

# Journées techniques SINP 20-21 mars 2019

## Bilan de la table ronde

### Introduction

À l'occasion des deuxièmes journées techniques SINP, une table ronde est organisée sur le thème de « Produire, partager et utiliser des métadonnées de biodiversité : Enjeux et perspectives ».

Cette table ronde s'est déroulée grâce au concours de 7 intervenants :

- Grégoire Lois - Directeur-adjoint du programme de sciences participatives Vigie-Nature - UMS PatriNat (AFB/CNRS/MNHN)
- Franck Binjamin - Chargé d'études « ingénierie de la donnée flore et habitats » au Pôle Coordination technique des conservatoires botaniques nationaux - AFB/DRED/PCCBN
- Annie Birolleau - Cheffe de projet Coordination des données du milieu marin- AFB/DRED/DSOD
- Simon Chagnoux - Responsable des applications scientifiques - Direction des Systèmes d'Information - MNHN/DSI
- Yvan Le Bras - Chef de projet Pôle National des Données de Biodiversité - UMS PatriNat (AFB/CNRS/MNHN)
- Isabelle Witté – Analyse de données Inventaires nationaux - UMS PatriNat (AFB/CNRS/MNHN)
- Pascal Dupont – Chef de l'équipe « Connaissance espèce – référentiel et inventaires » - UMS PatriNat (AFB/CNRS/MNHN)

Chaque intervenant a apporté son regard sur le thème de la table ronde à travers une présentation courte (présentation disponible sur Nature France, à la suite des journées techniques) et des échanges avec les participants.

Animée par Laurent Poncet (UMS PatriNat) pour la première partie de la table ronde sur la production, la gestion et le partage des métadonnées, et Julien Touroult (UMS PatriNat) pour la deuxième partie de la table ronde sur l'utilisation des métadonnées, cette table ronde a réservé une place importante aux échanges avec les participants. La présente restitution formalisée par Thomas Milon (UMS PatriNat) a été réalisée en fin de table ronde afin de partager la synthèse avec les participants du 21 mars.

### Éléments clés de cette table ronde

#### Le concept de métadonnées

La métadonnée

- c'est « ce qui est implicite pour le producteur de la donnée ». Il y a une forme d'expertise de la donnée dans les métadonnées. Cette expertise est partagée (On « sait » des choses sur les données/métadonnées) mais pas toujours explicitée dans les métadonnées.
- « c'est pas un truc d'informaticien ». Au contraire, elle permet d'environner la donnée d'un point de vue « métier ».
- ça permet de décrire les processus (ex : les données portent la validation, les métadonnées portent le processus de validation)



**Enjeu** : En fonction des systèmes d'information, les définitions autour des concepts de métadonnées sont différentes (ex : collecte (sur le terrain) et acquisition (notamment l'informatisation)). Pour assurer une interopérabilité sémantique, un travail d'harmonisation semble intéressant à aborder.

### Production des métadonnées

**Enjeu** : Difficile de motiver des producteurs à renseigner le cadre d'acquisition = besoin d'argument fort. Quels sont les moteurs de la production des métadonnées ? Il faut de l'intérêt pour le producteur.

Quelques moteurs ou solutions ont été cités :

- masquer la complexité de la métadonnée à travers la mise en place d'outil simple
- adapter les standards de métadonnées pour répondre aux besoins des producteurs.
- mettre en avant les jeux de données « de qualité » pour inciter les producteurs à mieux renseigner les JDD.

**Enjeu** : Harmonisation des pratiques dans le renseignement des métadonnées. Deux acteurs peuvent, pour un même jeu de données, décrire 2 fiches de métadonnées différentes. Granularité = comment découper les cadres d'acquisition et jeux de données ? De plus, la frontière entre données et jeux de données est parfois difficile à identifier.

