

GT Bati CNIG 8 10-03-2023

Date: 10.03.2023

- Lien vers le support Mural pdf
- Lien vers l'enregistrement vidéo
- Rappel - Lien vers la définition & annexe

Participants:

- Pierre PAGES (Toulouse Métropole)
- Stéphanie Alibert (Toulouse Métropole)
- Sébastien Wehrlé (Eurométropole de Strasbourg)
- Régis Haubourg CSTB
- François ROLLO (DGFIP) Bureau du cadastre
- Sébastien LEGER (DGFIP) Bureau du cadastre.
- Loïc REBOURS (Enedis)
- Benoît Génot (U.R.B.S)
- Mylène Ritter (SDIS 38)
- Olivier Lefebvre (INSEE)
- Julieta Barado (DGALN)
- Filipe Afonso (Stonal)
- Thibaud Le DOrze GéoVendée
- Benoit David (MTECT/Ecolab)

Actus projet Bat-ID

- Volet réglementaire :
 - Étude construction base légale
 - High Value Data Set
- Communes partenaires :
 - Avancées Grenoble & Toulouse
- Futures expé : Rennes, ERP, Mission Rénovation
- Interopérabilité des bases : BAN, BAL, BD Topo
- CR & travail en asynchrone

Update Volet Reglementaire

- Géométrie: on se réfère aux différentes géométries possibles dans Inspire - pas bloquant d'avoir plusieurs types de géométries
- Nombre d'étages: Il faut prévoir un attribut mais ok si pas encore parfait, complet
- Usage : Ok si on a indirectement l'information par le biais des locaux plutot que des bâtiments - dès lors que le lien est fait

- Possible de remonter des commentaires au groupe de maintenance technique d'Inspire
- Besoin de réaliser un tableau de correspondance du modèle national avec modèle INSPIRE
- Valeurs indiquant l'élément pris en considération pour saisir une géométrie horizontale.
 - Bord du toit
 - Combinaison
 - Emprise
 - Enveloppe
 - Enveloppe au dessus du sol
 - Point d'entrée
 - Point à l'intérieur de la parcelle cadastrale
 - Point à l'intérieur du bâtiment
 - Étage le plus bas au dessus du sol

Source ## Format ID Unique

L'issue Github dédiée résume les enjeux :

NanoID qui permet de la génération décentralisée avec risque de collision minimal.

- Minuscule, Majuscule : peut importe tant que l'ID n'est pas sensible à la casse
- L'utilisation des tirets (du bas ou du haut)

Question : Pourquoi être passé en majuscule ? De devoir mettre un caps locks c'est pénible. Donc potentiellement les UUID peuvent être bloqué si on prend en compte le cas d'usage le post-it sur lequel on rédige à la main. Le plus important c'est de ne pas avoir des minuscules ET des majuscules. Un autre argument pour les majuscules pourrait être que c'est plus lisible dans les formulaires.

Sur la signifiante : - Facilité de lecture et de saisie de l'ID : argumentaire pour construire l'identifiant de manière à ce qu'il soit rattaché à des clés de type "code commune" "code département" partant du principe que la chaîne de numéro donne bien un ID non signifiant lorsqu'il est lu dans sa totalité. Ce n'est que quand on le découpe qu'on trouve un sens à la numérotation. - Autre argument : ce type d'identifiant est utilisé dans un grand nombre de référentiel (n_carte d'identité, permis de construire, sécurité sociale, données relatives à l'environnement) - ce sont des choses qui existent déjà et son utilisé. Et elles permettent de favoriser la lecture, la réutilisation et la re-saisie par des utilisateurs moins habitués aux ID non compréhensibles. Mettre en avant la lisibilité au profit des utilisateurs plutôt qu'un ID complètement aléatoire

- A la rigueur peut importe si le code commune change (fusion, scission de commune), ce qui est intéressant c'est le fait de dire qu'à la création on a un code INSEE par exemple. Une fois que l'ID a été saisi à une certaine

date, il doit être considéré comme définitif, peut importe les évolutions. C'est vraiment l'ensemble de la chaîne qu'on prend en compte dans sa "non signifiante".

