



SINP

Système d'Information
sur la Nature et le Paysage



SINP de l'Occitanie

1^{er} CoTech

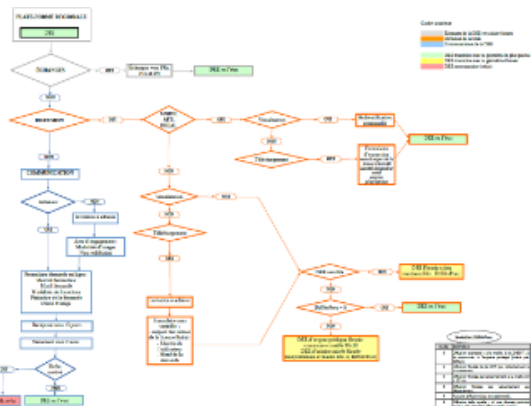
DREAL - 19 mars 2019



- 
- 2. Mise à disposition de la DEE**
 - 3. Plateforme régionale**
 - 4. Format d'import de données GINCO**
 - 5. Travail à réaliser sur les métadonnées**
 - 6. Problématique du regroupement de données**

Rappel des 3 modes (Échange, Diffusion et Communication)

Annexe 6 de la Charte



- Echange = transfert des DEE entre plateformes, PR<=>PN (INPN) pour ce qui nous concerne
- Diffusion = mise à disposition différenciée Services Etat (accès à la DEE sans floutage producteur ou sensibilité) vs. Grand public (accès à la DEE avec floutage producteur et sensibilité)
- Communication = accès à la DEE sans floutage producteur ou sensibilité soit par la voie de l'adhésion soit par un acte d'engagement (non adhérents)

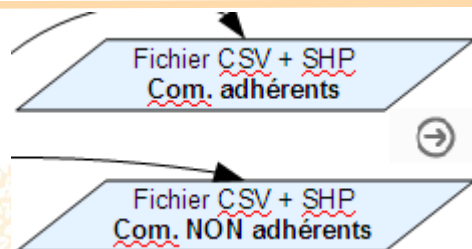
Propositions de principes :

- Le tour de table est proposé pour les demandes :
 - Pour les observations sensibles au sens du référentiel de sensibilité
 - Concernant une superficie importante (seuil à débattre)
 - Destinées à une étude ou publication scientifique
 - Émanent d'un non adhérent
- Pour toute demande on transmet systématiquement les données non sensibles non floutées + données sensibles floutées
- La transmission des données sensibles non floutées est soumise à l'approbation de la seule tête de réseau pour ce qui concerne le groupe taxinomique dont il a la responsabilité => pas de droit de veto sur les autres groupes taxinomiques

Propositions de principes :

- Le pôle taxinomique peut décider de ne pas activer l'approbation systématique la communication des données sensibles non floutées
- Un délai « technique » de trois semaines maximum pour la livraison des données
- En cas d'absence de réponse du pôle : non livraison et avis d'absence d'autorisation de livraison OU livraison ?
- Ces principes sont formalisés dans un « règlement intérieur » du CoTech

