



Procédure de conformité et de cohérence des données et des métadonnées circulant entre les plateformes du SINP - Thématique: Occurrences de taxon, Version 1.0

Solène Robert, Rémy Jomier, Thomas Milon, Judith Panijel, Frédéric Vest, Marie Barneix, Paul Fromage

► To cite this version:

Solène Robert, Rémy Jomier, Thomas Milon, Judith Panijel, Frédéric Vest, et al.. Procédure de conformité et de cohérence des données et des métadonnées circulant entre les plateformes du SINP - Thématique: Occurrences de taxon, Version 1.0. PatriNat (AFB - CNRS - MNHN). 2017, pp.26. mnhn-04287656

HAL Id: mnhn-04287656

<https://mnhn.hal.science/mnhn-04287656>

Submitted on 15 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Procédure de conformité et de cohérence des données et des métadonnées circulant entre les plateformes du SINP

-

Thématique : Occurrences de taxon

Version 1.0

-

04/12/2017

Documents de référence

Jomier R., Chataigner J., Poncet L., Lebeau Y., Robert S., Bourgoïn T., Chagnoux S., Saltr   A., Borremans C., Archambeau A-S., Lecoq M-E, Pamerlon S., Just A., Milon T., Cousin J-L., Viel N., Barreau S. 2016. *Standard de donn  es SINP Occurrences de taxons*, version 1.2.1, MEDDE/MNHN, Paris, 93 pp.

Jomier, R., Robert S. 2017. *Extension pour occurrences de taxons v1.2.1. Extension validation v1.0*, MNHN, Paris, 21 pp.

Chataigner J., Poncet L., Lecoq M.E., Chagnoux S., Vest F. 2014. *Identifiant permanent de la DEE. D  finition op  rationnelle dans le cadre du SINP pour la th  matique occurrence de taxon*. Rapport pour le SINP, 5 pp.

Robert S., Barneix M., Body G., Castanet J., Caze G., Cellier P., Desse A., de Mazi  res J., Fromage P., Gourvil J., Jomier R., Juste A., Landry P., Lebeau Y., Lecoq M.E., Lescure J., Marage D., Meyer D., Pamerlon S., Papacotsia A., Poncet L., Quintenne G., Saltr   A. & Touroult J. 2016. *Guide m  thodologique pour la conformit  , la coh  rence et la validation scientifique des donn  es et des m  tadonn  es du SINP – Volet 1 : occurrences de taxons, Version 1*. Rapport pour le SINP, rapport MNHN-SPN 2016-77, 63 pp.

Jomier R. M  tadonn  es du SINP v1.3.8, UMS 2006 Patrimoine Naturel, Paris, 51 pp.

Citation propos  e

Robert S., Jomier R., Milon T., Panijel J., Vest F., Barneix M., Fromage P. 2017. *Proc  dure de conformit   et de coh  rence des donn  es et des m  tadonn  es circulant entre les plateformes du SINP - Th  matique : Occurrences de taxon, Version 1.0*. Rapport pour le SINP. UMS 2006 Patrinat (AFB/CNRS/MNHN), Paris, 26 pp.

Historique des versions

Version	Date	R��daction	Contributions et relectures
0.1	10/04/2017	Sol��ne Robert UMS 2006 Patrinat	UMS 2006 – Patrinat : Judith Panijel, R��my Jomier, Thomas Milon et Fr��d��ric Vest
0.2	08/08/2017	Sol��ne Robert UMS 2006 Patrinat	OAFS : Paul fromage et Marie Barneix DREAL Grand Est : St��phane Guillin PPNMEFC : Franck Grossiord CEN PACA : Julie Delauge UMS 2006 – Patrinat : Judith Panijel

Table des matières

I.	Détection des doublons.....	10
1.	Règles	10
II.	Conformité des données	11
1.	Attributs obligatoires ou obligatoires conditionnels.....	11
2.	Attributs facultatifs.....	16
3.	Règles	17
III.	Conformité des métadonnées.....	17
1.	Attributs obligatoires ou obligatoires conditionnels.....	18
a)	Niveau « cadre d'acquisition »	18
b)	Niveau « jeu de données »	19
2.	Attributs facultatifs.....	20
3.	Règles	21
IV.	Cohérence	21
1.	Intra-fichier.....	22
a)	Données.....	22
b)	Métadonnées	24
2.	Inter-fichiers	25
V.	Dépendance entre données et métadonnées.....	26
1.	Règles	26
VI.	Rapport de conformité et de cohérence	26

Préambule

Les plateformes régionales et thématique occurrences de taxon ont pour mission d'assurer la production des données élémentaires d'échange (DEE) et des métadonnées associées à partir des informations transmises par les producteurs de manière à les partager, notamment avec la plateforme nationale. Elles portent la responsabilité de la conformité et de la cohérence des DEE et des métadonnées de DEE qu'elles produisent et rendent disponibles. Les données et les métadonnées échangées entre les plateformes du SINP doivent être conformes et cohérentes.

La plateforme nationale mettra en place des contrôles de conformité et de cohérence sur les données et les métadonnées transmises par les plateformes régionales et thématique. Elle alertera les plateformes sur les éventuelles erreurs détectées par le biais d'un rapport national de conformité et de cohérence.

Ce document présente les contrôles qui seront mis en place à l'entrée de la plateforme nationale pour les occurrences de taxon. Il vise à définir le cadre minimal de conformité et de cohérence des données et métadonnées mises en circulation. Les contrôles décrits dans ce document se veulent automatisables et adaptés à la réalité du réseau. Les contrôles trop complexes à implémenter ou dont la mise en place induirait un rejet massif des données (en général en lien avec les difficultés à assurer une mise au standard) ne sont ainsi pas retenus dans cette première version. Par ailleurs, les contrôles proposés ont été allégés sur les attributs du standard pour lesquels une évolution est prévue pour une prochaine version (suppression de l'attribut, réorganisation des concepts ou levé du caractère obligatoire).

La conformité et la cohérence des DEE sont liées à la conformité et la cohérence des données transmises aux plateformes régionales et thématique. Il est donc recommandé que les plateformes produisant des DEE à partir des données reçues se basent sur les contrôles applicables aux DEE pour construire leurs procédures de contrôles respectives à l'entrée des plateformes. **Chaque plateforme dispose en outre de toute liberté pour mettre en place des contrôles complémentaires sur les données qu'elle reçoit.**

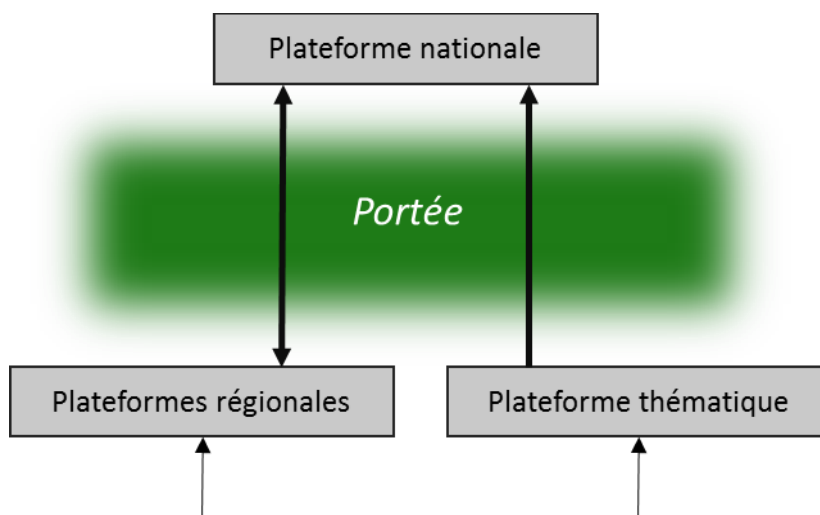
Sémantique

La conformité désigne le respect des règles fixées par les standards d'échange (renseignement des champs obligatoires, utilisation des référentiels et des listes de valeurs/nomenclatures).