- Est ce qu'il y a d'autres personnes qui partagent cet usage ?
- On a une certaine "habitude" avec ces identifiants (sur la partie n° de commune, département notamment) mais aller sur des éléments additionnels comme des n° d'année par exemple cela pourrait être plus difficile.
- Exemple du référentiel des codes cours d'eau : une fois que le code est donné on est sensé ne pas le changer. MAIS quand il y a un changement de code commune, et que l'ID du bâtiment ne serait plus raccord et bien certaines personnes risquent de faire des raccourcis. Et cela peut entraîner des problèmes importants dans les chaînes de traitement. Donc pour le gain en utilisabilité -> on a potentiellement beaucoup de problème dans les SI derrière.
- Mettre le focus sur l'utilisabilité de l'ID : pas trop long, pas trop complexe. On ne pouvait pas faire avant ce type d'UUID ou Nano-ID, donc on faisait une codification qui représentait la hiérarchie administrative pour éviter les recouvrements. Avec les nouvelles technologies, on peut en profiter.
- Vrai risque que l'ID soit détourné pour commencer à faire des filtres et ça peut entraîner des soucis.
 - A Toulouse ils l'ont identifié, que ça invite les utilisateurs à se servir des codes pour les requêtes. C'est une limite, mais il y a peut être moyen de préciser que ça reste un ID et que ça ne peut pas être utilisé comme clé de requête -> peut être de la pédagogie à faire ?
- Risque vs. gain ? Le gain est la facilité de l'expression AU MOMENT où il est généré uniquement. Sauf que si le code commune a changé cela risque en réalité de complexifier et créer de l'ambiguïté.
- Chez ENEDIS : pleins de problème avec de la signifiante. On se rend compte que dès qu'on parle des codes INSEE c'est une manière de rechercher ou de garantir qu'on identifie un bâtiment au bon endroit. Dans notre MDD on a l'adresse qui est associée, DONC on a déjà la relation avec certaines bases de données qui sont géographiques. Donc pas besoin d'en rajouter dans l'ID Bâtiment.
- D'accord également sur le "non signifiant". Et la CNIL peut ne pas apprécier le fait que l'identifiant soit signifiant pour ne pas relier à des données personnelles (il faut être attentif à ça). Il faut partir sur des ID non signifiant pour toutes les raisons de gestion qui ont été apporté. Quand on a besoin de faire le lien avec des éléments signifiant, c'est justement à ça que sert l'interopérabilité. Et les codes INSEE changent souvent. Et si on est encore par un n° de sécurité social signifiant ce n'est pas parce qu'on considère que c'est ce qui est le mieux, c'est uniquement

parce qu'aujourd'hui on ne peut plus changer. Ce serait trop compliqué. Donc oui, le n° de sécurité social est signifiant parce qu'on avait pas le choix à l'époque mais ajd s'il devait être créé, il serait non signifiant (pdv INSEE).

- Un ID signifiant ne doit-il pas l'être tout au long de la vie du bâti ? Quel intérêt s'il n'est signifiant qu'au moment où il est généré ?
- “On a toujours fait comme ça” -> le poids des habitudes parce qu'avant on n'avait pas le choix que d'avoir un ID signifiant pour différencier. Alors qu'aujourd'hui on a les Nano-ID.
- Exactement. Si on part sur du signifiant, il faut le rester, et on perd l'unicité dans le temps de l'id. Ce qui est une règle dure en système d'informations. (Inspire nous le rappelle d'ailleurs)
- Pour l'identifiant d'une personne : On peut considérer que c'est la commune à sa naissance je pense. Pour un identifiant sur un bâtiment, le risque est que les acteurs vont développer des règles sur la localisation
- D'expérience, je trouve que la crispation sur les identifiants se lève (un peu) lorsqu'on met en valeur qu'il y a 2 autres moyens d'identifications :
 - la position géographique
 - une désignation/“appellation” (adresse + entrée)
- Ce sont ces 3 infos qui peuvent servir dans les cas d'usage sur 'l'identification' => il faudra bien travailler sur ces 2 autres aspects aussi

Question de “A quel moment est généré l'ID ?” - Est ce qu'il pourrait y avoir deux temps, au moment de création de la donnée ? C'est incrémental donc presque pas d'intérêt à avoir quelque chose de signifiant.

