
NOTE POUR LE GROUPE DE PILOTAGE SANDRE

Réflexions sur l'utilisation des codes Sandre gelés par les acteurs de l'eau du SIE

Titre : Réflexions sur l'utilisation des codes Sandre gelés

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 : Dimitri Meunier

Date : 13/03/2014

Version : 1

Type : Texte

Format : PDF

Identifiant : urn:sandre:note-methodologique:GelEtGenealogie::::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.. Contexte.....	2
2.. Objectif.....	2
3.. Jeux de données de référence.....	3
4.. Objet.....	3
4.1Identifiant de l'objet.....	3
4.2Attribut de l'objet.....	3
4.3État de l'objet.....	4
5.. Administration des objets.....	5
6.. Utilisation des objets.....	6
7.. Perspective.....	8

📌 1. Contexte

Le service d'administration nationale des données et des référentiels sur l'eau (Sandre) est chargé de diffuser les jeux de données de référence pour le système d'information sur l'eau (SIE) en application du [schéma national des données sur l'eau \(SNDE\)](#) (cf. [l'arrêté interministériel publié au JO du 24 août 2010](#)). Ces jeux de données de référence sont des listes d'objets relatifs au monde de l'eau clairement identifiés (des barrages, des ouvrages de prélèvement, des lieux d'analyses d'eau, des tronçons de rivière, des substances chimiques, des poissons (ie. taxons), etc). Ils sont disponibles librement sur le [site internet du Sandre](#) ; les acteurs de l'eau sont chargés de les employer dans leurs systèmes d'informations respectifs selon [l'architecture technique informatique du SIE](#).

📌 2. Objectif

Chaque objet principal du SIE est identifié par un code unique – exemple le barrage de Rimer qui a comme code 63829. Les acteurs de l'eau rattachent à chaque code un ensemble de données métiers qu'ils produisent tels les débits d'eau journaliers mesurés sur un barrage. Au cours du temps, un objet peut subir des changements d'état pour des raisons diverses. Prenons l'exemple d'un barrage détruit partiellement pour rénovation complète et remis en service 1 an après. Faut-il conserver le code du barrage rénové ou en attribuer un autre ? Dans ce dernier cas, faut-il supprimer l'ancien code du barrage ? Faut-il rattacher les anciennes données métiers au nouveau code... Selon la nature même de l'objet, les raisons qui ont conduit au changement d'état de l'objet et la manière dont les données sont gérées, les réponses sont multiples. Sans réponse à ces questions, les modalités d'utilisation des codes diffèrent d'un acteur à un autre. Cette note fournit des pistes de réflexions au groupe de pilotage du Sandre (GPS).

3. Jeux de données de référence

Ensemble de données de référence homogènes, codifiées et répertoriées. Les jeux de données de référence établis par le Sandre sont des données alphanumériques ou géographiques régulièrement évolutives. Les données alphanumériques sont des listes d'objets codifiés (par exemple, les codes des paramètres comme la demande en Chlore (1397), les méthodes, les intervenants, les nomenclatures, les stations de mesure...). Les données géographiques sont des objets géolocalisés (par exemple, les masses d'eau, les stations de mesure, les zones hydrographiques, les zonages réglementaires...). Chaque objet est associé à une liste d'informations complémentaires (code/identifiant, nom, localisation...).

4. Objet

L'objet constitue un ensemble d'informations homogènes du point de vue métier. Un objet métier correspond à un concept partagé au sein d'un dictionnaire de données établi par le Sandre. Un objet peut être concret c'est à dire issu du monde réel comme un barrage ou abstrait au sens issu d'un concept (exemple : le découpage des masses d'eau).

4.1 Identifiant de l'objet

Un objet référentiel du monde de l'eau (station hydrométrique, nappe souterraine, station d'épuration...) est identifié par un *code Sandre* unique pour ce type d'objet (pour ce concept) – exemple le barrage de Rimer n°63829. L'unicité d'un code signifie qu'un code d'un objet n'est pas réutilisable par un autre objet du même référentiel. Pour distinguer un objet vis-à-vis des autres du même type, un ensemble d'informations le caractérisant (ie. attributs) lui sont associées. Pour le barrage n°63829, il s'agit de son nom : « Barrage de Rimer », de sa localisation définie par des coordonnées géographiques (X=450322.364825123 et Y=6764019.5133874) et d'une projection, etc. Les attributs discriminants de l'objet constituent sa *clé d'identification*. Si la *clé d'identification* est toujours respectée au sein d'un jeu de données de référence, l'absence de doublon est théoriquement garantie. Les attributs secondaires apportent donc des connaissances complémentaires sur l'objet alors que les attributs discriminants permettent de vérifier les risques de doublons (deux codes différents pour un même objet).

