

Architecture de table de données et métadonnées associées



Une mise en forme du modèle d'architecture

1. Le contexte
2. Le projet
3. Référentiels communs

Éléments contextuels

4. Un modèle pour une architecture de table de données naturalistes
5. Un modèle pour une architecture des métadonnées associées à un lot de données naturalistes

Éléments descriptifs

6. Fiabilité des données
7. Sensibilité des données

Précisions complémentaires



PDF

tableur

	nom du champ	format	contenu	commentaire
obligatoire	<u>CodeNom</u>	entier	code du taxon* selon le référentiel utilisé par le producteur de la donnée	*une donnée de type nom de genre (ex. « puffin sp. ») est gérée par tout « bon » référentiel
obligatoire	<u>NomScientifique</u>	texte	nom scientifique du taxon selon le référentiel utilisé par le producteur de la donnée	un nom scientifique sans ambiguïté cite le(s) nom(s) d'auteur(s)
optionnel*	<u>NomFrançais</u>	texte	nom français du taxon	*le nom français n'existe pas toujours ; d'où le caractère optionnel, mais fortement recommandé pour la lisibilité de la table par les non spécialistes
obligatoire	<u>ReferentielNom</u>	texte	citation du référentiel <u>nomenclatural</u> utilisé par le producteur de la donnée	l'utilisation d'un référentiel est très fortement recommandée et si possible un référentiel déjà existant
obligatoire	<u>CodeNomTaxRef</u>	texte	code du taxon selon le référentiel national TaxRef du MNHN en utilisant le champ CD_NOM de TaxRef	si le producteur utilise TaxRef pour son référencement, alors CodeNom = CodeNomTaxRef ; ce champ permet d'agréger des tables qui utiliseraient des ReferentielNom différents ;
obligatoire	<u>Presence</u>	texte	seules 2 valeurs possibles : oui / non	Valeur « non » = non observé ; cf. le cas échéant <u>DenombComplement</u> pour des précisions
optionnel	<u>Denombrement</u>	texte*	la quantité dénombrée	*valeurs possibles : valeur entière, valeur décimale, fourchette de valeur... d'où le format texte
optionnel	<u>DenombComplement</u>	texte	toutes spécifications nécessaires à la compréhension de <u>Denombrement</u>	valeurs possibles : grandeur mesurée (la métrique), ordre de grandeur, niveau de précision, niveau d'estimation...

Architecture de table de données et métadonnées associées

accessibles sur GéoBretagne

Première approche

← réunion 17 décembre 2015

→ support réunion

Proposition

← consultation courriel mars 2016

→ questionnaire

Consolidation

← réunion 21 avril 2016

→ relevé de décisions

Proposition

← consultation courriel août-septembre 2016

→ mise en forme v1

Validation

← réunion du 29 septembre 2016

→ mise en forme finale
post réunion du 29/09« **Fiabilité** » & « **Sensibilité** »