Question sur le nombre de caractères ? - Est ce que le fait d'avoir 12 caractères (en série de 3x4) - Ou est ce que 15 caractères (qui ont une probabilité très faible de collision)

- Triplé sont plus facile à retenir -> mais 12 caractères est un peu juste en terme de probabilité.
- Il faut pouvoir découper le code en groupe pour pouvoir l'énoncer - même si le traitement restera largement informatique.
- 16 caractères c'est un numéro de carte bleue

Sur la probabilité de collision : - Générateur d'ID : - <https://zelark.github.io/nano-id-cc/> - Le plus on évite le problème le mieux c'est :

Alphabet:
 123456789ABCDEFHJKLMNPQRSTVWXYZ 32/256

ID length: characters

Speed: IDs per hour/second

~347 years needed, in order to have a 1% probability of at least one collision.

```
const { customAlphabet } = require('nanoid');
const alphabet = '123456789ABCDEFHJKLMNPQRSTVWXYZ';
const nanoid = customAlphabet(alphabet, 12);
nanoid() //=> "66SPBG5NR92H"
```

- A priori baisser de 12 caractères ça semble gérable (risque faible de collision, longueur acceptable d'usage par un humain et permet la génération du stock)

Quel impact de générer le stock ? - Ca représente 20 ans, donc ça reste large. Donc il ne faut pas trop baisser non plus.

Déjà eu ce type de débat dans les BAL : - A l'époque on voulait du non-signifiant - Et le système BAL actuel devait être uniquement temporaire - la clé d'interopérabilité BAN n'est pas stable en effet. On s'est fait mordre aussi

Comment décider ? - On ouvre un ticket large avec un vote - On prend le temps de tester et ce n'est pas arrêté -> il faudra tester avec des utilisateurs pour voir comment ils s'en saisissent.

Question à regarder en intercession : est ce qu'on a besoin d'avoir une clé de contrôle (les 2 chiffres du n° de sécu par ex). -> A creuser. - A regarder aussi au regard de la criticité du référentiel - Est ce que les contrôles asynchrone de fiabilité seraient suffisants ?

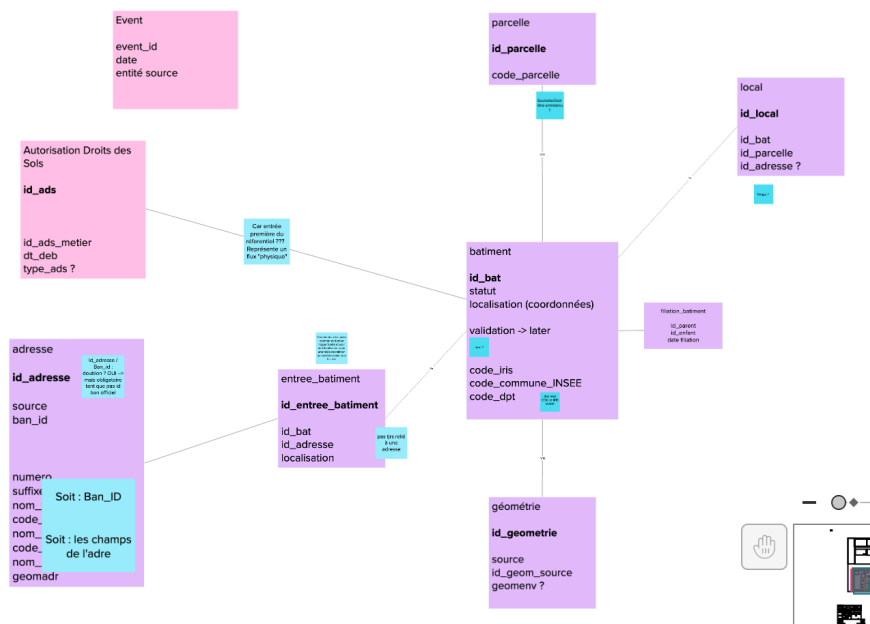
Statuts bâtiment

- Périmé dans le RIL : Quand le permis est annulé OU quand il y a une autorisation de démolir
- Pour le RNB:
 - Ne pas rajouter trop d'étape
 - Déjà savoir quand un chantier commence, quand il se termine
- Quelque soit les valeurs qu'on prend, est ce qu'on regarde celles d'INSPIRE :