4.2 Attribut de l'objet

L'attribut désigne ce qui est propre à un être ou à un objet. Les attributs (ou propriétés) sont des informations qui caractérisent un objet. Par exemple, l'objet « Station d'épuration » est caractérisé par des propriétés tels sa taille, son type... Les dictionnaires de données établis par le Sandre comportent des objets caractérisés par des attributs. Un attribut peut notamment être traduit physiquement sous la forme d'un champ d'une base de données.

L'objet - noté Obj - est défini comme suit :

Obj = (IDObj, AObj) où AObj = {A/Obj} où IDObj est l'identifiant de l'objet et AObj est l'ensemble de ses attributs discriminants et secondaires.

4.3 État de l'objet

La situation dans laquelle se trouve un objet à une date donnée représente son état. A sa création, l'objet est validé. Il porte le statut validé **StObj=validé**. Un objet peut être mis à jour au cours du temps sur tout ou partie de ses attributs **AObj** excepté sur la date de création de l'objet **DtCrObj** et sur son identifiant **IDObj**. Ces deux derniers restent invariants dans le temps. Notons qu'à chaque changement d'état de l'objet, sa date de mise à jour est actualisée **DtMajObj=[Date de mise à jour]**. Prenons l'exemple du barrage hydroélectrique dont le code est 3. En 2000 pour rendre hommage à un défunt député, le nom de l'objet **NomObj** et sa date de mise à jour **DtMajObj** ont été modifiés.

L'objet est *gelé* (**StObj=gelé**) dans au moins l'une des conditions suivantes :

- **La destruction** : C'est le fait que l'objet n'a plus d'existence dans le monde réel, ce qui amène à *geler* définitivement l'objet (**StObj=gelé**) ou à modifier la date de fin de vie de l'objet. Prenons un exemple, un paramètre chimique qui a été totalement *gelé* car ses caractéristiques (ses attributs **AiOBJ**) ne permettent pas d'identifier de façon équivoque ce paramètre, par exemple pour le paramètre d'**IDObj=2522** dont le nom est **NomObj=2,4-D-ester**. Il peut s'agir aussi de doublon comme 2 zones de pêche ayant les mêmes noms et localisées au même endroit, mais avec des codes différents. Notons que **la réapparition** (équivalent au *dégel*), est le fait de *valider* un objet qui a été *gelé*. Prenons l'exemple des plans d'eau 'Etang du mur' (code L9005103) et 'Etang du glandier' (code U1236023) présents dans la BD Carthage 2011 (Fond cartographique des rivières et lacs), disparus en 2012 puis réapparus dans la version de 2013.

Parfois un objet n'est plus utilisé mais il n'est pas détruit (notion de cycle de vie de l'objet) ; son statut reste validé (**StObj=validé**). Dans ce cas, un attribut de l'objet signale qu'il n'est plus employé comme la date d'arrêt d'activité de la station de mesure qualité. Cet attribut signale si la station est toujours en activité ou pas.

- **La recodification** : C'est le fait de *geler* un objet parce qu'il comporte au moins une erreur de cohérence sur des attributs **AiOBJ** discriminants. Prenons l'exemple du paramètre **IDObj=5348** **NomObj=Sulfonate de perfluorooctane** et sa formule chimique brute erronée **FbObj=C8-H-F17-O3-S**. Cet objet est *gelé* au profit du paramètre **IDObj=6561** **NomObj=Sulfonate de perfluorooctane** avec la bonne formule. Il existe aussi des contres exemples, quand l'objet de change pas d'identifiant malgré la modification de ses attributs. C'est le cas des communes qui peuvent changer de géométrie tout en conservant leurs codes (**IDObj**) ou aussi les contours des bassins qui changent sans que leurs codes (**IDObj**) soient modifiés.
- **La division** : C'est le fait de *geler* un objet au profit de n nouveaux. Une masse d'eau peut avoir été découpée en deux masses d'eau distinctes. Par exemple, les masses d'eaux côtières IC11 (Les Saintes) et IC02 (Pte du vieux fort-Sainte-Marie) dans la dernière version du référentiel MDO (rapportage 2010) sont issues d'une seule masse d'eau côtière dans la précédente version: IC2 (2-Pte du vieux fort-Sainte-Marie).
- **La fusion** : C'est le fait de *geler* n objets pour créer un nouvel objet. C'est le cas des communes, il arrive que 2 communes fusionnent en une seule.