### Perspectives :

- Pousser à une meilleure utilisation des standards (notamment nomenclature et référentiels). **Attention**, il est important de faire évoluer le standard de métadonnées pour que celui-ci soit plus proche des besoins des producteurs
- Implémenter des ontologies et thésaurus existants
- Le guide de production des métadonnées, qui sera prochainement accessible, propose des bonnes pratiques pour découper et renseigner les métadonnées

**Enjeu** : Qui doit produire les métadonnées ? Tout est dans la tête du producteur (implicite). C'est lui qui est le plus à même de produire ces métadonnées. Cependant, il faut trouver les moteurs pour le motiver à les produire. Il n'a pas toujours le temps

### Perspectives :

- Identifier clairement que c'est le producteur qui produit le jeu de données et maître d'ouvrage qui produit le cadre d'acquisition
- Calibrer le projet pour qu'il y ai des moyens concernant la production des métadonnées (au titre de la documentation du projet)

**Enjeu** : Les métadonnées peuvent être produites à partir des données. Certains champs sont même remis en cause, car calculable à travers les données. Ces champs sont historiques et ont vocation à être utilisés dans le cadre d'échange entre plateformes. Certains champs ne sont pas calculables (souvent dans le cadre d'acquisition) et doivent de toute façon être renseignés par le producteur.

**Enjeu** : Faut-il cadrer la production des cadres d'acquisition au niveau régional pour assurer une certaine harmonisation.

**Perspective** : C'est le porteur du programme qui doit le décrire et communiquer pour qu'il soit utilisé.



**Enjeu** : « Standard de métadonnées = contraint mais harmonisé » VS « Format spécifique = Richesse et flexibilité mais difficilement consolidable ? »

### Gestion des métadonnées

**Enjeu** : les producteurs sont plus ou moins avancés dans le processus de productions de métadonnées. Comment gérer cela de manière cohérente ? Comment structurer mes jeux de données si les producteurs ne produisent pas ?

Plusieurs administrateurs de données régionaux choisissent de conserver cette hétérogénéité des jeux de données et travaillent à accompagner dans l'amélioration dans la production des métadonnées (ex : promouvoir le modèle « Un JDD = un protocole »)

**Enjeu** : Où mettre les métadonnées ?

**Perspectives** : Plusieurs solutions ont été présentées ;

- Utilisation de fiches de métadonnées et de catalogue pour les publier.
- Description directement dans les données pour assurer une meilleure traçabilité
- Format « Datapackage » qui rassemble les données et métadonnées

**Enjeu** : Comment gérer un grand nombre de métadonnées ? dans le temps ? Comment gérer les évolutions des métadonnées et des jeux de données ? Obsolescence des fiches de métadonnées :

### Partage des métadonnées

**Enjeu** : le partage successif de métadonnées peut engendrer la dégradation de ces métadonnées. Les métadonnées sont différentes dans le SINP et dans les producteurs.

Perspectives :

- Unifier la publication du jeu de données
- Travailler le standard pour que les différents acteurs de la donnée s'y retrouvent
- Améliorer l'appropriation des standards.

### Utilisation des métadonnées

Quelles sont les utilisations des métadonnées ?

- **Pour le suivi** : pour s'assurer que les données sont déposées, permet le suivi de programme (notamment pour les financeurs)
- **Pour faciliter la réutilisation** : notamment dans le cadre de la recherche, pour filtrer les jeux de données d'intérêt (importance des informations sur le protocole)
- **Pour la visibilité** : pour bien référencer les organismes. Il y a une attente d'amélioration sur ce point, dans la réutilisation, de mettre plus en avant tous les acteurs et procédures autour des jeux de données.

**Enjeu** : Comment savoir qui utilisent les métadonnées ? (Mesure d'impact)

Une réponse passe par la mise en place d'identifiants uniques (DOI, UID) – solution réalisée par le GBIF

**Enjeu** : Pousser l'utilisation de métadonnées. Dans certains cas, les métadonnées ne sont pas utilisées (ex : certaines études d'impacts)