiques et les moyens de surmonter certains défis particuliers liés à la diffusion de toponymes, par exemple inclure plusieurs noms d'entités officiels et des noms en plusieurs langues autochtones. L'objectif est de fournir des toponymes uniformes, faisant autorité et à jour de la BDTC dans le cadre de la série de données ouvertes géospatiales du Canada.