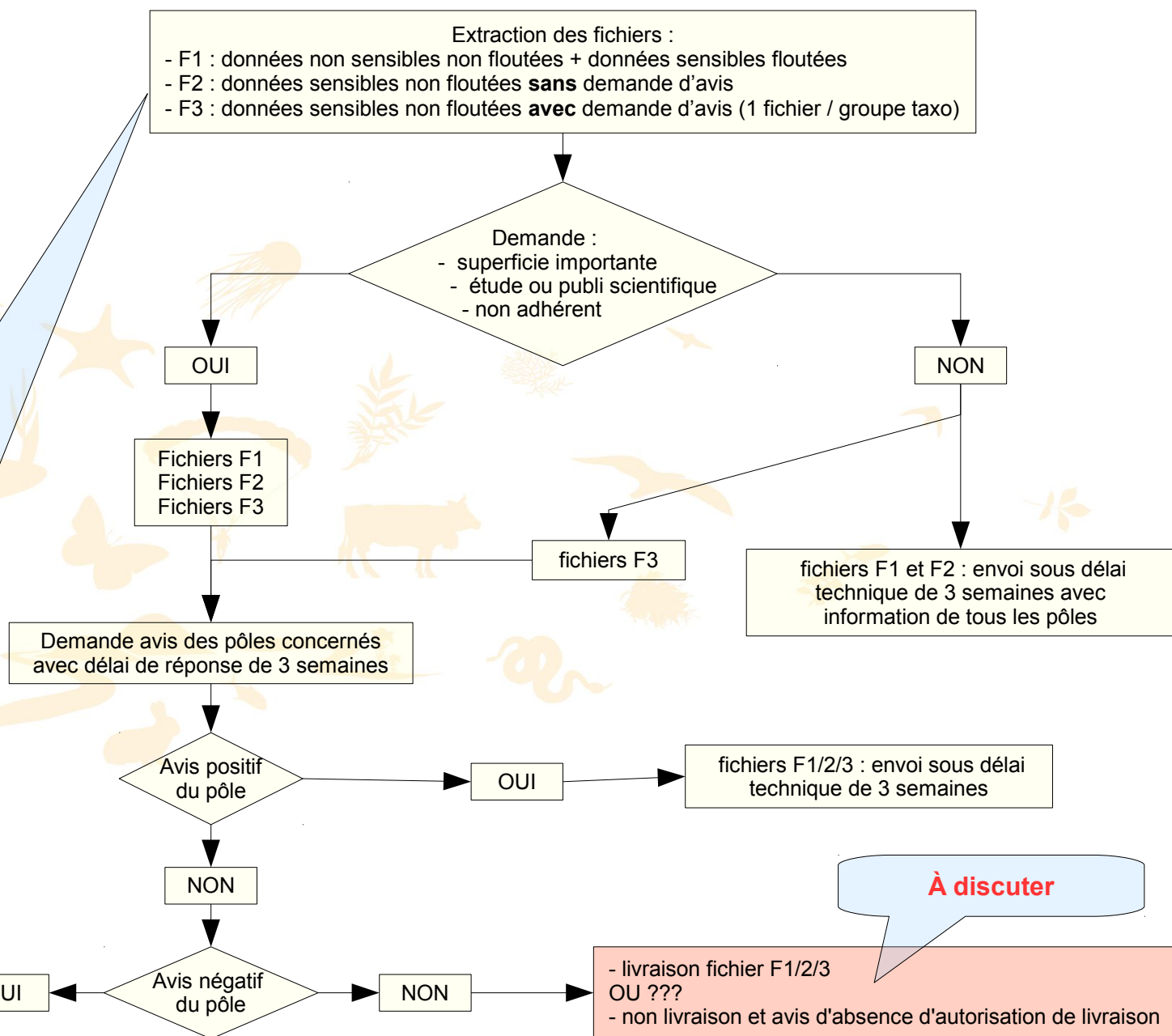
Modalités du 'tour de table' pour la communication des données



Fichier F1 : délivré sur simple information des animateurs de pôle sous un délai technique de trois semaines maximum. Ainsi le pétitionnaire disposerait déjà un bon niveau d'alerte.

Fichier F2 : délivré sur simple information des animateurs de pôle sous un délai technique de trois semaines maximum, et seulement pour les groupes taxinomiques pour lesquels les animateurs de pôles **ne souhaitent pas** donner leur approbation systématique.

Fichiers F3 : délivrés après tour de table des animateurs de pôles, sans droit de veto, et seulement pour les groupes taxinomiques pour lesquels les animateurs de pôles **souhaitent** approuver systématiquement la communication des données sensibles non floutées..