La cohérence désigne le respect de la logique combinatoire des informations transmises, au sein des données ou au sein des métadonnées (on parlera de cohérence intra-fichiers) ou entre données et métadonnées (on parlera de cohérence inter-fichiers).

Portée

Ce document porte sur la conformité et la cohérence des DEE et des métadonnées (cadres d'acquisition et jeux de données) de DEE circulant entre plateformes du SINP.



Perspectives

Ce document sera susceptible d'évoluer en lien avec la mise en place progressive du SINP. Des contrôles pourront être ajoutés dans l'objectif d'améliorer progressivement la qualité des données échangées.

Les contrôles présentés étant basés sur les standards en vigueur, la procédure devra être mise à niveau dès lors qu'une version majeure d'un standard sera publiée et que les modifications concerneront des attributs soumis à un contrôle.

Schéma général du standard de données

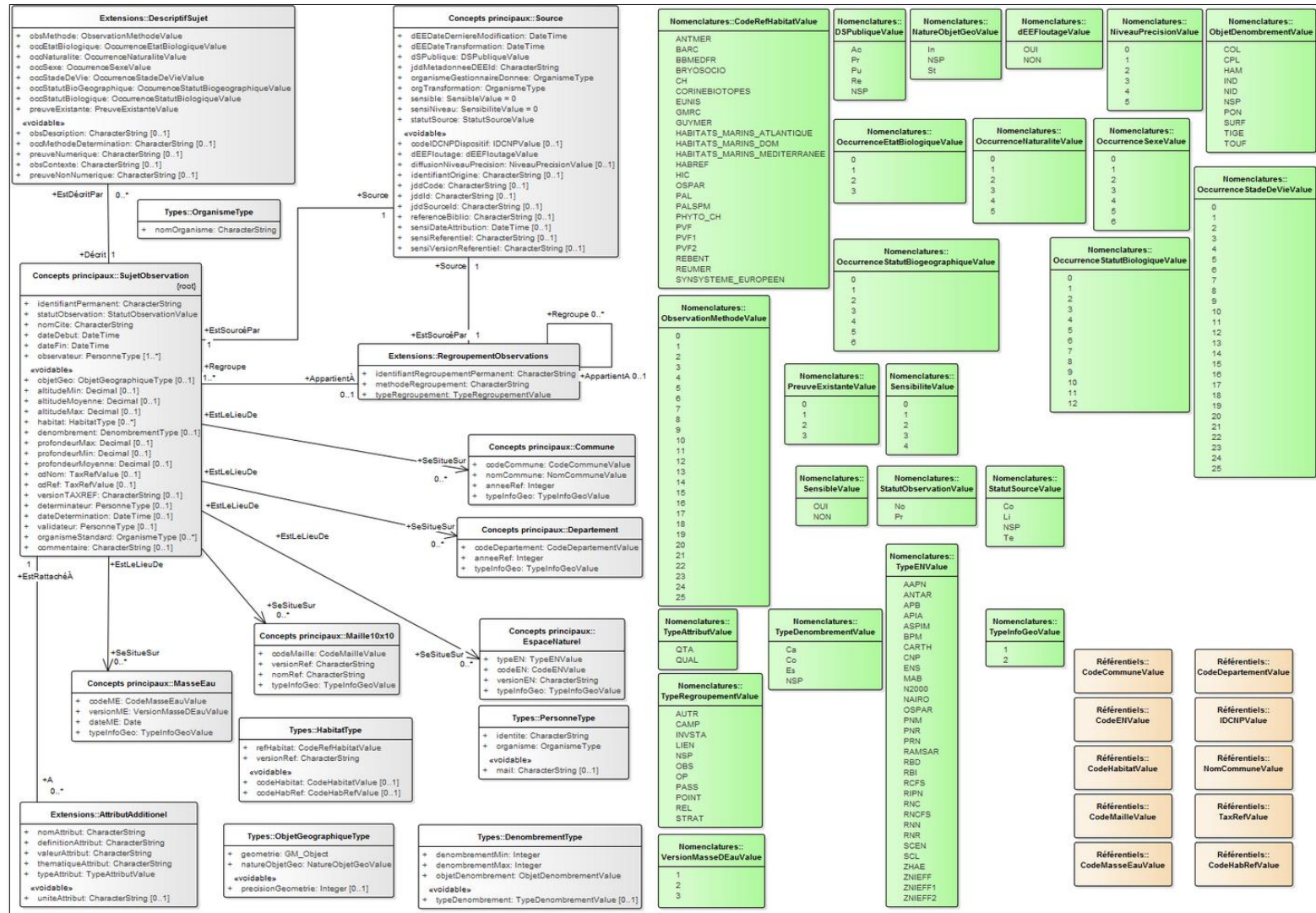


Schéma général de l'extension validation du standard de données

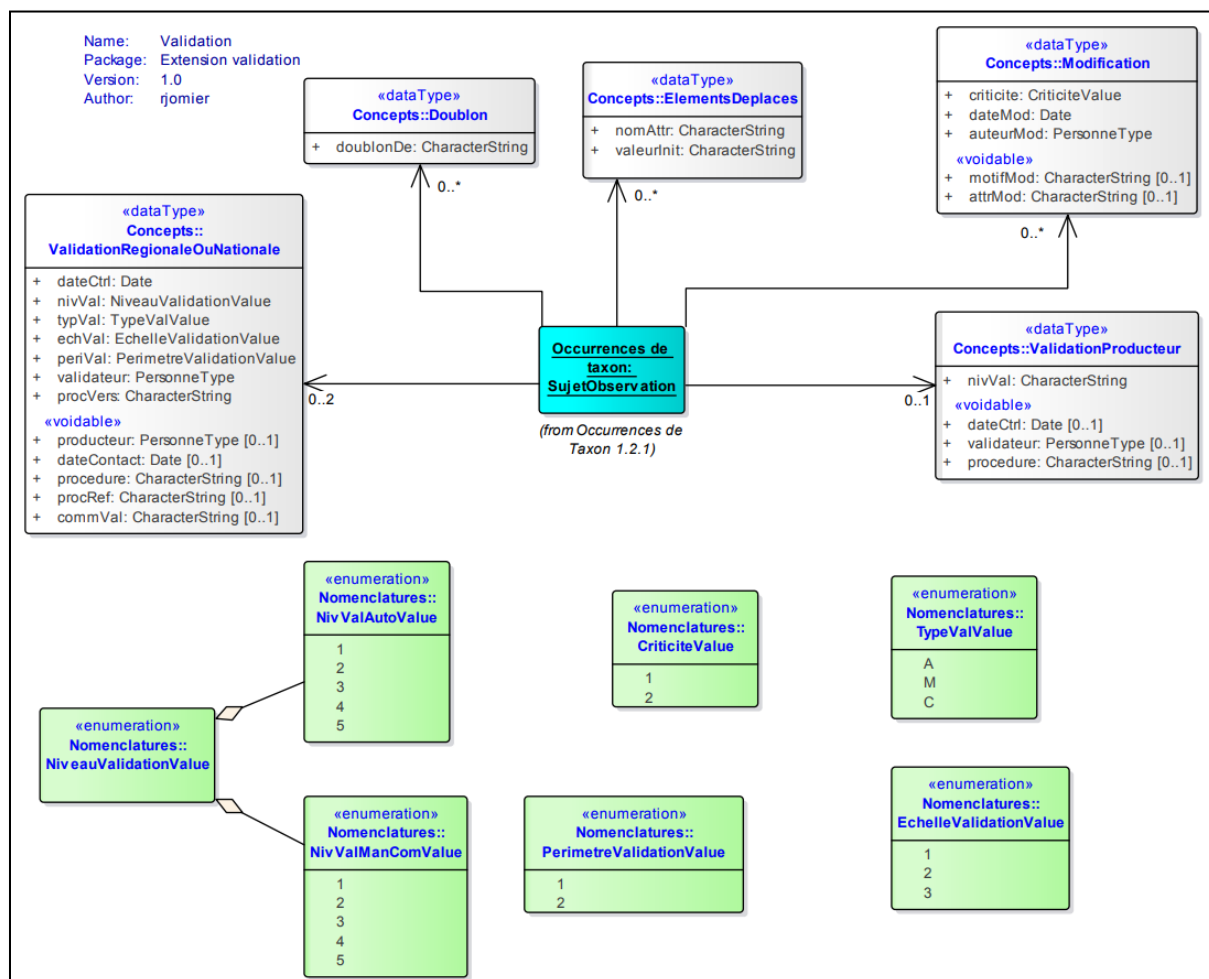


Schéma général du standard de métadonnées pour les cadres d'acquisition

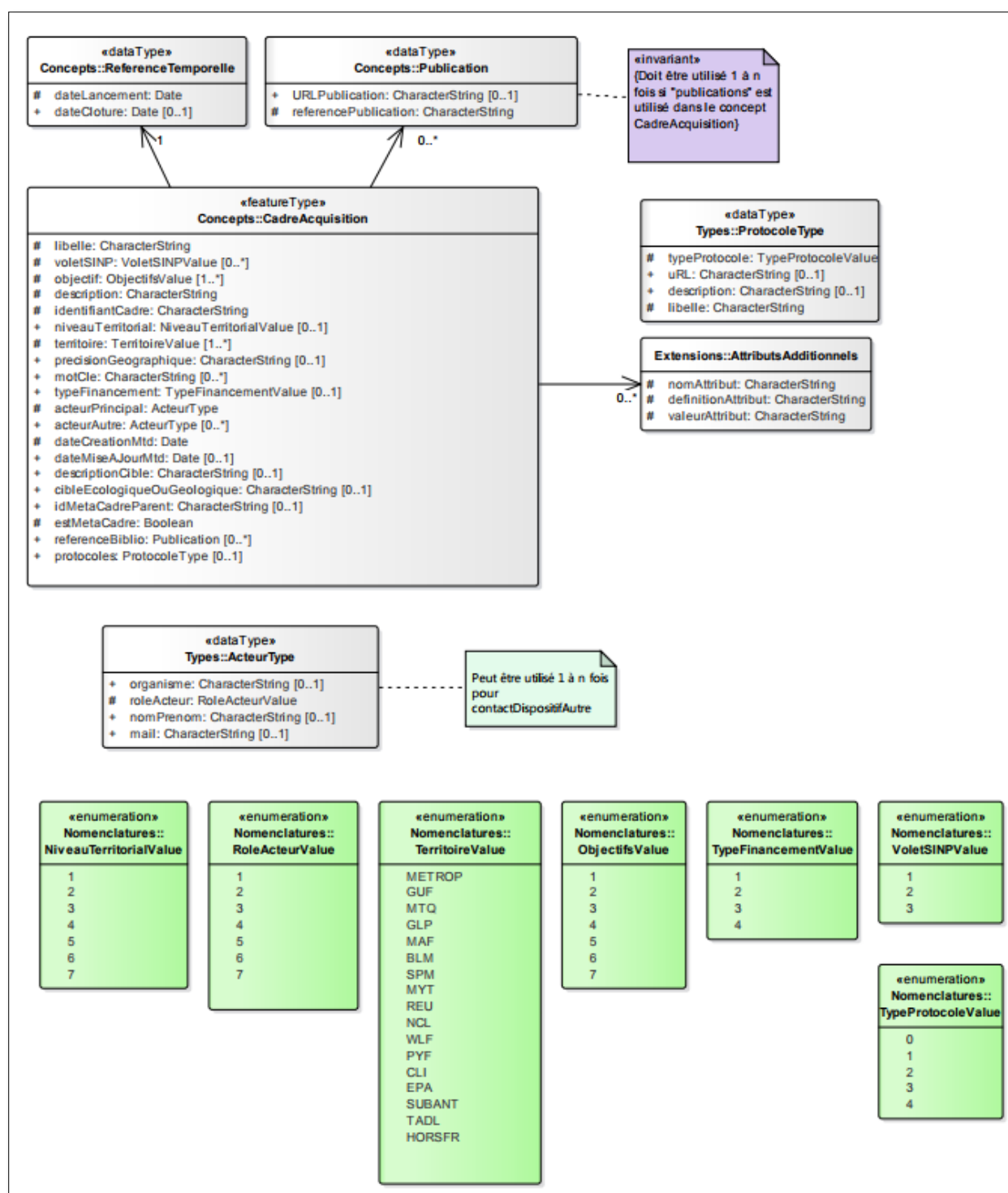
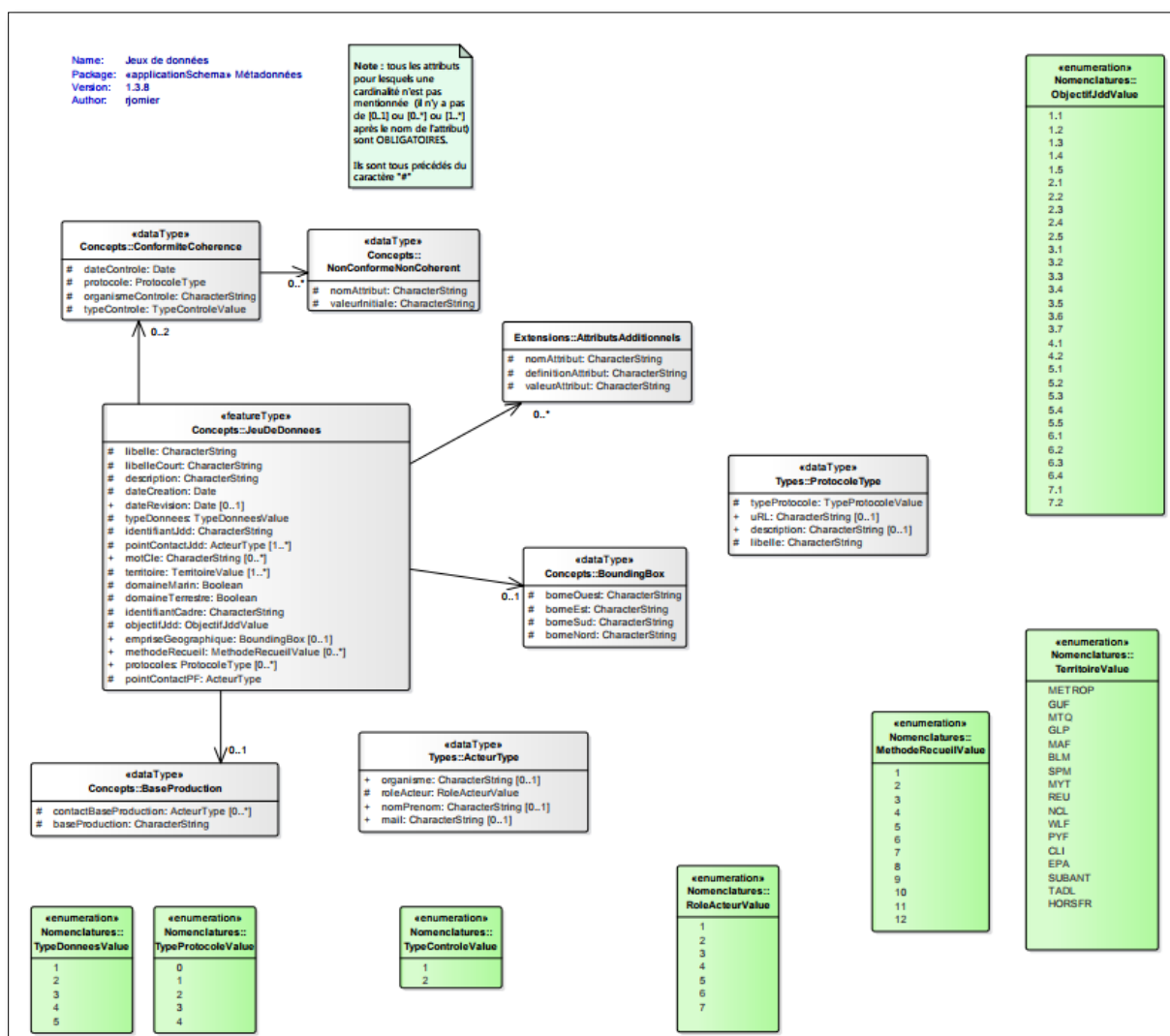
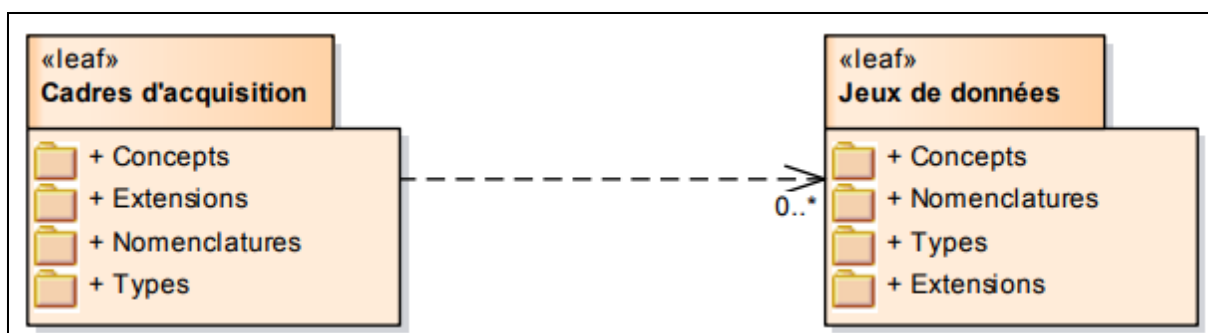