L'état de l'objet - noté EtatObj - est défini comme suit :

EtatObj \neq (IDObj, AObj, RAObj, t) où t est le temps, **RAObj** = {RAObj} Où RAObj est l'ensemble des règles d'administration d'objet conduisant au changement d'état de l'objet.

5. Administration des objets

Les différents changements d'état d'un objet sont maîtrisés grâce à l'application de l'ensemble des règles d'administration d'objet conduisant au changement d'état de l'objet (**RAObj**). Cela garantit la cohérence et la qualité de l'ensemble des objets. L'ensemble des règles doit répondre aux questions relatives à l'administration des objets à savoir par exemple « Faut-il conserver le code du barrage rénové ou en attribuer un autre ? », « Dans ce dernier cas, faut-il supprimer l'ancien code du barrage ? »... L'ensemble des règles d'administration d'objets (**RAObj**) définit les modalités de collecte, de gestion et de diffusion des objets. Au sein du SIE, ces règles sont définies dans les documents de référence Sandre - exemple :

http://www.sandre.eaufrance.fr/IMG/pdf/sandre_doc_reference_FAN_2-0.pdf

Ces règles doivent obligatoirement hériter des contraintes sur les objets qui figurent dans les dictionnaires et scénarios d'échange Sandre à savoir (cf. Figure 1 – Méta modèle d'un objet) :

- le ou les attributs clé de l'objet,
- les multiplicités des attributs de l'objet,
- les types des attributs de l'objet,
- les liens (relations) entre l'objet et les autres
 - les natures des relations,
 - les types des relations,
 - les multiplicités des relations,
 - les rôles des relations,
 - les sens des relations.

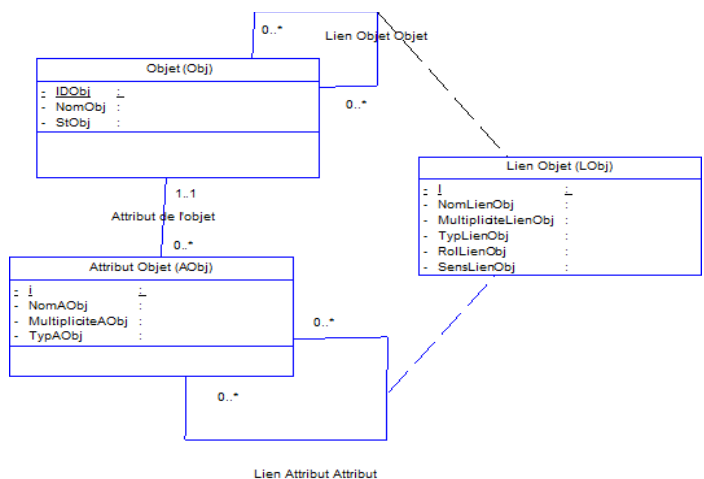


Figure 1 – Méta modèle d'un objet

Ces règles sont appliquées par l'administrateur d'objets. Au sein du SIE, les administrateurs sont multiples selon les domaines de l'eau à savoir :

- pour les stations hydrométriques le service central d'hydrométéorologie et d'appui à la prévention des inondations (Schapi),
- pour les stations de mesure de la qualité des eaux superficielles et continentales les Agences de l'eau...

