

GT Standards espèces

Compte-rendu 1ère réunion

3 JUIN 2022 / 09:00-12:00 / VISIOCONFÉRENCE

PARTICIPANTS (CERCLE 1)

ARB-Bourgogne-Franche-Comté
CBN du Bassin parisien
DEAL Réunion
DREAL Hauts-de-France
FAUNA
PatriNat (OFB-CNRS-MNHN)
Picardie Nature

Déborah BECHTEL
Silvère CAMPONOVO
Rémi BOUILLY
Valérie RAEVEL
Paul FROMAGE
Pascal DUPONT
Mathieu MANCEAU
Noëlie MAUREL
Olivier NORVEZ
Sophie PAMERLON
Kevin PLAETEVOET
Solène ROBERT
Frédéric VEST
Chloé VINET
Adèle HURABIELLE

EXCUSÉS (CERCLE 1)

OFB
LPO
MTE/DEB
PatriNat (OFB-CNRS-MNHN)
PatriNat (OFB-CNRS-MNHN)

Julie CHATAIGNER
Laurent COUZI
Jean-Jacques RICHARD
Jeanne DE MAZIÈRES DE CHAMBON
Eva RODINSON

ORDRE DU JOUR

Standardisation hors SINP

- Le rapprochement avec le Darwin Core a déjà été initié avec la version 4 du standard de données.
- Cette approche permet également de participer à l'évolution du Darwin Core au niveau international (participation aux conférences).

Bilan enquête et audit 2022

- Concepts/extensions/attributs utilisés et les informations véhiculées :
 - Dépend des données reçues. Des concepts ne vont être indispensables que pour certains jeux de données particuliers.
 - Proposition : analyser le contexte d'utilisation (jeu de données) de ces concepts/extensions/attributs avant de se poser la question de leur suppression.

Statistiques sur la base de données INPN

- notion de méta-cadre -> voir travaux du GT Métadonnées engagés sur le sujet.
- Attributs additionnels:
 - constat : les attributs additionnels sont beaucoup utilisés pour renseigner de l'information qui pourrait être renseignée avec des attributs existants (donc standardisable).
 - Le concept Attribut Additionnel doit véhiculer de l'information **non standardisée/standardisable**.
 - Proposition : création de scénario d'échange ou de 'profils' de données pour les attributs additionnels qui seraient récurrents.
- Nomenclature de l'attribut "objectifJdd" (métadonnées jeu de données) :
 - remarques sur la complexité de la liste de valeurs.
 - toutes les valeurs de la nomenclature ont été utilisées au moins une fois.
 - Propositions :
 - avoir une sémantique derrière chaque valeur de nomenclature pour éviter les incompréhensions des non-experts (exemple : "inventaires généralisés et exploration" et "inventaires de répartition", quelle est la différence ?).

- regroupement de valeurs dont la sémantique n'est pas claire (cf proposition ci dessus) et dont le nombre d'utilisations est faible ?

Standards de données v4 et métadonnées 1.4 (non publiés, GT 2019)

- Le concept Évènement correspond à un relevé simple (remplace le regroupement d'observations).
- Propositions :
 - ajouter une récursivité au regroupement d'évènements.
 - renommer le concept regroupement évènements en "plan d'échantillonnage" pour une meilleure différenciation avec le concept évènement.
 - ajout d'un concept site d'étude (-> voir avec les travaux du GT standard Habitats).
 - métadonnées du jeu de données :
 - Suppression des attributs restrictionDonneesPersonnelles et diffusionPrecise.
- l'attribut dSPublique peut être conservé mais ne peut pas être utilisé pour déterminer un niveau de floutage.

Propositions d'évolutions pour les nouvelles versions 2022

Axe A - Rapprochement avec le Darwin Core

- BasisOfRecord : qualification de la donnée qui correspond à plusieurs attributs du standard SINP.
 - Propositions :
 - faire évoluer le mapping GBIF<->SINP actuel
 - proposer au GBIF/TDWG une évolution de l'attribut dans le Darwin Core
 - proposer un nouvel attribut dans le Darwin Core qui remplirait plus de critères du SINP
- PathWay : c'est une information que l'on a très rarement au niveau de la donnée d'observation, plutôt au niveau de l'espèce. Son utilisation pour le standard données SINP n'est pas retenue.
- DegreeOfEstablishment/EstablishmentMeans/OccurrenceStatus/LifeStage :
 - correspondance avec des nomenclatures SINP.
 - nomenclatures hiérarchisées. Possible de reprendre cette hiérarchisation mais est-elle seulement pour présentation ou gérée comme ça ? -> Sophie se renseigne pour le GBIF.
 - Si ces vocabulaires contrôlés sont repris dans le SINP, il faudra gérer les correspondances avec les anciennes versions pour les données historiques.
 - DegreeOfEstablishment/EstablishmentMeans : plutôt liées à l'espèce.

