
NOTE POUR LE GROUPE DE PILOTAGE SANDRE

Principes pour la définition des identifiants

Titre : Principes pour la définition des identifiants

Créateur : Système d'Information sur l'Eau / Sandre

Éditeur : Ministère en charge de l'écologie et du développement durable

Contributeur : Office International de l'Eau / Sandre

Auteur : NC

Date : 23/02/2001

Version : 1

Type : Texte

Format : PDF

Identifiant : urn:sandre:note-methodologique:Identifiant::::ressource:1:::pdf

Langue : fra

Couverture spatiale : France métropolitaine

Couverture temporelle :

Droits d'usage : <http://creativecommons.org/licenses/by-nc-sa/2.0/fr>

Table des matières

1.. Principe généraux.....	2
1.1 Qu'est-ce qu'un identifiant.....	2
1.2 Types d'identifiant.....	2
2.. Contraintes dans la mise en place d'identifiant commun.....	3
2.1 Définir le code :.....	4
2.2 Définir l'organisation de la gestion du code :.....	4
2.3 Définir l'organisation de la mise à disposition du code :.....	5

📍 1. Principe généraux

➤ 1.1 Qu'est-ce qu'un identifiant

Un identifiant est un caractère ou un groupe de caractères constituant une donnée, employé pour désigner de façon unique un élément dans un ensemble et ainsi le distinguer des autres éléments de ce même ensemble. Dans un contexte informatique, la valeur d'un identifiant est la clé qui de sélectionner dans un fichier ou une table, un enregistrement parmi tous ceux que contient le fichier ou la table.

➤ 1.2 Types d'identifiant

Un identifiant relève de plusieurs typologies. La suite du paragraphe ne prétend pas être complet. Il ne vise qu'à présenter les principales catégories.

Identifiant simple / composé :

Un identifiant signifiant simple est un identifiant composé d'une seule donnée. Un identifiant composé est un identifiant qui résulte de la concaténation de plusieurs données. Exemple : l'identifiant de la commune est la concaténation du code du département avec le code de la commune au sein du département.

Identifiant signifiant / non signifiant :

Un identifiant signifiant est un identifiant qui comporte tout ou partie de l'information contenu dans les éléments de données décrivant l'objet auquel il se rapporte.

Un identifiant non signifiant est un identifiant qui ne comporte qu'une seule information : celle nécessaire à distinguer l'objet qu'il cible de son environnement.

Attention

Les identifiants signifiants sont à proscrire !

En effet, un système de codification signifiant présente notamment les défauts suivants :

- il ne prévoit jamais tous les cas. L'application d'un système peut s'avérer parfois impossible. Quel code attribue-t-on alors ? Comment faire la distinction entre les vrais codes signifiant et les autres ?
- il arrive toujours que l'on se trompe dans l'affectation d'un code. Ainsi, si il n'est pas possible de le modifier, on a un code en décalage avec sa signification. La banque Hydro codifie ses stations sur la base de la zone hydro où elles se situent. La banque a connu plusieurs erreurs de codification de sorte qu'elle contient plusieurs stations dont le code est différent de celui de la zone hydro où elle se situe. Si l'utilisateur se base alors toujours sur le code, il est trompé sur ces quelques stations.
- il est l'image à un instant t d'une organisation et d'une approche qui peut varier dans le temps. Or on demande à un code d'être pérenne. On peut donc être sûr, comme pour le point précédent, qu'au bout d'une période qui peut être courte (<10 ans), que le code soit en décalage avec sa signification (ex. les codes réseaux de l'INP).
- en résumé, le risque de tromper un utilisateur sur un code signifiant est élevé, et cette erreur cause à terme plus souvent d'inconvénients qu'un code non signifiant.

Exemple de codes identifiants signifiants mis en défaut

- **n° de sécurité sociale** : plusieurs personnes peuvent avoir le même code de sécurité sociale. Plusieurs cas existent de personnes qui sont nées à 100 ans d'intervalles et qui possèdent le même numéro. Les cas sont beaucoup plus fréquents pour les français nés à l'étranger (code département 99).
- **code département** : le désordre fut total dans les bases de données quand la Corse fut dotée de deux départements : 2A et 2B.
- **n° de sécurité social canadien** : le numéro de sécurité social canadien (comme le français) comporte une indication sur le sexe. Or rien qu'au Québec, officiellement plus de 6000 personnes changent de sexe.
- ...

2. Contraintes dans la mise en place d'identifiant commun

L'identifiant est l'un des outils indispensables à l'échange de données.

Le transfert de données ou la consolidation de données ne peuvent s'effectuer sans la mise en commun entre les partenaires de l'échange des identifiants.

Cependant, le partage d'un identifiant introduit plusieurs contraintes dont les principales sont décrites ci-après.