Schéma général du standard de métadonnées pour les jeux de données



Articulation des métadonnées des cadres d'acquisition / jeux de données



I. Détection des doublons

L'architecture du SINP prévoit que les données ou métadonnées ne doivent entrer dans le système que par un point unique. C'est-à-dire qu'une même donnée ou métadonnée ne doit, en principe, pas être livrée par le producteur à plusieurs plateformes d'entrée dans le système (plateformes régionales et plateforme thématique occurrences de taxon).

La plateforme nationale effectuera un contrôle de doublons. L'enjeu de ce traitement consiste en l'identification des doublons mais le présent document n'arbitre pas sur le devenir de ces doublons.

En cas d'identification de doublons potentiels, la plateforme nationale avertira la plateforme qui a fourni les éléments afin de définir l'organisation à mettre en œuvre. Les plateformes devront alors, avec les producteurs de données concernés, définir la marche à suivre et clarifier l'organisation des flux (à savoir une seule livraison au SINP : soit par la plateforme régionale, soit par la plateforme thématique occurrences de taxons).¹

A noter :

L'animation des réseaux et la connaissance partagée des flux du système global sont les premiers garants de la limitation des doublons dans le SINP.

1. Règles

Règle Doublon-1A : La donnée ou la métadonnée ne doit pas déjà être présente dans le système. L'identification du doublon se base sur l'identifiant permanent et la date de mise à jour.

Non-respect de la règle :	Conséquence
D1A	Signalement au fournisseur. Ajout d'un attribut identifiant le doublon potentiel (voir extension validation du standard occurrence de taxons)

¹ Le sujet des doublons est lié à la question des flux de données dans le SINP. Ce sujet sera traité dans le cadre du GT Architecture.

II. Conformité des données

Dans le cas où un attribut **obligatoire ou obligatoire conditionnel** du standard est contrôlé et ne répond pas aux exigences de conformité, la donnée est considérée comme « Non conforme » et la plateforme SINP habilitée qui produit les DEE doit procéder à la correction des informations.

Dans le cas où un attribut **recommandé ou facultatif** ne répond pas aux exigences de conformité, la donnée est considérée comme « Non conforme ». Pour être considérée conforme tout en gardant la complétude de l'information transmise, les attributs non conformes doivent avoir été préalablement déportés dans un attribut dédié² et avoir été mis à blanc.

NB :

Les modifications sur les DEE doivent être réalisées par la plateforme SINP habilitée qui produit la DEE et non par la plateforme qui les reçoit.

Si cela implique une modification des données sources transmises par le producteur à la plateforme, celle-ci réalisera les corrections/modifications avec l'aval du producteur des données. Les principes de bonnes pratiques voudraient que les modifications soient effectuées en tout premier lieu sur la base de données source.

1. Attributs obligatoires ou obligatoires conditionnels

Les tableaux ci-après reprennent les éléments des standards d'échange en listant les attributs obligatoires ou obligatoires conditionnels devant être renseignés dans la DEE transmise à la plateforme nationale. Ils précisent les formats attendus pour chaque attribut et indiquent si les valeurs de ces attributs sont fixées par l'usage d'un référentiel ou par un vocabulaire contrôlé (nomenclatures). Pour toute précision sur les attributs, le lecteur est invité à se référer au standard d'échanges.

Attribut : les attributs sont les propriétés, les caractéristiques des concepts.

NB : les attributs de concept de rang inférieur sont indiqués de la manière suivante dans les tableaux : « conceptSup.attributInf ».

Exemple : « identité » est un attribut du concept « PersonneType » qui doit être utilisé pour renseigner l'attribut « observateur » du concept « SujetObservation » : l'attribut est noté observateur.identité dans le tableau.

Format des attributs : Texte, Entier, Date, Décimal.

Valeur : précise si le contenu de l'attribut est régi par une nomenclature spécifique ou une norme particulière. Les valeurs sont ici marquées « Sans norme » lorsqu'aucune norme particulière ne régit l'attribut ou « Avec norme non contrôlable » lorsque la norme en question ne peut être vérifiée de manière automatique (liste de valeurs ouverte, etc.).