← à venir pour fin 2016

doctrine régionale à finaliser

définition des critères d'applicabilité

TABLE DE DONNEES

	nom du champ	format	contenu	commentaire
obligatoire	CodeNom	entier	code du taxon* selon le référentiel utilisé par le producteur de la donnée	*une donnée du type nom de genre (ex. « puffin sp. ») est gérée par tout « bon » référentiel
obligatoire	NomScientifique	texte	nom scientifique du taxon selon le référentiel utilisé par le producteur de la donnée	un nom scientifique sans ambiguïté cite le(s) nom(s) d'auteur(s)
optionnel*	NomFrancais	texte	nom français du taxon	*le nom français n'existe pas toujours ; d'où le caractère optionnel, mais fortement recommandé pour la lisibilité de la table par les non spécialistes
obligatoire	ReferentielNom	texte	citation du référentiel nomenclatural utilisé par le producteur de la donnée	l'utilisation d'un référentiel est très fortement recommandé et si possible un référentiel déjà existant
obligatoire	CodeNomTaxRef	texte	code du taxon selon le référentiel national TaxRef du MNHN en utilisant le champ CD_NOM de TaxRef	si le producteur utilise TaxRef pour son référencement, alors CodeNom = CodeNomTaxRef ; ce champ permet d'agrèger des tables qui utiliseraient des ReferentielNom différents ;
obligatoire	Presence	texte	seules 2 valeurs possibles : oui / non	Valeur « non » = non observé ; cf. le cas échéant DenombComplement pour des précisions
optionnel	Denombrement	texte*	la quantité dénombrée	*valeurs possibles : valeur entière, valeur décimale, fourchette de valeur... d'où le format texte
optionnel	DenombComplement	texte	toutes spécifications nécessaires à la compréhension de Denombrement	valeurs possibles : grandeur mesurée (la métrique), ordre de grandeur, niveau de précision, niveau d'estimation...
obligatoire	DateDebut	entier	valeur ISO8601 de la date d'observation soit annéemoisjour	ex. : 20160530 pour 30 mai 2016
obligatoire*	DateFin	entier	valeur ISO8601 de la date d'observation soit annéemoisjour	*si la donnée concerne une date unique DateFin = DateDebut

Pour rappel, l'essentiel des modifications apportées en réunion du 21 avril :

- consolidation de l'aspect nomenclatural ;
- simplification de l'aspect métrique.

TABLE DE DONNEES

	nom du champ	format	contenu	commentaire
obligatoire	EntiteGeographique	texte*	code spécifique à l'entité selon le référentiel utilisé ; si aucun référentiel n'est utilisé, c'est un numéro classant	si le fichier est au format SIG, ce champ est aussi dans la table attributaire ; *texte : le codage peut être de nature textuelle d'où le format générique texte
obligatoire optionnel* conditionné	TypeGeographique	texte	type d'entité codée dans le référentiel : maille, commune... ou secteurproducteur si le type n'est pas référencé	optionnel* conditionné: si ils ne sont pas cités dans la table, ils sont cités dans la métadonnée
obligatoire optionnel* conditionné	ReferentielGeographique	texte	citation du référentiel géographique utilisé ; valeur « néant » si aucun référentiel n'est utilisé	le référentiel peut être institutionnel (commune, cours d'eau...) ou propre au producteur ; optionnel* conditionné: si ils ne sont pas cités dans la table, ils sont cités dans la métadonnée
obligatoire*	X	décimal	coordonnées métriques en Lambert93 du point	*obligatoire si la géométrie est ponctuelle et que le fichier n'est pas au format SIG
obligatoire*	Y			
obligatoire	Sensibilite	texte	seules 2 valeurs possibles : oui / non	se référer au référentiel régional « Sensibilité » (en cours de réflexion) pour catégoriser la donnée
optionnel	Fiabilite	entier	seules 3 valeurs possibles : 1 / 2 / 3	hiérarchie : 3 > 2 > 1 ; se référer au référentiel régional « Fiabilité » (en cours de réflexion) pour catégoriser la donnée
optionnel	TypeObservation	texte	seules 3 valeurs possibles : terrain / littérature / collection	
obligatoire	Observateur	texte	la personne (ou l'organisme) à créditer de l'observation ; peut être complété avec l'auteur de l'identification du specimen	valeurs possibles : une personne, un organisme...anonyme, inconnu...
optionnel* conditionné	Producteur	texte	organisme ayant collecté l'observation auprès de l'observateur	ces trois champs permettent de gérer les multiples cas ayant conduit à l'élaboration de la table ; la notion de maître d'ouvrage permet de gérer une compilation de données issues de divers producteurs
optionnel* conditionné	Maitredouvrage	texte	organisme ayant compilé l'ensemble des données de la présente table	
optionnel* conditionné	Commanditaire	texte	organisme ayant commandité l'élaboration de la table	

Pour rappel, l'essentiel des modifications apportées en réunion du 21 avril :

- Précision concernant les opérateurs.