6. Utilisation des objets

Les acteurs de l'eau reposent leurs données d'observation et d'évaluation sur les identifiants des objets (**IDObj**) produits par les administrateurs d'objets et distribués par le Sandre. Ayant la responsabilité de la cohérence de leurs données, ils sont chargés de respecter des règles métiers de bonne utilisation des objets. Les acteurs de l'eau doivent :

1. obligatoirement bancariser les codes des objets (**IDObj**) employés par leurs données métiers.
2. éventuellement bancariser tout ou partie des attributs des objets (**AObj**).
3. éventuellement bancariser les états des objets (**EtatObj**).
4. obligatoirement rattacher à leurs données métiers des codes d'objets (**IDObj**) connus.
Pour toute donnée nouvellement acquise, le dernier **IDObj** en date doit être utilisé.
Pour toute donnée acquise ou commandée anciennement, soit l'**IDObj** connu aux dates d'acquisition est utilisé, soit de préférence le dernier **IDObj** en date est utilisé.

Si l'objet est *gelé* (**StObj**=*gelé*), la raison du *gel* doit être prise en compte, dans le cas de :

- **la destruction** : Pour les données acquises ou commandée après la date du *gel*, l'usage de l'**IDObj** est interdit. En effet, il est inconcevable de rattacher des données à un objet théoriquement inexistant du monde réel. Pour les données acquises avant la date du gel, l'usage de l'**IDObj** est autorisé sans restriction.
- **la recodification** : L'usage de l'**IDObj** est interdit à compter de la date du gel. Pour les données acquises ou commandée avant la date du gel, l'usage de l'**IDObj** est toléré mais des restrictions particulières peuvent l'interdire.
- **la division** : L'usage de l'**IDObj** est interdit à compter de la date du gel. Pour les données acquises ou commandée avant la date du gel, l'usage de l'**IDObj** est autorisé mais des restrictions particulières peuvent l'interdire.
- **la fusion** : L'usage de l'**IDObj** est interdit à compter de la date du gel. Pour les données acquises ou commandée avant la date du gel, l'usage de l'**IDObj** est autorisé mais des restrictions particulières peuvent l'interdire.

Par conséquent, il est toléré de bancariser et d'échanger des identifiants d'objets (**IDObj**) *gelés* (**StObj**=*gelé*) selon certaines règles d'utilisation d'objets notées dans les **RUObj**. En effet, il est difficile d'empêcher l'usage d'identifiants d'objets (**IDObj**) *gelés* pour des raisons parfois techniques, organisationnelles et/ou thématiques – par exemple :

- L'acteur de l'eau saisit des données métiers dans son système informatique temporairement en panne d'Internet ; il n'est pas encore synchronisé avec celui du Sandre.
- L'acteur de l'eau diffuse des données métiers historiques qui n'ont pas encore été corrigées sur la base des derniers états des objets (**EtatObj**) car son expert est en congés.
- L'acteur de l'eau attend de nouveaux objets (**EtatObj**) demandés dans le cadre d'un texte réglementaire. Dans l'attente de sa publication, il utilise les anciens objets pour ses données métiers.

Les règles d'utilisation d'objets (**RUObj**) permettent notamment de répondre à la question « Faut-il rattacher les anciennes données métiers au nouveau code ». Mais à ce jour, le Sandre ne spécifie pas les règles d'utilisation d'objets (**RUObj**) ou de façon partielle au sein des dictionnaires et des scénarios d'échange. Pour information, le dictionnaire relatif à la description des stations de mesures de qualité des eaux superficielles et continentales indique notamment cette prescription d'usage : « pour une exploitation cartographique, statistique ou autre, des mesures effectuées, les données obtenues sur la station sont ramenées à un point : le point caractéristique de la station ». L'ensemble des acteurs de l'eau est donc amené à faire ses propres choix d'utilisation des objets. Il est donc courant de trouver à une même date, des données métiers de même nature employant des identifiants (**IDObj**) différents !

A ce jour, les acteurs disposent des différents états des objets (**EtatObj**) du SIE sur le site Sandre. De cette manière, ils peuvent analyser leurs données métiers avec les objets connus aux mêmes périodes. Ils disposeront - en plus en fin d'année 2013 via le Sandre - du chaînage structuré des états des objets (**EtatObj**) de sorte qu'un système informatique puisse le traiter de manière automatique ou semi-automatique selon la [nomenclature n°590](#) et une relations « père/fils » - cf. Figure 2.