² Voir l'extension validation du standard de données.

Pour être conforme, le fichier standard transmis à la plateforme nationale doit contenir au minimum :

Attribut	Format	Valeur
Concept -> SujetObservation		
identifiantPermanent	Texte	Forme UUID
statutObservation	Texte	Vocabulaire contrôlé (voir nomenclature n°18 sur le site des standards du SINP)
nomCite	Texte	Sans norme
dateDebut	Date	Norme ISO8601
dateFin	Date	Norme ISO8601
observateur.identite	Texte	Avec norme non contrôlable
observateur.organisme.nomOrganisme	Texte	Sans norme
Concept -> Source		
dEEDateDerniereModification	Date	Norme ISO8601
dEEDateTransformation	Date	Norme ISO8601
jddMetadonneeDEEId	Texte	Forme UUID
organismeGestionnaireDonnee.nomOrganisme	Texte	Sans norme
orgTransformation.nomOrganisme	Texte	Sans norme
statutSource	Texte	Vocabulaire contrôlé (voir nomenclature n°19)
sensiNiveau	Entier	Vocabulaire contrôlé (voir nomenclature n°16)

Le fichier standard transmis à la plateforme nationale doit contenir au minimum l'une des informations de localisation suivante par observation/non observation :

Attribut	Format	Valeur
Concept -> SujetObservation		
objetGeo.geometrie	Texte	Avec norme non contrôlable
objetGeo.natureObjetGeo	Texte	Vocabulaire contrôlé (voir nomenclature n°3)
OU		
Concept -> Commune		
codeCommune	Texte	Vocabulaire contrôlé (INSEE)
nomCommune	Texte	Vocabulaire contrôlé (INSEE)
anneeRef	Entier	Sans norme
typeInfoGeo	Entier	Vocabulaire contrôlé (voir nomenclature n°23)
OU		
Concept -> Departement		
codeDepartement	Texte	Vocabulaire contrôlé (INSEE)
anneeRef	Entier	Sans norme
typeInfoGeo	Entier	Vocabulaire contrôlé (voir nomenclature n°23)
OU		
Concept -> Maille10x10		
codeMaille	Texte	Vocabulaire contrôlé (voir standard et site de l'INPN)

typeInfoGeo	Entier	Vocabulaire contrôlé (voir nomenclature 23 sur le site des standards du SINP)
OU		
Concept -> EspaceNaturel		
typeEN	Texte	Vocabulaire contrôlé (voir standard et site de l'INPN). Voir nomenclature 22 du site des standards du SINP.
codeEN	Texte	Vocabulaire contrôlé (voir standard et site de l'INPN)
versionEN	Texte	Sans norme
typeInfoGeo	Entier	Vocabulaire contrôlé (INSEE)
OU		
Concept -> MasseEau		
codeME	Texte	Vocabulaire contrôlé (voir standard et site du SANDRE)
versionME	Texte	Sans norme
dateME	Date	Norme ISO8601
typeInfoGeo	Entier	Vocabulaire contrôlé (voir nomenclature 23 sur le site des standards du SINP)

Pour une donnée, si sensiNiveau est renseigné avec une valeur différente de « 0 », alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> Source		
sensiDateAttribution	Date	Norme ISO8601
sensiReferentiel	Texte	Sans norme
sensiVersionReferentiel	Texte	Sans norme

Pour une donnée, si CdNom ou cdRef sont renseignés, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> SujetObservation		
determineur.identite	Texte	Avec norme non contrôlable
determineur.organisme.nomOrganisme	Texte	Sans norme

Pour une donnée, si le dénombrement est utilisé, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> DenombrementType		
objetDenombrement	Texte	Vocabulaire contrôlé (voir nomenclature n°6)
denombrementMin	Entier	Sans norme
OU		
denombrementMax	Entier	Sans norme

Pour une donnée, si le concept d'habitat est utilisé, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> HabitatType		
refHabitat	Texte	Vocabulaire contrôlé (voir nomenclature n°12)
versionRef	Texte	Sans norme

Pour une donnée, si le concept d'habitat est utilisé et que la valeur de refHabitat est HABREF, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> HabitatType		
codeHabRef	Texte	Vocabulaire contrôlé (voir référentiel HabRef sur le site de l'INPN)

Pour une donnée, si le concept d'habitat est utilisé et que la valeur de refHabitat est différente de HABREF, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> HabitatType		
codeHabitat	Texte	Sans norme

Pour une donnée, si le concept d'organismeStandard est utilisé, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> SujetObservation		
organismeStandard.nomOrganisme	Texte	Sans norme

Pour une donnée, si statutSource est renseigné avec la valeur « Li », alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> Source		
referenceBiblio	Texte	Sans norme

Si le concept d'AttributAdditionnel est utilisé pour une donnée, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> AttributAdditionnel		
nomAttribut	Texte	Sans norme
definitionAttribut	Texte	Sans norme
valeurAttribut	Texte	Sans norme
thematiqueAttribut	Texte	Sans norme
typeAttribut	Texte	Vocabulaire contrôlé (voir nomenclature n°20)

Pour une donnée, si l'attribut preuveExistante est renseigné avec la valeur « 1 » (=une preuve existe) et que preuveNonNumerique n'est pas renseignée (Null), alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> DescriptifSujet		
preuveNumerique	Texte	Sans norme

Pour une donnée, si preuveExistante est renseigné avec la valeur « 1 » (=) et que preuveNumerique n'est pas renseigné (Null), alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> DescriptifSujet		
preuveNonNumerique	Texte	Sans norme

Si le concept RegroupementObservations est utilisé pour une donnée, alors la donnée doit également contenir :

Attribut	Format	Valeur
Concept -> RegroupementObservations		
identifiantRegroupementPermanent	Texte	Avec norme non contrôlable
methodeRegroupement	Texte	Sans norme
typeRegroupement	Texte	Vocabulaire contrôlé (voir nomenclature n°24)
Concept -> Source		
dEEDateDerniereModification	Date	Norme ISO8601
identifiantOrigine	Texte	Sans norme

Si l'extension validation est utilisée, alors le fichier doit contenir au moins l'un des concepts suivants :

- ➔ ValidationRégionaleOuNationale
- ➔ ValidationProducteur
- ➔ ElementsDeplaces
- ➔ Doublon
- ➔ Modification

Si le concept ValidationRégionalOuNationale est utilisé, les attributs suivants doivent au minimum être transmis :

Attribut	Format	Valeur
Concept -> ValidationRégionaleOuNationale		
dateCtrl	Date	Norme ISO8601
nivVal	Entier	Vocabulaire contrôlé (voir nomenclature n°79 ou 80)
typVal	Texte	Vocabulaire contrôlé (voir nomenclature n°82)
echVal	Entier	Vocabulaire contrôlé (voir nomenclature n°78)
periVal	Entier	Vocabulaire contrôlé (voir nomenclature n°83)
validateur.identite	Texte	Avec norme non contrôlable
validateur.organisme.nomOrganisme	Texte	Sans norme
procVers	Texte	Sans norme

Si le concept ValidationProducteur est utilisé, les attributs suivants doivent au minimum être transmis :

Attribut	Format	Valeur
Concept -> ValidationProducteur		
nivVal	Texte	Sans norme

Si le concept ElementsDeplaces est utilisé, les attributs suivants doivent au minimum être transmis :