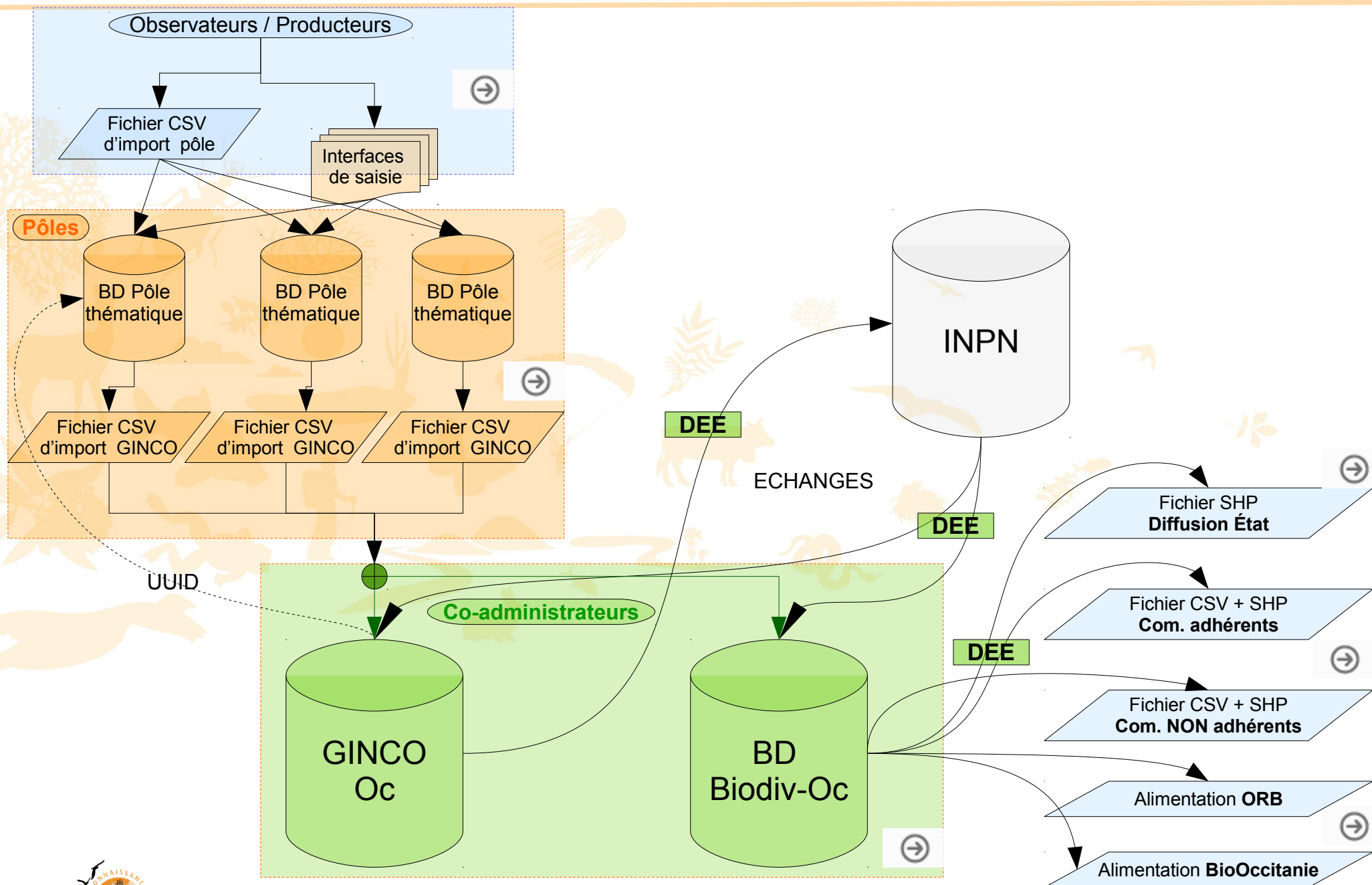
Plateforme régionale : Principes de la co-administration

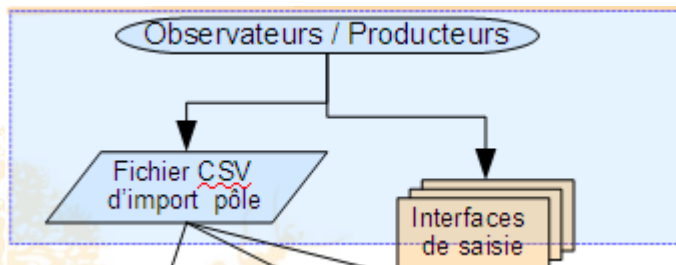
- Introduire un certain niveau de redondance afin d'éviter que le système soit inopérant si l'administrateur unique vient à faire défaut
- Proposition de partager l'administration entre la DREAL, les CENs et CBNs. La DREAL est administrateur principal et assure le guichet unique
- Objets du partage :
 - gestion des jeux de données
 - ✓ récupération auprès des animateurs de pôles
 - ✓ versement dans GINCO
 - Traitement des demandes de communication
- Une convention définit les tâches qui doivent être accomplies par les coadministrateurs ainsi que leurs modalités de partages