METADONNEES

	nom du champ	format	contenu	commentaire
obligatoire	TitreLotDonnee	texte	nom donnée au lot de donnée	
obligatoire	DescriptionLotDonnee	texte	description sémantique du contenu du lot de donnée	
obligatoire	IdentifiantLotDonnee	texte	code identifiant de manière unique* le lot de donnée	*une procédure possible pour assurer l'unicité : FR + n° Siren + nom du fichier (cf. wiki GéoBretagne)
obligatoire	ThemeISO	texte	autant de valeurs que nécessaire parmi le lot de mots clés fermés dans le référentiel ThemeISO de GéoBretagne	
obligatoire	ThemeInspire	texte	autant de valeurs que nécessaire parmi le lot de mots clés fermés dans le référentiel ThemeINSPIRE de GéoBretagne	
optionnel	MotClef	texte	valeur(s) au choix du producteur	
optionnel	ExtensionGeographique	texte	liste des limites administratives correspondant à l'extension géographique des données, ainsi que le référentiel administratif utilisé (ex. communes Geofla2015)	ne s'utilise que si l'extension correspond à une limite administrative : « Bretagne », « Départements 22 et 29 »...; seules les limites de la plus grande dimension sont citées (ex. les départements mais pas les communes contenues)
obligatoire	LatitudeN	décimal	les coordonnées métriques x,y de chacun des 4 coins du rectangle de l'emprise maximale du lot de données, en Lambert93	l'automatisation du calcul est possible quand la fiche de métadonnée est directement remplie dans GéoNetWork
obligatoire	LatitudeS			
obligatoire	LongitudeE			
obligatoire	LongitudeO			
obligatoire	DateCreation	entier	valeur ISO8601 année-mois-jour de la date de création du lot de données	ex. : 20160530
obligatoire	DatePublication	entier	valeur ISO8601 année-mois-jour de la date de publication du lot de données	
obligatoire	DateRevision	entier	valeur ISO8601 année-mois-jour de la date de révision du lot de données	révision : correction apportée a posteriori, ajout de champs complémentaire, ajout de lignes de données, etc.
obligatoire	DateDebut	entier	valeur ISO8601 année-mois-jour de la première date de données	
obligatoire	DateFin	entier	valeur ISO8601 année-mois-jour de la dernière date de données	

METADONNEES

	nom du champ	format	contenu	commentaire
obligatoire	Methode	texte	concerne les aspects méthodologiques d'acquisition sur le terrain ainsi que de traitement et d'analyse de la donnée	il est possible de citer un lien vers une documentation externe
optionnel	LimiteUtilisation	texte	citer les limites méthodologiques éventuelles liées à l'utilisation du lot de données	ne concerne que les aspects méthodologiques : « données non pertinentes dans telles conditions », etc.
obligatoire	EchelleUtilisation	texte	citer la gamme d'échelle pour laquelle le lot de donnée reste pertinent	
optionnel	ContrainteUtilisation	texte	citer les contraintes éventuelles (autres que méthodologiques) liées à l'utilisation du lot de donnée	ex. : « usage libre sous réserve des mentions obligatoires sur tout document de diffusion... », « ne pas diffuser ce lot de donnée en l'état car contient des données sensibles précises », etc.
optionnel	AccesDonnes	texte	lien pour accéder à la donnée	cas où la donnée peut être téléchargée via un site, une plateforme
obligatoire	Contact	texte	personne à contacter pour tout renseignement sur le lot de donnée ou sur son accès	l'information doit permettre d'accéder le plus directement possible à la personne ressource
obligatoire	Responsable	texte	organisation(s) ou personne(s) responsable(s) de la création, gestion, maintenance et diffusion du lot de données	ce champ permet d'exprimer aussi selon les cas : le/les producteurs et/ou le maître d'ouvrage et/ou le commanditaire