Attribut	Format	Valeur
Concept -> ElementsDeplaces		
nomAttr	Texte	Sans norme
valeurInit	Texte	Sans norme

Si le concept Doublon est utilisé, les attributs suivants doivent au minimum être transmis :

Attribut	Format	Valeur
Concept -> Doublon		
doublonDe	Texte	Avec norme non contrôlable (=identifiant SINP de la donnée identifiée comme doublon)

Si le concept Modification est utilisé, les attributs suivants doivent au minimum être transmis :

Attribut	Format	Valeur
Concept -> Modification		
criticite	Entier	Vocabulaire contrôlé (voir nomenclature n°81)
dateMod	Date	Norme ISO8601
auteurMod.identite	Texte	Avec norme non contrôlable
auteurMod.organisme.nomOrganisme	Texte	Sans norme

2. Attributs facultatifs

Les attributs facultatifs sont les informations pouvant être véhiculées dans la DEE mais dont la présence n'est pas obligatoire pour que la donnée soit conforme. Cependant, lorsque ces attributs sont utilisés, ils doivent alors respecter les formats et les vocabulaires contrôlés définis par le standard d'échange. Sinon, ils doivent être déportés dans un champ dédié. Pour toute précision sur ces attributs, leur format et les valeurs attendues, le lecteur est invité à se référer directement au standard d'échange.

3. Règles

Conformité des données - 1A : Les attributs obligatoires et obligatoires conditionnels cités dans les tableaux précédents doivent être présents.

Conformité des données - 1B : Les « types » ou « formats » et les vocabulaires contrôlés des attributs obligatoires et obligatoires conditionnels cités dans les tableaux précédents doivent être respectés.

Conformité des données - 2A : Les « types » ou « formats » et les vocabulaires contrôlés des attributs facultatifs doivent être respectés.

Non-respect de la règle :	Conséquence
CD1A	La donnée n'est pas conforme. Elle doit être complétée/corrigée.
CD1B	
CD2A	La donnée n'est pas conforme. Au choix de la plateforme source : - elle doit être complétée/corrigée - ou bien les attributs non conformes doivent être déportés dans l'attribut spécifiquement dédié.

III. Conformité des métadonnées

Dans le cas où un attribut **obligatoire ou obligatoire conditionnel** du standard est contrôlé et ne répond pas aux exigences de conformité, la métadonnée est considérée comme « Non conforme » et la plateforme qui produit les métadonnées de DEE doit procéder à la correction des informations.

Dans le cas où un attribut **facultatif** ne répond pas aux exigences de conformité, la métadonnée est considérée comme « Non conforme ». Pour être considérée conforme tout en gardant la complétude de l'information transmise, les attributs non conformes doivent avoir été préalablement déportés dans un attribut dédié et avoir été mis à blanc.

NB :

Les modifications sur les métadonnées de DEE doivent être réalisées par la plateforme qui produit la métadonnée et non par la plateforme qui les reçoit.

Si cela implique une modification des métadonnées sources transmises par le producteur à la plateforme, celle-ci réalisera les corrections/modifications avec l'aval du producteur des métadonnées. Les principes de bonnes pratiques voudraient que les modifications soient effectuées en tout premier lieu sur la base de données source.

1. Attributs obligatoires ou obligatoires conditionnels

Les tableaux ci-après listent les attributs obligatoires ou obligatoires conditionnels devant être renseignés dans les métadonnées de DEE échangées. Ils précisent les formats attendus pour chaque attribut et indiquent si les valeurs de ces attributs sont fixées par l'usage d'un référentiel ou par un vocabulaire contrôlé (nomenclatures). Pour toute précision sur les attributs, le lecteur est invité à se référer au standard d'échange.

Attribut : les attributs sont les propriétés, les caractéristiques des concepts.

Nb : les attributs de concept de rang inférieur sont indiqués de la manière suivante dans les tableaux : « concept.attribut »

Format des attributs : Texte, Entier, Date, Décimal.

Valeur : précise si le contenu de l'attribut est régi par une nomenclature spécifique ou une norme particulière. Les valeurs sont ici marquées « Sans norme » lorsqu'aucune norme particulière ne régit l'attribut ou « Avec norme non contrôlable » lorsque la norme en question ne peut être vérifiée de manière automatique (liste de valeurs ouverte, etc.).

a) Niveau « cadre d'acquisition »

Les fichiers standards de métadonnées doivent contenir au minimum, pour les cadres d'acquisition :

Attribut	Format	Valeur
Concept -> CadreAcquisition		
libelle	Texte	Sans norme
voletSINP	Entier	Vocabulaire contrôlé (voir nomenclature n°90)
objectif	Entier	Vocabulaire contrôlé (voir nomenclature n°85)
identifiantCadre	Texte	Forme UUID
territoire	Texte	Vocabulaire contrôlé (voir nomenclature n°87)
dateCreationMtd	Date	Norme ISO8601
estMetaCadre	Booléen	true/false
acteurPrincipal.roleActeur	Entier	Doit être égal à 1 (=contact principal)
acteurPrincipal.organisme	Texte	Sans norme
OU		
acteurPrincipal.nomPrenom	Texte	Avec norme non contrôlable

Si acteurAutre est utilisé, alors, le fichier doit contenir :

Concept -> CadreAcquisition		
acteurAutre.roleActeur	Entier	Vocabulaire contrôlé (voir nomenclature n°86)
acteurAutre.organisme	Texte	Sans norme
OU		
acteurAutre.nomPrenom	Texte	Avec norme non contrôlable

Pour un cadre d'acquisition, si le concept « Publication » est utilisé, alors le cadre d'acquisition doit également contenir :

Attribut	Format	Valeur
Concept -> Publication		
referencePublication	Texte	Avec norme non contrôlable / ISO 690

Si l'attribut « protocoles » est utilisé, alors le fichier doit contenir :

Attribut	Format	Valeur
Concept -> CadreAcquisition		
Protocoles.typeProtocole	Texte	Vocabulaire contrôlé (voir nomenclature n°89)
Protocoles.libelle	Texte	Sans norme

Pour un cadre d'acquisition, si l'extension AttributsAdditionnels est utilisée, alors doivent être transmis :

Attribut	Format	Valeur
nomAttribut	Texte	Sans norme
definitionAttribut	Texte	Sans norme
valeurAttribut	Texte	Sans norme

b) Niveau « jeu de données »

Au minimum, les fichiers transmis à la plateforme nationale doivent contenir, pour les jeux de données :

Attribut	Format	Valeur
Concept -> JeuDedonnees		
libelle	Texte	Sans norme
libelleCourt	Texte	Sans norme
description	Texte	Sans norme
dateCreation	Date	Norme ISO8601
typeDonnees	Entier	Vocabulaire contrôlé (voir nomenclature n°94)
identifiantJDD	Texte	Forme UUID
territoire	Texte	Vocabulaire contrôlé (voir nomenclature n°87)
domaineMarin	Booléen	true/false
domaineTerrestre	Booléen	true/flase
identifiantCadre	Texte	Avec norme non contrôlable
objectifJdd	Texte	Vocabulaire contrôlé (voir nomenclature n°92)

pointContactJDD.roleActeur	Entier	Au moins un des rôles pour le jeu transmis doit être égal à 1
pointContactJDD.nomPrenom	Texte	Avec norme non contrôlable
OU		
pointContactJDD.organisme	Texte	Sans norme
pointContactPF.roleActeur	Entier	Doit être égal à 1 (=contact principal)
pointContactPF.nomPrenom	Texte	Avec norme non contrôlable
OU		
pointContactPF.organisme	Texte	Sans norme

