

SIMM

système
d'information sur
le milieu marin



Point d'information : référentiel des « paramètres » du
SIMM : P01 du BODC

2 février 2021

● Historique

- Juin 2020 : choix du Groupe de pilotage du Langage Commun **d'adopter le P01 du BODC** comme référentiel des « paramètres » pour le SIMM

↳ Mais : nécessité de **traduire** le référentiel et d'assurer le lien avec le **référentiel analytique du SANDRE** (interopérabilité inter-SI fédérateurs)

- Fin 2020 : mise en place d'un **projet avec le BODC** de traduction et transcodage du P01

↳ Choix de travailler pour commencer sur les PSFMU du référentiel analytique du SANDRE utilisés dans **Quadrigé**

- 1- **Transcodage** entre les PSFMU Sandre et les vocabulaires contrôlés du BODC à partir de l'extraction de Quadrigé
- 2 – **Traduction** des termes du BODC en français
- 3 – **Intégration** des labels en français dans la base du BODC, et **création des P01** équivalents aux associations PSFMU de Quadrigé
- 4 – Développement d'une **API/web-service** d'accès aux P01 en français sur le portail du SAR avec le transcodage avec les codes SANDRE



- Support de **l'équipe du BODC** tout le long du projet
- Arrivé en mars d'un **CDD SIMM**, dont 4 mois sur ce projet avec le SAR

SIMM

système
d'information sur
le milieu marin



Nouvelle version du document : « Principes généraux de qualification des données du SIMM »

2 février 2021

● Principes généraux de qualification des données du SIMM version 0.5

- Rôle des producteurs et banques du SIMM dans la qualification des données diffusées
 - Communication sur les procédures d'assurance et de contrôle qualité
 - Qualification des données du SIMM par les producteurs et banques
- Communication des informations de qualification dans le cadre du SIMM
 - A l'échelle du lot de données
 - A l'échelle de la donnée



Principes généraux de qualification des données du SIMM version 0.5

➤ Nouveautés par rapport à la version précédente :

- Ré-organisation du document
- Ajout d'encarts type « boîte à outils » présentant des outils/bonnes pratiques concrètes
- Ajout d'exemples
- Ajout de schémas synthétiques



SIMM

Le SIMM pourra s'en inspirer, avec certains ajustements selon les domaines métiers, si nécessaire (par exemple, discrimination de plusieurs cas où la donnée peut être jugée comme douteuse.

année	année	année	année	année	année	année
2010	2011	2012	2013	2014	2015	2016
2017	2018	2019	2020	2021	2022	2023

Exemple d'utilisation des valeurs de qualité dans le SI métier. Niveau du SI. La valeur associée à chaque valeur informe d'un coup d'œil l'utilisateur sur la qualité de la donnée (si vert = correct).

Schéma synthétique du cycle de qualification d'une donnée

Une réflexion doit être menée sur la manière de communiquer les informations relatives à la qualité de la manière la plus complète possible pour les utilisateurs « experts », mais également d'une manière suffisamment accessible pour le « grand public ».

10

SIMM

dans le cadre du SIMM. Il sera considéré comme le niveau le plus à jour aux vues des connaissances scientifiques et techniques et celui qui capitalise le plus d'expertise.

Lorsque des données ont été produites il y a plus d'une dizaine d'années et qu'elles n'ont pas encore été qualifiées, il faut admettre qu'elles ne le seront probablement jamais. En effet, il sera trop difficile d'obtenir toutes les informations nécessaires à leur qualification, qui n'ont pas forcément été bien renseignées ou ont pu être perdues.

Le Service d'Administration des Références (SAR) s'engage à produire des standards qui renseignent sur la manière de structurer les données pour les diffuser dans le cadre du SIMM. Ces standards seront définis en fonction des besoins, selon les domaines métiers ou les rapports aux politiques publiques. Le SAR pourra y inclure des champs de type « métadonnées » et de type « qualification » comme cités précédemment afin d'assurer la communication des informations sur la qualité des données à l'échelle de chaque enregistrement. **Les producteurs devront être accompagnés dans le renseignement de ces champs afin de s'assurer de leur complétude.** Le SAR devra pour cela créer des tutoriels voire des formations à destination des producteurs, et proposer des outils pour faciliter l'utilisation des standards : test d'un fichier d'échange en ligne, générateur de base de données, etc.

Le SAR pourra utiliser, lors de la création des standards du SIMM, les standards internationaux, développés notamment par l'OGC, ISO et INSPIRE. En effet, ils intègrent des attributs sur la qualité des données, comme par exemple *Observation and Measurement* de l'OGC (ci-contre).

Exemple :

Dans le standard *Observation and Measurement*, l'observation est l'élément central du modèle et correspond à l'action de mesure d'une valeur.

Elle est liée à différents éléments qui vont venir apporter les informations nécessaires à la compréhension des données (le résultat, l'objet d'intérêt, la méthode employée, la propriété mesurée, etc.).

Agg Observation schématisée d'un exemple d'utilisation de l'OGC

10