

SIMM

système
d'information sur
le milieu marin



SERVICE D'ADMINISTRATION
DES RÉFÉRENTIELS MARINS



Note de cadrage pour la création de modèles UML

Historique du document

Date de création : 30 avril 2021
Dernière modification : 01 juillet 2021
Version : non définitive

Version	Date	Auteurs	Action
0.1	30/04/2021	A. Rouyer, C. Rabevolo	Initialisation du document
1	31/05/2021	A. Rouyer, C. Rabevolo, S.Piel, C.Bonnet, G. Judeau	Précision des règles et ajout d'exemples
1.1	01/07/2021	A. Rouyer, C. Rabevolo	Précision sur les tables hiérarchiques et les tables associées

Contacts des membres du service d'administration des référentiels marins (SAR) :

Nom	Téléphone	Adresse électronique	Organisme
Steven Piel	0298338745	steven.piel@ofb.gouv.fr	Office français de la biodiversité Pôle maritime de Brest Direction surveillance, évaluation, données 16 quai de la douane 29229 Brest
Clémence Rabévol	0298224695	clemence.rabevolo@ifremer.fr	Service SISMER – Ifremer centre Bretagne ZI de la pointe du Diable 29280 Plouzané
Armelle Rouyer	0298224058	armelle.rouyer@ifremer.fr	Service SISMER – Ifremer centre Bretagne ZI de la pointe du Diable 29280 Plouzané

Table des matières

Préambule	4
1. Modélisation de la classe	4
1.1. Généralités	4
1.2. Notation	5
1.3. Langue	6
2. Modélisation des attributs	6
2.1. Formats de données des attributs (DataType)	6
2.2. Conventions de nommage	7
3. Associations	8
4. Stéréotypes des classes	9

Préambule

La mutualisation des pratiques entre les 3 systèmes d'informations fédérateurs de l'Etat (SI Eau, SI Milieu Marin et SI Biodiversité) est une composante importante permettant d'assurer l'interopérabilité.

Le SAR a opté pour une **modélisation UML** (Unified Modeling Language) avec le logiciel Enterprise Architect qui est le logiciel utilisé par de nombreux organismes de standardisation et normalisation.

Les **modèles conceptuels de données** créés par le SAR doivent :

- être compréhensibles et réutilisables facilement par les utilisateurs du SIMM,
- permettre la conception de la base de données par le groupe urbanisation,
- appliquer les recommandations des groupes de travail internationaux lorsque cela est nécessaire et réalisable par le SAR.

Suite aux interrogations du Service d'Administration des Référentiels (SAR) du SIMM pour la création de modèles conceptuels de données, l'équipe a décidé de rédiger une **note de cadrage** sur les pratiques à adopter. Le présent document définit donc les règles à suivre dans les modèles UML que le SAR diffusera. **Tous les modèles du SAR s'appuieront sur ces règles et permettront de générer les dictionnaires de données associés.**

1. Modélisation de la classe

1.1. Généralités

Le diagramme de classes est un schéma pour présenter les classes et les interfaces des systèmes ainsi que leurs relations.

Une **classe** est un ensemble de fonctions et de données (attributs) qui sont liées ensemble par un champ sémantique. Elle est caractérisée par des propriétés (attributs et méthodes) communes à un ensemble d'objets et permet de créer des objets ayant ces propriétés. Elle représente un concept au sein du système. Le nom de la classe doit évoquer le concept décrit par la classe.

NomDeLaClasse
- nomAttribut1 : type 1
- nomAttribut2 : type 2
- ...

Les **attributs** sont les composantes élémentaires de la description d'une classe. Ils définissent des informations qu'une classe ou un objet doivent connaître. Chaque attribut est décrit par un texte précisant sa définition, son format de données (datatype) et la liste éventuelle de valeurs possibles. Des informations complémentaires peuvent être ajoutées à la description de l'attribut : longueur, règles de nomenclature, nom du champ correspondant dans la table attributaire, rôle de l'attribut (identifiant), etc.

Chaque classe et chaque attribut sera décrit dans le dictionnaire de données.