Pour un jeu de données, si le concept BaseProduction est utilisé, alors le jeu doit également contenir :

Attribut	Format	Valeur
Concept -> BaseProduction		
baseProduction	Texte	Sans norme
ContactBaseProduction.roleActeur	Entier	Doit être = 7 (Point de contact sur la base de données de production)
ContactBaseProduction.nomPrenom	Texte	Avec norme non contrôlable
OU		
ContactBaseProduction.organisme	Texte	Sans norme

Pour un jeu de données, si l'attribut protocoles est utilisé, alors le jeu doit également contenir :

Attribut	Format	Valeur
Concept -> ProtocoleType		
typeProtocole	Entier	Vocabulaire contrôlé (voir standard de métadonnées) – ne peut cependant pas être = à 3 (conformité et cohérence porté par un concept dédié)
libelle	Texte	Sans norme

Pour un jeu de données, si le concept ConformiteCoherence est utilisé, alors le jeu doit aussi contenir :

Attribut	Format	Valeur
Concept -> ConformiteCoherence		
dateControle	Date	Norme ISO8601
organismeControle	Texte	Sans
typeControle	Entier	Vocabulaire contrôlé (voir standard de métadonnées)
Protocole		
typeProtocole	Entier	Doit être égal à 3
libelle	Texte	Sans

Pour un jeu de données, si l'extension AttributsAdditionnels est utilisée, alors doivent être transmis :

Attribut	Format	Valeur
nomAttribut	Texte	Sans
definitionAttribut	Texte	Sans
valeurAttribut	Texte	Sans

2. Attributs facultatifs

Les attributs facultatifs sont les informations pouvant être véhiculées dans la métadonnée de DEE mais dont la présence n'est pas obligatoire pour que la métadonnée soit conforme. Cependant, lorsque ces attributs sont utilisés, ils doivent alors respecter les formats et les vocabulaires contrôlés définis par le standard d'échange. Pour toute précision sur ces attributs, leur format et les valeurs attendues, le lecteur est invité à se référer au standard d'échange.

3. Règles

Conformité des métadonnées - 1A : Les attributs obligatoires et obligatoires conditionnels cités dans les tableaux précédents doivent être présents.

Conformité des métadonnées - 1B : Les « types » et les vocabulaires contrôlés des attributs obligatoires et obligatoires conditionnels cités dans les tableaux précédents doivent être respectés.

Conformité des métadonnées - 2A : Les « types » et les vocabulaires contrôlés des attributs facultatifs doivent être respectés.

Non-respect de la règle :	Conséquence
CM1A	La métadonnée n'est pas conforme. Elle doit être complétée/corrigée.
CM1B	
CM2A	La métadonnée n'est pas conforme. Elle doit être complétée/corrigée ou bien les attributs non conformes doivent être déportés dans l'attribut spécifiquement dédié.

IV. Cohérence

Dans le cas où un attribut **obligatoire ou obligatoire conditionnel** du standard est contrôlé et ne répond pas aux exigences de cohérence, la donnée ou la métadonnée est considérée comme « Non conforme » et la plateforme qui produit les données et les métadonnées de DEE doit procéder à la correction des informations.

Dans le cas où un attribut **facultatif** ne répond pas aux exigences de conformité, la donnée ou métadonnée est considérée comme « Non conforme ». Pour être considérée conforme tout en gardant la complétude de l'information transmise, les attributs non conformes doivent avoir été préalablement déportés dans un attribut dédié et avoir été mis à blanc.

NB :

Les modifications sur les données et métadonnées de DEE doivent être réalisées par la plateforme qui produit la DEE ou la métadonnée et non par la plateforme qui les reçoit.

Si cela implique une modification des données ou des métadonnées sources transmises par le producteur à la plateforme, celle-ci réalisera les corrections/modifications avec l'aval du producteur.

Si cela implique une déportation des informations, considérant que la plateforme ne peut statuer seule sur l'information à retenir (quel attribut est celui qui fait référence), les deux attributs incohérents sont mis à blanc et déportés dans l'attribut spécifiquement dédié à l'opération.

1. Intra-fichier

a) Données

Cohérence des attributs temporels

Cohérence des données - 1A : Les dates de début et de fin d'observation ou de « non observation » doivent être cohérentes entre elles : dateDebut doit être < ou = à dateFin.

Cohérence des données - 1B : Les dates de début et de fin d'observation ou de « non observation » doivent être cohérentes avec la date du traitement de la donnée. dateDebut et dateFin doivent être < ou = à la date de l'intégration (sysdate).

Non-respect de la règle :	Conséquence
CHD1A	La donnée n'est pas cohérente. Elle doit être complétée/corrigée.
CHD1B	

Cohérence des attributs géographiques

Cohérence des données - 2A : Au moins une donnée du jeu de données transmis doit être située sur le territoire français (Métropole et Outre-Mer, terrestre ou marin).

Cohérence des données - 2B : Si un fichier « EN » est transmis, le code d'espace naturel doit être retrouvé dans la liste des valeurs du référentiel spécifié (selon le type d'espace naturel).

Cohérence des données - 2C : Si un objet géographique d'acquisition (X/Y ou objetGeo) est transmis, alors il doit être cohérent avec la ou les communes associées sur la base du code INSEE (si commune(s) renseignée(s)). Test de contiguïté. Tolérance de 500 m.

Cohérence des données - 2D : L'objet géographique d'acquisition doit être cohérent avec le ou les espace(s) naturel(s) mentionnés (si espace(s) renseigné(s)). Test de contiguïté. Tolérance de 500m.

Cohérence des données - 2E : L'objet géographique d'acquisition doit être cohérent avec la ou les maille(s) mentionnée(s) (si maille(s) renseigné(s)). Test de contiguïté. Tolérance de 500m.

Cohérence des données - 2F : L'objet géographique d'acquisition doit être cohérent avec la ou les masse(s) d'eau mentionnée(s) (si masse(s) d'eau renseignée(s)). Test de contiguïté. Tolérance de 500m.

Cohérence des données - 2G : Si les fichiers « Commune » et « EN » sont transmis, alors les deux informations doivent être cohérentes (test de contiguïté des objets – basé sur le code INSEE des communes transmis). Tolérance de 500m.

Cohérence des données - 2H : Si les fichiers « Commune » et « maille » sont transmis, alors les deux informations doivent être cohérentes (test de contiguïté des objets – basé sur le code INSEE des communes transmis). Tolérance de 500m.

Cohérence des données - 2I : Si les fichiers « Commune » et « ME » sont transmis, alors les deux informations doivent être cohérentes (test de contiguïté des objets – basé sur le code INSEE des communes transmis). Tolérance de 500m.

Cohérence des données - 2J : Si les fichiers « EN » et « Maille » sont transmis, alors les deux informations doivent être cohérentes (test de contiguïté des objets). Tolérance de 500m.

Cohérence des données - 2K : Si les fichiers « EN » et « ME » sont transmis, alors les deux informations doivent être cohérentes (test de contiguïté des objets). Tolérance de 500m.