Propositions :

- CBNs : données fournies par le pôle Flore-Fonge-Habitats naturels, Poissons
- CENs : données fournies par les pôles Invertébrés, Mammifères, Reptiles-Amphibiens
- DREAL : données fournies par les pôles Oiseaux

Plateforme régionale : Organisation globale des flux de données





L'**observateur** : celui qui observe ou collecte l'information.

Le **producteur** : celui qui traite de manière informatique les données à la source.

Ces deux rôles sont souvent le fait d'une même personne.

Saisie :

- à partir d'un outil interne de saisie,
- sur une plateforme métier : ex. SILENE, Faune-Occitanie, etc. intégration directe dans la base du pôle mais ne couvre pas tous les groupes taxinomiques ,
- sur CardObs : respect des référentiels nationaux,
- sur GéoNature : respect des référentiels nationaux. Pourrait-il être mis en place une plateforme régionale GéoNature ? Hébergement / maintenance ? Étudier si la DREAL (/SG/UPSI) est en mesure d'opérer une telle plateforme.

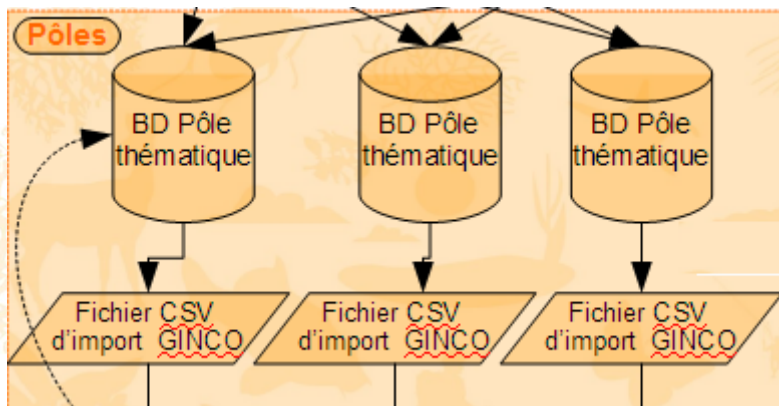
Transmission des données :

L'observateur/producteur transmet ses données numérisées aux pôles thématiques. Il transmet en même temps les éléments de métadonnées correspondants, les éventuelles directives de floutage à la diffusion, les éventuels souhaits d'anonymisation.

Transmission dans des formats stables, aux gabarits connus et publiés sur le portail du SINP-OC.

L'idéal serait un format unique. *A minima* formats connus du pôle thématique qui adapte des mécanismes d'import.

Possibilité que l'observateur/producteur produise et transmette l'identifiant unique de son observation (UUID).



Le pôle thématique maintient une base de données destinée à compiler toutes les données SINP-OC dont il a la charge.

Enregistrer :

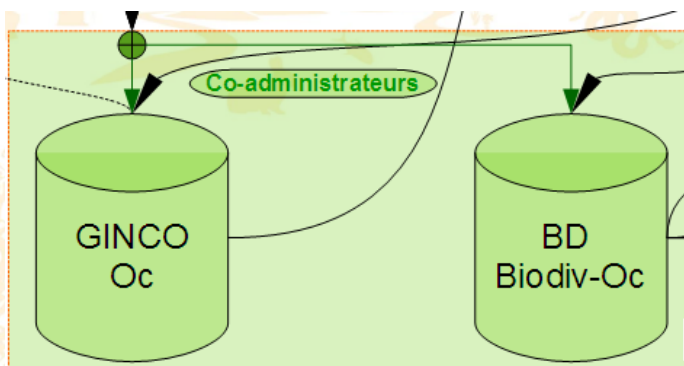
- identifier les données SINP
- récupérer et enregistrer les éléments de métadonnées
- enregistrer les souhaits d'anonymisation
- discriminer données publiques / données privées
- pour les données privées obtenir et enregistrer les éventuelles directives de floutage à la diffusion

Valider :

- mettre en place un processus de validation des données SINP

Transmettre :

- transférer aux co-administrateurs de GINCO les données SINP au format d'import de GINCO : données + éléments de métadonnées.