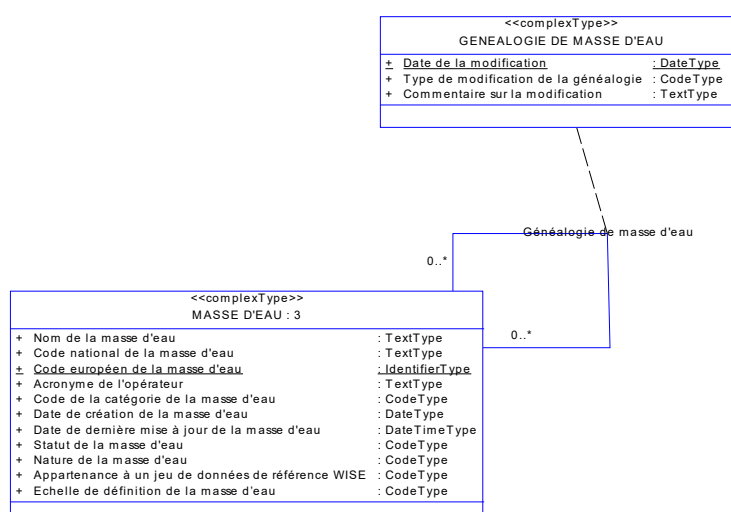


Figure 2 - Concept généalogie des masses d'eau

Notons que cette modélisation en UML – ici spécifique aux objets masses d'eau - est en faite théoriquement généralisable à l'ensemble des objets d'où une modélisation générique selon la figure 3 ci-dessous extraite du dictionnaire matrice reposant sur la [nomenclature n°695](#) :

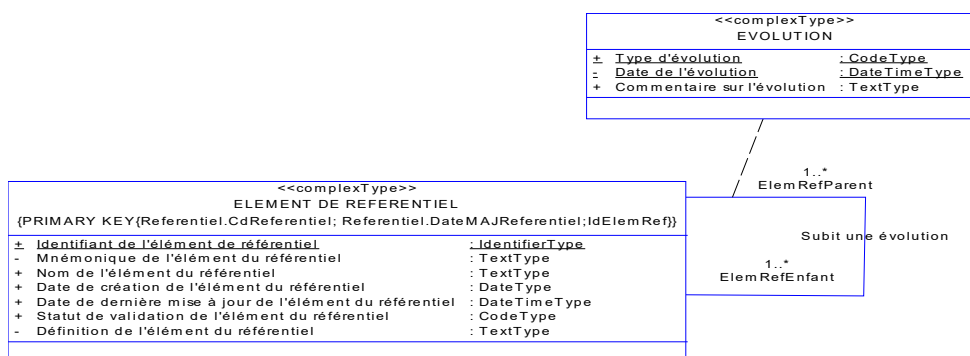


Figure 3 - Concept d'évolution d'un élément de référentiel

Notons que le 03/12/2013, les Administrateurs de données Sandre ont décidé de conserver la modélisation de la figure 3 avec quelques ajustements dont l'utilisation de la nomenclature n°590.

🔗 7. Perspective

Lorsque les données métiers de sources différentes ne respectent pas les mêmes règles d'utilisation des objets, l'utilisateur de ces données est alors amené à transformer les données pour les exploiter.

Prenons l'exemple du processus de commandes d'analyse d'eau entre commanditaires et laboratoires. Le commanditaire programme une liste de substances chimiques à mesurer à différents endroits d'une rivière. Le laboratoire réalise les analyses des substances puis transmet ses résultats au commanditaire. Imaginons que pendant le déroulement du processus, des **IDObj** (codes de substances chimiques) ont été *gelés*. Le commanditaire va devoir traiter le cas où le laboratoire a pris en compte ou pas ces changements. Prenons un autre exemple de processus, celui d'un indicateur de biologie comme l'indice biologique diatomée (IBD) qui est calculé à partir d'une liste d'**IDObj** (codes de diatomées) définie. Imaginons qu'entre deux phases de calcul de cet indice, des **IDObj** (codes de diatomées) ont été *gelés*. Il est probable que son résultat soit différent si le moteur de calcul ne prend pas en compte ces changements.

Pour définir les règles d'utilisation d'objets (**RUObj**), il faudrait notamment connaître tous les processus d'utilisation des objets. Face au nombre important d'acteurs de l'eau, d'applications informatiques et de pratiques différentes, commencer ce travail de définition sans méthodologie préalable apparaît risqué !