



Identifiants numériques sur des données océanographiques



Philippe Téchiné, Gaël Alory, LEGOS/OMP Toulouse
avec le support du SEDOO/OMP Toulouse



Digital Object Identifier (DOI)

- ✓ DOI = Identifiant numérique unique attribué de manière pérenne à un objet : Publication, jeu de données, rapport, documentation, etc.
- ✓ L'INIST (Institut de l'Information Scientifique et Technique, <https://www.inist.fr>) du CNRS est habilité à délivrer des DOI
 - ➔ Membre du consortium international DataCite (<https://datacite.org>) = Agence d'enregistrement des DOI
 - ➔ Tout organisme ayant passé un contrat avec l'INIST peut attribuer des DOI
- ✓ Attribution de DOI par l'INIST
 - ➔ Forfait annuel de 180 € HT pour un nombre illimité de DOI
 - ➔ Fourniture d'un préfixe unique de DOI
 - ➔ Accès à des outils de test et de création des DOI

3 conditions pour poser un DOI sur des données

1. Un nom de DOI = **préfixe** (unique, donné par l'organisme d'attribution du DOI) + **suffixe** (à choisir librement)
Exemple : **10.6096 / SSS-LEGOS**
 2. Remplir un fichier de métadonnées au format XML contenant 6 propriétés obligatoires (figure 1) + d'autres recommandées ou facultatives (figure 2)
 3. Créer une *landing page* qui contient un descriptif des données + une URL stable pour accéder aux données
- ✓ La *landing page* peut contenir des figures, préciser les conditions d'accès aux données, la façon de citer les données, etc.
 - ✓ Le DOI n'est pas modifiable mais il est possible de modifier la *landing page* et les informations qu'elle contient

ID	Property	Obligation
1	Identifier (with mandatory type sub-property)	Mandatory
2	Creator (with optional given name, family name, name identifier and affiliation sub-properties)	M
3	Title (with optional type sub-properties)	M
4	Publisher	M
5	PublicationYear	M
10	ResourceType (with mandatory general type description sub-property)	M

ID	Property	Obligation
6	Subject (with scheme sub-property)	Recommended
7	Contributor (with optional given name, family name, name identifier and affiliation sub-properties)	R
8	Date (with type sub-property)	R
9	Language	Optional
11	AlternateIdentifier (with type sub-property)	O
12	RelatedIdentifier (with type and relation type sub-properties)	R

13	Size	O
14	Format	O
15	Version	O
16	Rights	O
17	Description (with type sub-property)	R
18	GeoLocation (with point, box and polygon sub-properties)	R
19	FundingReference (with name, identifier, and award related sub-properties)	O

Figure 2. Propriétés recommandées (R) ou facultatives (O).

Figures 1-2 extraites du schéma de métadonnées de DataCite, version 4.3, 2019 (<https://schema.datacite.org/meta/kernel-4.3/>).

La citation recommandée du jeu de données contient tous les attributs obligatoires : Creator (PublicationYear): Title. Publisher. (resourceTypeGeneral). Identifier

Exemple du Service National d'Observation Sea Surface Salinity (SNO SSS)

- ✓ Dans le cadre du SNO SSS (voir poster ROSAME / SSS dédié), nous avons demandé la mise en place de DOI au Service de Données de l'OMP (SEDOO), habilité par l'INIST à en attribuer, sur les données de salinité temps différé validées et sur les produits grillés SSS distribués à la communauté scientifique.

En pratique, après avoir renseigné les métadonnées dans le fichier XML

- ✓ Premier DOI **10.6096/SSS-LEGOS** créé par le SEDOO en 2015 sur les données SSS temps différé validées
- ✓ *Landing page* (figure 3), hébergée sur l'interface web de diffusion des données SSS, reflète les métadonnées disponibles sur DataCite
- ✓ Choix d'un DOI « permanent » plutôt que de demander un nouveau DOI à chaque mise à jour des données
- ✓ Après chaque modification ou ajout de nouvelles données
 - ➔ Mise à jour des métadonnées recommandées ou facultatives
 - **Version** (propriété 15), en modifiant le chiffre avant la virgule si ajout de données, après la virgule si correction des données
 - Période d'acquisition (propriété 8, **Date**; Temporal coverage dans la figure 3)
 - **Description** (propriété 17; Abstract dans la figure 3)
 - Liste des partenaires si besoin (propriété 7, **Contributor**)
 - ➔ Archivage pour retrouver l'état du jeu de données utilisé dans les publications
 - ➔ Mise à jour d'un historique des ajouts et corrections sur les données, pour les utilisateurs (champ « History »)
- ✓ Trace des utilisateurs et données téléchargées dans les statistiques
- ✓ Dernière version des données disponible en libre accès, versions précédentes disponibles sur demande
- ✓ Sur le même principe, des DOI ont été posés sur les produits grillés et données historiques (<http://sss.sedoo.fr>)

Attribuer un DOI à un objet permet d'augmenter sa visibilité

- ✓ L'attribution de ces DOI permettra de mieux évaluer le nombre et la nature des publications utilisant les données et produits grillés SSS qui sont en accès libre, sans restriction. Origine et traçabilité des données utilisées dans les publications sont de plus en plus demandées par les éditeurs et les reviewers.
- ✓ Lecture et références
 - ➔ Dernier schéma de métadonnées DataCite : <https://schema.datacite.org>
 - ➔ Documentation : DataCite Metadata Working Group. (2019). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 4.3. DataCite e.V. <https://doi.org/10.14454/7xq3-zf69>

The screenshot shows the landing page for the SSS data base web access. The URL is <https://dx.doi.org/10.6096/SSS-LEGOS>. The page title is "SSS data base web access". The main content area displays the following metadata:

- Title:** Sea Surface Salinity data from Voluntary Observing Ships Network
- DOI:** 10.6096/SSS-LEGOS
- Version:** 11.0
- Temporal coverage:** 1993-07-01 - 2019-06-28
- Publisher:** The French Sea Surface Salinity Observation Service (SSS OS)
- Publication year:** 2002

Below the metadata, there is a list of creators and an abstract. The abstract describes the SSS data base and its use. The History section shows a list of updates and corrections to the data base, with red arrows pointing to the Version, Temporal coverage, and History sections.

Figure 3. Landing page des données SSS temps différé validées. Les flèches rouges indiquent les métadonnées mises à jour.