- Il y a des listes de valeurs traduites (codes listes en anglais avec les valeurs de traduction en français)
- Il faut avoir en tête qu'on doit être capable de faire la traduction
- Sur “en construction” et “construit”
 - Dans le flux, on a les ADS → l'idée est de créer l'ID le plus tôt possible. Donc il faut au moins un statut initial (avant la construction). Ce qui semble très important c'est de pouvoir passer d'un statut initial à l'état de “construit”
 - “En projet” est ce que ce serait dès la demande du permis ? Ou s'il est autorisé ?
 - * A voir si c'est important d'avoir la granularité “En projet” ou “Instruit” ?
 - “En construction” → DOC
 - “Construit” → DAACT → mais aussi probable que ça puisse remonter par d'autres biais (DGFIP, INSEE)
 - Mais oui globalement, on souhaite coller aux processus d'ADS mais il faut que ça soit plus large par ailleurs pour réussir à avoir cette information par d'autres biais que les ADS (ex. DGFIP, INSEE, Autre)
 - Un bâtiment en rénovation sera considéré comme “en construction” ?
- Attention : Quelle capacité des différents acteurs de l'alimentation à renseigner les différents postes qu'on aura ?
- Autre manière de se mettre d'accord sur le statut :
 - Est ce qu'on peut regarder les dates d'INSPIRE : construction, démolition, rénovation
- En fait si on gère un historique, on devrait comprendre que si c'est un “2e” construction → c'est une rénovation puisqu'il y a que la première construction qui vaut pour faire sortir le bâtiment de terre.
 - Mais si on met “date de rénovation” on risque d'avoir très peu de dates qui remontent dans INSPIRE.
 - Rénovation = sujet sur lequel on a pas vraiment la source d'information. On aura du mal à qualifier ce qui rentre dans la notion de rénovation (beaucoup de rénovation ne passent pas par le permis ou les aides d'état) → donc c'est un flux qui est très compliqué à attraper.
 - Il y a très peu de flux d'information sur la rénovation qui seraient exhaustif et propre aujourd'hui.

- OSM : Modèle intéressant pour qualifier la généalogie, historique de chaque attribut -> Il est essentiel d'avoir la source de l'information. Donc ce serait bien de pouvoir qualifier d'où vient l'information et de manière.
 - oui, on va toujours préciser le statut "selon xx acteur"
- Il faut bien distinguer "Annulé" (le permis n'a pas abouti) et "Erreur en base / Correction inventaire"
- Quelle différence entre "en projet" et "en construction" -> a priori le flux de demande de permis vs. DOC
- Quid du statut "En ruine" -> a voir comment on arrive à le collecter. Point d'action avec OSM pour comprendre comment ils pourraient remonter ce flux d'information.

Modèle de donnée



- Violet : Element socle
- Rose : Éléments types "événements"
- Bleu : remarques en vrac

Questions : - Référentiels externes : On s'attend à ce que eux référence l'ID Bâtiment. Mais pas demandé à la saisie dans le flux de création de bâtiment (sauf peut être locaux & parcelles ?). - Sur les adresses : est ce que l'adresse BAN est obligatoire ou pas ? Sur un immeuble qui a 11 adresses - qu'est ce qu'on fait ? Est ce qu'un bâtiment a 3 adresses principale? - Un enjeu stratégique : il faut qu'on ai un socle très solide adresse-bati-local -> la cible est l'unicité

de ces 3 référentiels. La feuille de route pour y arriver, c'est toute la technicité du sujet. - Attention : grosse marche déjà pour avoir des adresses partout en France - et donc difficile de descendre jusqu'au logement - Dans le groupe de travail BAN : on a pour l'instant des lois qui mettent les communes jusqu'aux voies non closes. Il y a des endroits où les entrées de bâtiments sont loin, et on ne les aura pas dans la BAN car loin du périmètre des communes. - Il y a bien une entrée, qui a un ID Ban si possible (et peut être des adresses qui viendront d'ailleurs) et optionnel une position de l'entrée et un localisant de l'entrée de l'immeuble, de la maison etc... - Il faut que la localisation d'entrée soit optionnelle et complémentaire à la BAN.