Les coadministrateurs importent les données dans GINCO et organisent la communication des données auprès des utilisateurs.

Q ? : que fait-on des (quelques) données non rattachées à TAXREF ?

lb_lot	grlr	no_taxref	% total
Données SILENE Flore	0	7	0.00 %
Poissons	710	710	3.87 %
ZNIEFF points	107	114	0.34 %
Chiroptères	1220	1220	2.63 %
Papillons	0	72	0.03 %
Oiseaux	968	1721	0.07 %
Reptiles-Amphibiens	5587	6048	4.47 %

Importer dans GINCO :

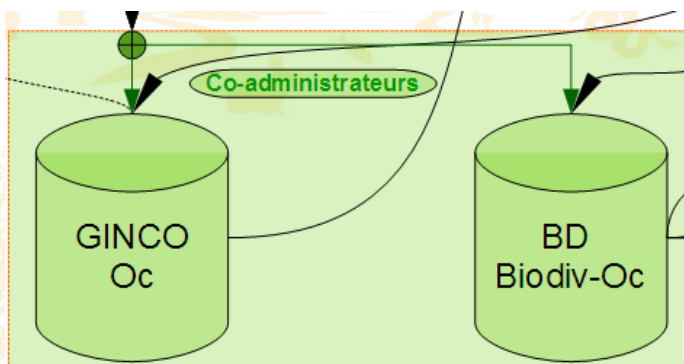
- récupérer auprès des pôles thématiques les lots de données au format d'import de GINCO (CSV...)
- créer la métadonnée sur la plateforme nationale de métadonnées
- importer les lots de données :
 - opérations d'import
 - retours vers les pôles des éventuelles erreurs d'importation pour résolution
 - retours vers les pôles des identifiants uniques UUID créés lorsque ceux ci ne sont pas générés par le pôle.

Organiser la communication :

En attendant que GINCO offre les fonctionnalités adéquates :

- Constituer une base de donnée 'miroir' destinée aux exports : hébergée à la DREAL
- Cette BD est accessible via Internet par PHPpgAdmin (Port 80) aux co-administrateurs
- Alimentée par les fichiers CSV d'extraction des bases des pôles au format d'import dans GINCO
- Contient des éléments permettant d'enrichir les DEE communiquées (statuts des taxons, calcul du groupe taxinomique, référentiel taxinomique régional, calcul de la sensibilité).
- Gestion maintenance de la structure et des données par la DREAL.

Format d'import de données GINCO



Modèles GINCO

ChD1	ChD2	ChD3	ChD4	ChDN	ChR1	ChR2	ChRN
enr11		enr13		enrN	enr16	enr17	enrN
enr21		enr23		enrN	enr26	enr27	enrN

Modèle de données

ChD1	ChD3	ChDN	ChR1	ChR2	ChRN
enr11	enr13	enrN	enr16	enr17	enrN
enr21	enr23	enrN	enr26	enr27	enrN

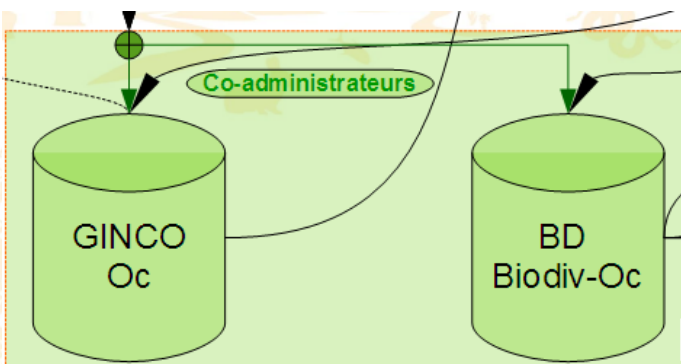
Modèle d'import

```

1 ChD1;ChD3;ChDN;ChR1;ChR2;ChRN
2 enr11;enr13;enrN;enr16;enr17;enrN
3 enr21;enr23;enrN;enr26;enr27;enrN
    
```