Axe B - améliorer et mettre à jour les standards

- Restructuration du standard de données
 - Propositions retenues :
 - création du concept SourceType à la place des concepts Source, SourceEvenement, SourceRegroupement
 - création du concept Localisation à la place de Commune, Mailles, Département.
Un évènement ne peut avoir qu'une seule localisation.
 - Autres propositions :
 - Faut-il conserver la valeur Maille10x10 ? -> Vérifier s'il y a de l'acquisition à cette échelle.
 - Un évènement peut avoir une seule localisation -> la localisation la plus précise ?
 - typeInfoGeo à supprimer ?
- Identification du taxon et validation scientifique (rapprochement avec le Darwin Core et aide à la validation scientifique)
 - Propositions retenues :
 - création d'un concept pour l'identification du taxon reprenant des attributs de SujetObservation et DescriptifSujet
 - occMethodeDetermination : transformation de l'attribut en liste fermée -> voir avec proposition Pascal (morphologie externe sans et avec manipulation, morphologie interne, génétique, chimie, son)
- Mise à jour des nomenclatures (ce sujet fera l'objet d'un document de travail partagé et de réunions spécifiques au besoin)
 - point d'attention sur les nomenclatures qui ont des valeurs séparées par slash
- Gestion des complexes taxonomiques
 - TaxRef v16 : agrégats même clé primaire que les espèces cdNom (également cdRef) -> changement sans impact sur le standard.

Les points suivants n'ont pas pu être traités au cours de la réunion, ils seront listés dans le document de travail partagé.

Axe C - améliorer la cohérence entre les standards du SINP

- Cohérence des concepts, nomenclatures et attributs
 - PersonneType/OrganismeType
 - DateType

- Lien avec les données habitat : concept LienHabitat
- Harmonisation de la documentation entre les standards d'échange du SINP

Axe D - améliorer l'accompagnement autour des standards

- Documentation plus pédagogique
- mieux définir le périmètre et expliquer les objectifs d'un jeu de données, une métadonnées, un cadre d'acquisition
- Foire aux questions
- Perspectives : création de tutoriels vidéos, organisation de webinaires sur le thème des standards d'échange

AUTRES SUJETS ABORDÉS / REMARQUES

- **point de vigilance sur la non diffusion de toute l'information (notamment des plateformes régionales vers l'INPN) :**
 - présupposition des usages de la donnée.
 - légalement si une donnée est remontée au SINP, elle doit être diffusée.
 - Le producteur pourrait reprocher la non diffusion d'informations qu'il juge importantes.
- **cardinalités multiples entre JDD et CA, entre CA et CA, entre JDD et JDD (+ données de synthèse et EBV) :** sort du périmètre de ce GT -> voir avec GT Métadonnées et nouveau GT sur le sujet en 2023.
- **Donnée d'absence :**
 - véhiculée par l'attribut 'evenementSansObservation' du concept Evenement et l'attribut statutObservation (valeur No - Non Observé) du concept SujetObservation (v4 du standard de données).
 - La prise en compte de l'information d'absence d'observation devient valorisable dans le cadre d'un protocole de collecte.
 - Propositions :
 - mieux promouvoir l'utilité de renseigner cette information d'absence dans la documentation d'accompagnement (ex: résultats non significatifs en recherche)
 - apposer une condition 'obligatoire conditionnel' ou 'recommandé' pour l'attribut 'evenementSansObservation' dans le standard de données lorsqu'un protocole de collecte est renseigné.

- **Attribut StatutSource** : comment définir une donnée qui a été numérisée depuis un article scientifique, mais :
 - qui est une donnée de terrain (article de compte-rendu d'un inventaire terrain).
 - qui est une donnée de seconde main (citation d'une collection ou d'un autre article).
 - Proposition : définir ces règles dans le standard pour qu'elles soient utilisées par tous les acteurs de la même manière.

SUITES DU GT (JUN-JUILLET-AOÛT)

Création d'un document de travail partagé reprenant les points ci-dessus et ceux qui n'ont pas pu être traités.

Création d'un document de travail partagé reprenant les nomenclatures SINP.

Ces documents vont permettre de poursuivre les discussions engagées lors de la réunion ainsi que de prioriser le travail de mise à jour des standards et de proposer l'organisation d'autres réunions en sous-groupe du cercle 1, sur des points spécifiques selon le besoin.