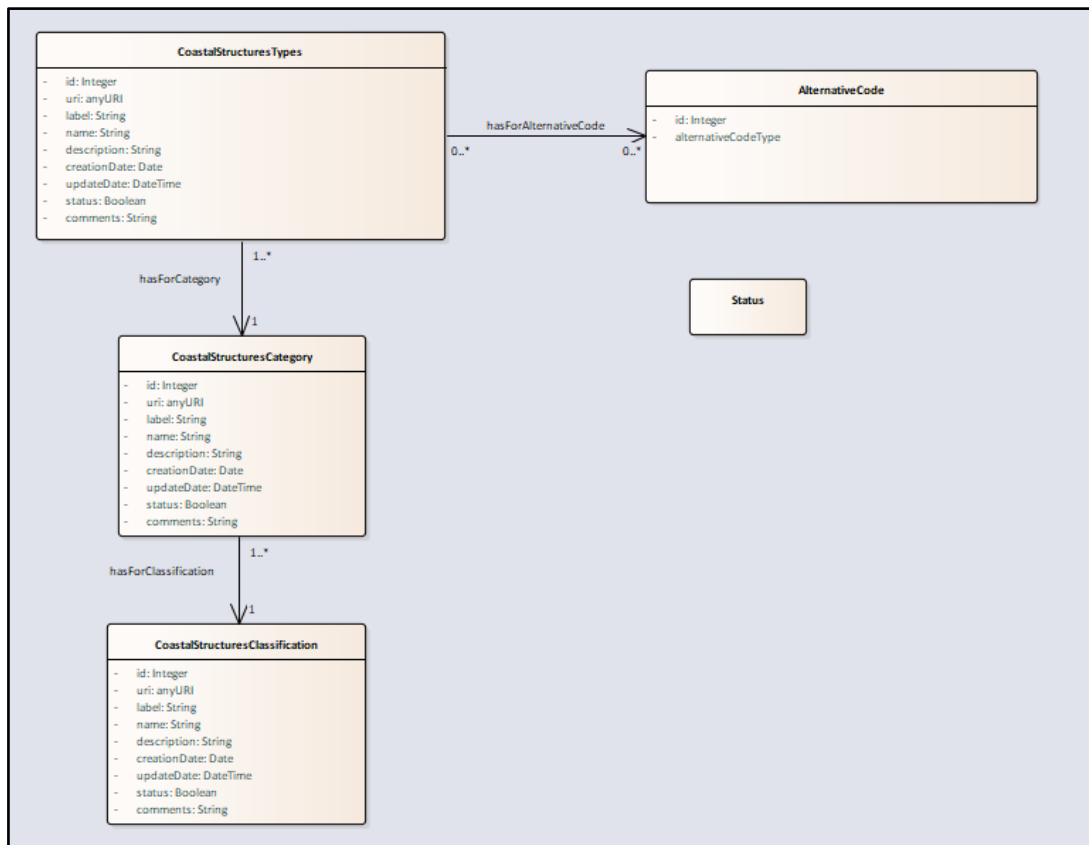


Figure 1: Exemple d'un modèle avec quatre classes.

1.2. Notation

Le SAR a pris connaissance de la norme ISO 19505/UML 2.4.1. Elle donne des recommandations sur les pratiques UML pour les modèles de données géographiques sur lesquels se basent les standards la majorité des communautés internationales autour de l'interopérabilité tels que OGC, INSPIRE, etc.

Par conséquent, le SAR utilisera ces préconisations dans les diagrammes de classe, qu'ils soient géographiques ou non. Il appliquera la **notation CamelCase** (sans '_') pour :

- **le nom de la classe devra débuter démarrer par du uppercase** : « NomDeLaClasse » (“Capitalize the first letter of class names (if the character set supports uppercase)”);
- **et l'attribut par du lowercase** : « nomAttribut » « nomRoleAssociation » (“Attribute names typically begin with a lowercase letter. Multiword names are often formed by concatenating the words and using lowercase for all letters, except for upcasing the first letter of each word but the first”).

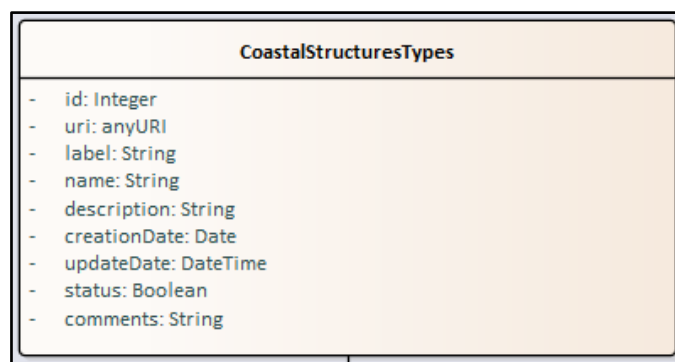


Figure 2 : Exemple d'une classe (nom en uppercase) avec des attributs (en lowercase).