Ma suggestion donc pour l'objet Entrée : - relation BAN - détail adresse si pas encore de lien BAN et en option (surtout site privé) : - position - libellé entrée

Remettre les relations communes : - Pas sûr que ça soit faisable. Ça fait beaucoup de gestion et qui fait la MAJ lors de la fusion. - Et on a des cas de bâtiments à cheval sur les deux communes. Et est ce que ça pourrait être MAJ via les flux d'ADS ? Mais quid des fusions de communes ? - Du coup, c'est gérer explicitement une relation bâtiment-commune. Les objets externes qui sont dans d'autres référentiel, il faudrait mieux les séparer pour ne pas créer des gestions ou des discussions. - Il faudrait qu'on puisse comprendre qu'elles sont les gestions propres du bâtiments et quelles sont les relations étrangères ?

Sur les référentiels externes : - Adresse BAN à utiliser dans Ref Bati - Mais Parcelle et local : ce sont ces référentiels qui utiliseront le référentiel Bâtiment ?

Quid d'une table d'historisation d'un bâtiment ? - Vrai question, relative à l'historisation du statut du bâtiment. - Il semble que ça serait important → dès que c'est collaboratif il faut pouvoir revenir en arrière et avoir des informations sur l'origine - Des communes on comprend qu'il peut y avoir des allers-retours aussi sur l'existence ou la nature d'un bâtiment - INSPIRE demande l'historisation de la filiation. - Est ce qu'on veut tout historiser ? Tous les champs c'est monstrueux. Et après on regarde comment on l'implémente techniquement - Idéalement on a des log historique et clé de valeurs qui indiquent le cycle de vie sur l'ensemble des attributs. - La table des évènements peut aussi faire office d'historisation (un évènement s'est passé à telle date et a modifié telle valeur)

- Est ce qu'on fait transparaître dans le modèle conceptuel le modèle historique ?
- Quid de l'objet "immeuble" inspiré des Suisse - pourquoi il n'y est plus ?
 - A défaut de mieux, on fait un lien avec un immeuble mais parce qu'ils n'arrivent pas à se coordonner.
 - Objet immeuble : pour le CSTB c'est important. On a besoin des deux niveau des représentations notamment pour la gestion.

- Dans ce cas ce n'est pas le même immeuble que les suisses. Pour eux, l'immeuble permettait de faire le lien entre les données du cadastre et donnée géographiques. Alors que pour le CSTB c'est une notion plus physique d'ensemble et d'agrégat de bâtiments. Plutôt au sens d'ensemble constructif cohérent.
- Différentes compréhension du coup de ce qu'est "l'immeuble" dans le MDD Suisse
- Du coup ça induit de se poser la définition d'immeuble -> dans la BDNB on a la notion de construction : si on a 3 bâtiments fiscaux qu'on ne sait pas spatialiser dans la BD Topo (à compléter)
- C'est sur que dans le stock/flux on aura du mal à descendre à la notion de bâtiment, MAIS on ramène à toutes les discussions que si j'ai deux objets alors il y aura beaucoup de nouveaux débats qu'on avait lors des discussions autour de la définition.
 - Si on ne le fait pas (à compléter)
- Pourquoi ne pas dire une table de relation pour dire qu'un bâtiment collé est lié à d'autres ?
 - Juste la relation topologique ne suffit pas, via la parcelle ça reste trop probabiliste.
 - Je pense qu'on a besoin d'un objet intermédiaire type "immeuble"
- La question ici est dire quand est ce qu'on a envie de les regrouper.

A faire

- Ouvrir un ticket avec vote ouvert sur le format de l'ID
- Regarder l'opportunité de clé de contrôle de l'ID
- côté OSM, voir avec François Lacombe qui a suivi le sujet Bâtiments
- Clarifier la notion d'immeuble et illustrer par des exemples d'usages/problématiques.