➤ 2.1 Définir le code :

2.1.1 Définir l'objet ciblé par l'identifiant :

Un identifiant ne peut être défini sans une définition claire de l'objet auquel il se rapporte et de son cycle de vie. En effet, de cette définition, on pourra en déduire au niveau de l'identifiant :

- **le champ d'application du code** : quels sont les objets qui peuvent être codés, quels sont ceux qui ne peuvent pas l'être, quels sont les limites de l'utilisation du code...

Par exemple, le SANDRE qui n'a pas de personnalité morale ne peut obtenir un code SIRET auprès de l'INSEE.

- **les règles de gestion du code** : le cycle de vie de l'objet permet de définir les règles et les critères de création, de modification et de suppression d'un code.

2.1.2 Définir la règle de construction d'un code :

La création, la modification et la suppression d'un code doivent suivre des règles quand bien même celles-ci sont rudimentaires.

Par exemple :

- débiter à 1 et incrémenter de 1 à chaque création,
 - les codes ne sont jamais supprimés, ils sont gelés...

➤ 2.2 Définir l'organisation de la gestion du code :

2.2.1 Définir la ou les instances en charge de la gestion du code

Un système de codification doit être mis en oeuvre par une ou plusieurs structures. Celles-ci doivent répondre à plusieurs conditions pour une codification pérenne :

- la structure doit maîtriser le système de codification et son mode de gestion,
- la structure doit être à même d'assurer le service de ce codification (dimensionner pour le nombre de demandes potentielles, règle de neutralité vis à vis des demandeurs potentiels...) et d'en assurer la continuité,
- la structure doit s'assurer que chaque code est unique et que chaque objet n'est codé qu'une fois,
- la structure communique à la structure de diffusion du code (cf. ci-après) les évolutions du système de codification sur une base qu'elles ont à définir...

Attention

L'identification d'une instance en charge de la gestion d'un code peut conditionner la définition du code lui-même

La gestion d'un code est lourde. Elle ne peut se faire raisonnablement si elle intervient dans le cadre d'une mission ou d'une action quotidienne de l'instance.

Identifier une instance de codification n'est pas une chose aisée. Plusieurs scénarios sont possibles :

- *identification d'un système de codification unique assuré par une ou plusieurs instances de codification dont c'est la mission* : rechercher dans la définition d'un identifiant à réutiliser au maximum ceux qui existent déjà : code INSEE des communes, Codes des forages du BRGM...

-
- *accepter plusieurs systèmes de codifications* en laissant à des communautés différentes la possibilité de désigner en leur sein un système de codification qu'elles utiliseront pour leurs échanges sans que ces systèmes aient la prétention d'être utilisés au delà de leur communauté notamment pour des raisons de coût et de légitimité que représenteraient la généralisation de leurs codes,
 - *accepter tous les systèmes de codifications* issus de toute structure qui souhaiterait gérer le code notamment pour ses besoins propres ; la difficulté de la gestion du code étant dans ce cas reportée à l'identification des structures.

2.2.2 Définir les modalités de saisie de cette structure pour la création/modification d'un code

Les demandes de création, de modification et de suppression de codes doivent faire l'objet de procédures claires et transparentes vis à vis de tout utilisateur.

Chaque procédure doit comporter au moins les indications suivantes :

- les acteurs et leur rôle,
- les règles de gestion du code,
- les critères à remplir,
- les droits et devoirs de chaque parties notamment les délais d'attribution d'un code ,
- les services associés à la création, modification et suppression (aide à la recherche des données indispensable à la création d'un code...).

➤ 2.3 Définir l'organisation de la mise à disposition du code :

2.3.1 Définir l'instance en charge de la diffusion du code

Pour être utilisé, un système de codification doit être accessible ou mis à disposition auprès des utilisateurs par une ou plusieurs structures (identiques ou non à celles qui gèrent le système de codification). Celles-ci doivent répondre à plusieurs conditions pour une utilisation pérenne :

- la structure doit maîtriser les modalités de la diffusion, les documenter si nécessaire et les promulguer ainsi que les système de codification lui même,
- la structure doit être à même d'assurer le service de mise à disposition (dimensionner pour le nombre de demandes potentielles, règle de neutralité vis à vis des demandeurs potentiels...) et d'en assurer la continuité,
- la structure récupère auprès de la structure gestion du code (cf. ci-avant) les évolutions du système de codification sur une base qu'elles ont à définir...

2.3.2 Définir les modalités de cette diffusion :

Les modalités de diffusion des codes doivent faire l'objet de procédures claires et transparentes vis à vis de tout utilisateur.

Chaque procédure doit comporter au moins les indications suivantes :

- les acteurs et leur rôle,
- les critères d'accès au code (gratuit/payant, libre/confidentiel, ...),
- les fréquences de mises à disposition des évolutions : période fixe, sur événement, ...,
- les vecteurs de diffusion : serveur Internet, liste de diffusion, diffusion papier...,
- les droits et devoirs de chaque parties notamment les délais d'obtention d'un code,
- les services associés à la diffusion...