1.3. Langue

Etant donné le public concerné par les modèles de données, la question de la langue utilisée s'est posée. Cependant, la volonté du SAR est de coller aux standards et pratiques de l'interopérabilité internationale, il a donc été décidé que les nommages des modèles de données créés seront en **anglais**.

Si les structures des extensions, modèles spécifiques qui graviteront autour sont en français cela créera un déséquilibre et au passage, empêchera de pousser/partager/enrichir ces structures de données définies par le SAR vers d'autres SI hors France (marins ou autre).

La définition des classes et attributs sera tout de même en français dans les dictionnaires de données (sauf si traduction de certains dictionnaires en anglais selon les besoins). Cette définition devra être la plus complète et claire possible afin d'éviter toute ambiguïté.

2. Modélisation des attributs

2.1. Formats de données des attributs (DataType)

Les formats ou types de données utilisés sont les suivants ceux des schémas XML, tels que recommandés par le W3C¹.

Format de données	Définition
string	Chaîne de caractères de longueur limitée
integer	Nombre entier
decimal	Nombre réel
date	Information sur le jour, mois et année

¹ Source : <https://www.w3.org/TR/2004/REC-xmlschema-2-20041028/datatypes#built-in-datatypes>

Format de données	Définition
dateTime	Information sur le jour, mois et année + heure et minute
boolean	Valeurs logiques (0, 1 ; true, false)
geometry	Format de données géométriques, (point, ligne, polygone, etc.) précisé dans la définition de l'attribut
anyURI	Référence à un identifiant uniforme de ressource (URI)

2.2. Conventions de nommage

Les attributs les plus utilisés dans les classes centrales des modèles seront les suivants.

Nom de l'attribut	Définition	Obligatoire	Format de données
id	Identifiant unique numérique incrémenté automatiquement dans chacune des classes	Oui	integer
uri	Identifiant d'une ressource sur le web	Oui	anyURI
label	Identifiant unique géré manuellement, ayant une forte signification pour les utilisateurs	Non	string
name	Libellé descriptif court	Oui	string
description	Définition de l'objet	Oui	string
comments	Commentaires	Oui	string
updateDate	Date de dernière mise à jour	Oui	dateTime
creationDate	Date de création	Oui	date
startDateValidity	Date de début de validité	Non	date

Nom de l'attribut	Définition	Obligatoire	Format de données
endDateValidity	Date de fin de validité	Non	date
position	Localisation de l'objet	Non	geometry
area	Surface de l'objet	Non	geometry

3. Associations

Les liens entre classes sont appelés associations. Chaque association est représentée par un trait simple surmonté à chaque extrémité d'une cardinalité. La cardinalité minimale (premier chiffre) et la cardinalité maximale (second chiffre) précisent l'implication de chaque classe dans la relation. Le signe * signifie que le nombre maximal peut être infini.

Le SAR retiendra les règles suivantes :

- La multiplicité des associations doit être renseignée (c'est-à-dire le nombre possible de connections). On notera 1 pour exactement une seule association, * ou 0..* pour plusieurs et 1..* pour au moins une.
- La direction des associations doit être précisée (avec une flèche) pour faciliter la lecture du modèle. On part de la classe « centrale » de notre modèle vers la classe liée.
- On peut indiquer un nom à l'association ou aux rôles dans le cas où il y aurait plusieurs associations ou en cas d'ambiguïtés.