Cohérence des données - 2L : Si les fichiers « Maille » et « ME » sont transmis, alors les deux informations doivent être cohérentes (test de contiguïté des objets). Tolérance de 500m.

***NB :** Une tolérance dans l'écart entre les objets est admise pour prendre en compte les éventuelles imprécisions de géoréférencement (pointage sur carte, imprécision du GPS, etc.) ainsi que les éventuelles écarts en lien avec des reprojections. Ces contrôles prennent en compte la précision géographique de l'objet indiqué par le producteur lorsqu'elle est transmise.*

Cohérence 2M : Pour une même donnée, il ne doit y avoir qu'un seul objet (géographique, administratif, maille, etc.) à TypeInfoGeo=1 (=géoréférencement). Si des données (exemple dans le cas de la bibliographie) présentent plusieurs objets géographiques à TypeInfoGeo=1 alors ces objets devront être fusionnés préalablement en un seul polygone.

***A noter :** une donnée peut ne pas être liée à un objet de TypeInfoGeo=1 lorsqu'il s'agit d'une donnée privée ayant été floutée.*

Non-respect de la règle :	Conséquence
CHD2A	Les données ne sont pas cohérentes avec le périmètre du SINP. Elles doivent être complétées/corrigée (potentielle erreur de décalage de coordonnées ou inversion). Si toutes les données sont réellement situées sur un territoire hors France, alors le fournisseur peut se tourner vers d'autres systèmes de diffusion (GBIF notamment).
CHD2B	La donnée n'est pas cohérente. Elle doit être complétée/corrigée.
CHD2C	
CHD2D	
CHD2E	
CHD2F	
CHD2G	
CHD2H	
CHD2I	
CHD2J	
CHD2K	
CHD2L	
CHD2M	La donnée n'est pas cohérente. Elle doit être complétée/corrigée ou la plateforme doit mettre en place une fusion des objets.

Cohérence des attributs taxonomiques

Néant. Les noms cités ne sont pas normalisés et un contrôle automatique avec le cd_nom ou cd_ref est donc impossible. Par ailleurs, un nom cité peut-être incohérent avec le cd_nom dès lors qu'il y a eu une détermination ultérieure sans qu'il s'agisse d'une erreur.

Cohérence des autres types d'attributs

Cohérence des données - 3A : Si les dénombrements sont renseignés, alors le dénombrement minimal doit être inférieur ou égal au dénombrement maximal. $\text{denombrementMin} < \text{ou} = \text{denombrementMax}$.

Non-respect de la règle :	Conséquence
CHD3A	La donnée n'est pas cohérente. Elle doit être complétée/corrigée ou bien les attributs incohérents doivent être déportés dans l'attribut spécifiquement dédié.

b) Métadonnées

Cohérence des attributs temporels

Cadre d'acquisition :

Cohérence 1A : La date de lancement du cadre d'acquisition doit être inférieure ou égale à la date de clôture du cadre d'acquisition (si cette dernière est renseignée).

Non-respect de la règle :	Conséquence
1A	La métadonnée n'est pas cohérente. Elle doit être complétée/corrigée.

Jeu de données :

Néant. Les bornes temporelles du jeu de données sont automatiquement calculées à partir des dates d'observation des données conformes et cohérentes.

Cohérence des attributs géographiques

Cadre d'acquisition :

Néant. Pas de double information géographique normalisée et obligatoire dans la fiche de cadre d'acquisition.

Jeu de données :

Néant. Pas de double information géographique normalisée et obligatoire dans la fiche de jeu de données.

Cohérence des aspects taxonomiques

Cadre d'acquisition :

Néant. Pas d'information normalisée pour le cadre d'acquisition sur la taxonomie.

Jeu de données :

Néant. Pas d'information normalisée pour le jeu de données sur la taxonomie.

2. Inter-fichiers

Cohérence des attributs contextuels

Entre cadre d'acquisition et jeu de données :

Les objectifs du cadre d'acquisition et des jeux de données doivent être cohérents (imbrication des valeurs).

Cohérence des attributs temporels

Entre cadre d'acquisition et jeu de données :

Néant. Un programme peut mobiliser des données antérieures à sa date de lancement (donc des jeux de données avec une emprise plus large que celle du cadre).

Entre cadre d'acquisition et données :

Néant. Un programme peut mobiliser des données antérieures à sa date de lancement.

Entre jeu de données et données :

Néant, les dates d'emprise temporelle du jeu de données sont calculées à partir des données.

Cohérence des attributs géographiques

Entre cadre d'acquisition et jeu de données :

Cohérence 2A : Les territoires déclarés pour les cadres et les jeux associés doivent être cohérents.

Non-respect de la règle :	Conséquence
2A	Les métadonnées ne sont pas cohérentes. Elles doivent être complétées/corrigées.

Entre cadre d'acquisition et données :

Cohérence 3A : Les données doivent être cohérentes avec les territoires déclarés au niveau du cadre d'acquisition (incluses).

Non-respect de la règle :	Conséquence
3A	La métadonnée n'est pas cohérente avec les données. Métadonnées ou données doivent être complétée/corrigée.

Entre jeu de données et données :

Néant, la bounding box (rectangle d'encombrement) du jeu de données est calculée à partir des données.

Sur les attributs taxonomiques :

Entre cadre d'acquisition et jeu de données :

Néant. Pas d'attributs obligatoires dédiés normalisés permettant un contrôle automatique.

Entre cadre d'acquisition et données :

Néant. Pas d'attributs obligatoires dédiés normalisés permettant un contrôle automatique.

Entre jeu de données et données :

Néant. Pas d'attributs obligatoires dédiés normalisés permettant un contrôle automatique.

V. Dépendance entre données et métadonnées

Pour circuler au sein du SINP, les DEE doivent être associées à une fiche de jeu de données qui doit être associée à une fiche de cadre d'acquisition.

De fait, lorsqu'une fiche de cadre d'acquisition n'est pas cohérente/conforme, la fiche de ce cadre et les fiches de jeux de données ainsi que les données qui y sont rattachés ne peuvent ni circuler, ni être diffusées.

De la même manière, lorsqu'une fiche de jeu de données est incohérente/non-conforme, la fiche du jeu de données et les données qui y sont attachées ne peuvent ni circuler, ni être diffusées.

1. Règles

Dépendance 1A : les données transmises doivent se rattacher à un jeu de données.

Dépendance 1B : les jeux transmis doivent se rattacher à un cadre d'acquisition.

Dépendance 1C : les données transmises doivent se rattacher à un jeu de données conforme et cohérent, lui-même lié à un cadre d'acquisition conforme et cohérent.

Non-respect de la règle :	Conséquence
DEP1A	Une fiche de jeu de données associée aux données doit être transmise.
DEP1B	Une fiche de cadre d'acquisition associée aux jeux de données doit être transmise.
DEP1C	Les informations incohérentes ou non conformes doivent être complétées/corrigées.

VI. Rapport de conformité et de cohérence

Suite à la réception de DEE et de métadonnées de DEE issues des plateformes régionales et thématiques, la plateforme nationale réalisera les contrôles de conformité et de cohérence et transmettra un rapport national de conformité et de cohérence aux plateformes sources.

Ce rapport comportera :

- ➔ La mention des attributs non conformes ou incohérents et la raison de la non-conformité ou de l'incohérence.
- ➔ La date de réalisation des contrôles
- ➔ La référence de la procédure de contrôle appliquée
- ➔ Les résultats des contrôles d'identification de doublons potentiels