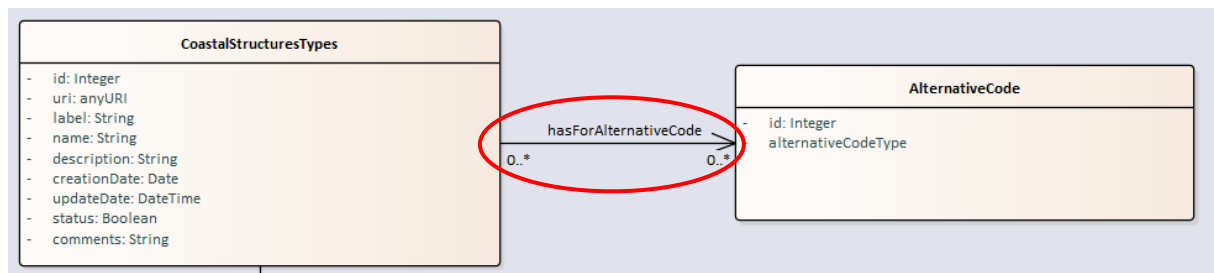
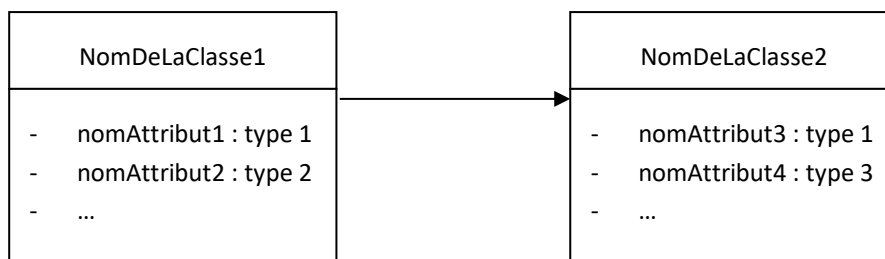
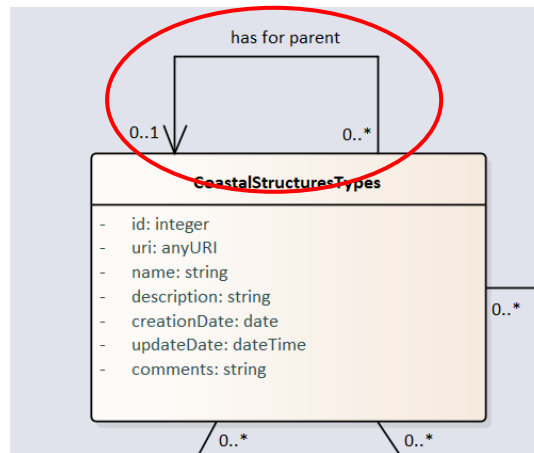
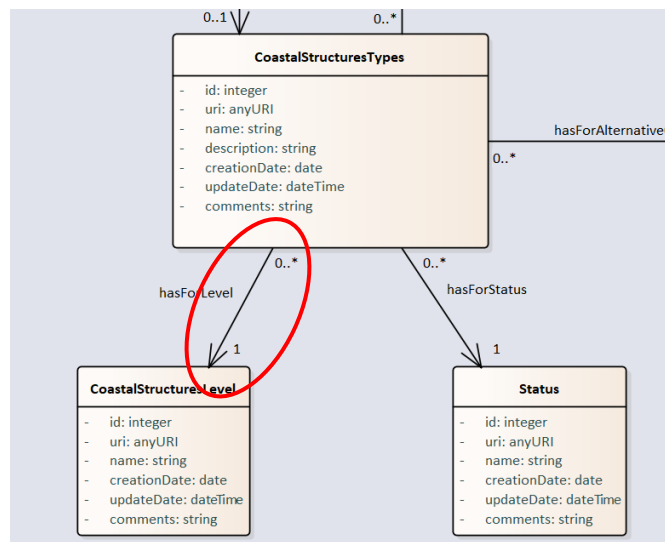


Figure 3 : Exemple d'association entre deux classes.

Pour modéliser une hiérarchie (par exemple type / classe / catégorie), on privilégiera une association bouclant sur une même classe plutôt que de faire une classe pour chaque niveau de la hiérarchie.



Pour modéliser des éléments présents dans une autre table (code liste), par exemple le statut de validité de l'élément. On ne crée pas un attribut avec comme datatype la classe, mais on crée une association entre notre classe d'intérêt et la classe qui contient la code liste.



4. Stéréotypes des classes

Au-dessus du nom de la classe, un stéréotype peut être renseigné entre cotes (<stereotype>). Un stéréotype permet de caractériser une classe d'un type particulier et de préciser implicitement les concepts qui y sont liés.

Pour le moment, il a été décidé de ne pas utiliser de stéréotypes dans les modèles du SAR afin qu'ils restent facilement compréhensibles pour les utilisateurs du SIMM. En effet, les stéréotypes ne sont

pas des concepts universels, il est parfois difficile d'en trouver facilement la définition et ne seront donc pas compréhensibles par tous.

Toutefois, il pourra être envisagé à l'avenir de stéréotyper certaines classes si le besoin de le faire est exprimé.