Fichier d'import

Format d'import de données GINCO



Modèles GINCO

Modèle de données :

modele_donnee_occitanie_v1

131 attributs répartis dans 11 concepts

concept	nombre d'attributs
SujetObservation	45
Source	21
DescriptifSujet	13
Commune	6
Maille10x10	5
Departement	4
EspaceNaturel	4
MasseEau	4
RegroupementObservations	3
Validation	24
AttributTechnique	2
	131

Défini à partir du modèle par défaut (obligatoire). Basé sur la version 1.2.1 du standard OCCTAX

http://standards-sinp.mnhn.fr/occurrences_de_taxons_v1-2-1/

Proposition de rajouter les enrichissements suivants :

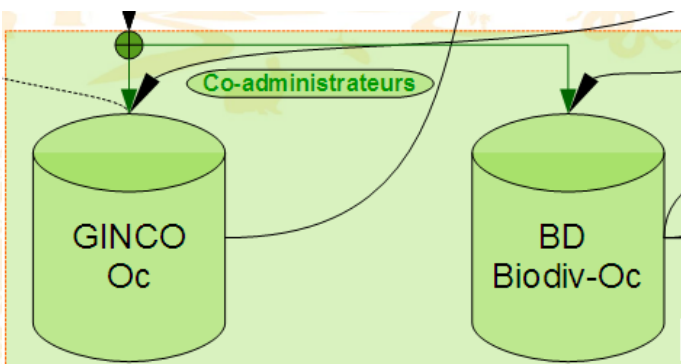
- Attributs décrivant les processus de validation régionale ou nationale selon les modalités du standard OCCTAX Version 2.0

<http://standards-sinp.mnhn.fr/occurrences-de-taxon-v2-0/>

- Attribut 'Groupe taxonomique' ventilation par groupes taxonomiques. Un groupe taxonomique s'étend au maximum sur l'emprise taxonomique d'un pôle thématique. Plusieurs groupes taxonomiques peuvent être définis pour un même pôle thématique (ex. : pour le pôle Invertébrés, au minimum un groupe 'Invertébrés' ou possibilité de décliner en plusieurs groupes, par exemple 'Insectes', 'Arachnides', 'Mollusques'.

D'autres ajouts sont-ils nécessaires ?

Format d'import de données GINCO



Modèles GINCO

Modèle d'import: modele_import_occitanie_v1
87 attributs répartis dans 9 concepts

Détail des champs retenus dans l'annexe 1 du document process_ginco_1.pdf transmis en amont de ce CoTech.

concept	nombre d'attributs
SujetObservation	34
Source	13
DescriptifSujet	13
Commune	4
Maille10x10	4
Departement	3
RegroupementObservations	3
Validation	12
AttributTechnique	1
	87

Format d'import de données GINCO



Requêteur

Type de données
modele_donnee_ocdanie_v1

Formulaires

Observation

Critères Sélectionner..

Résultats Sélectionner..

nomCite
observateurIdentite
observateurNomOrganisme
statutObservation

Localisation

Critères Sélectionner..

Résultats Sélectionner..

Regroupements

Critères Sélectionner..

Résultats Sélectionner..

Standardisation

Critères Sélectionner..

Résultats Sélectionner..

dSPublique
statutSource

Autres

Critères Sélectionner..

Annuler Réinitialiser **Rechercher**

Carte Résultats

Couches & Legendes

Couches Legendes

Résultats de la rech

Départements
Mailles
Communes
Géométries préc

Limites administrati

Régions
Départements
Mailles 10km
Communes

EN - Protections rég

EN - Protections cor

EN - Protections pa

EN - Protections au

EN - Natura 2000

EN - ZNIEFF

Fonds

Orthophctograp

Carte TGN

Format d'import de données GINCO



Requêteur

Type de données

modele_donnee_occitanie_v1

Formulaires

Observation

Critères Sélectionner...

Résultats Sélectionner...

nomCite

observateurIdentite

observateurNomOrganisme

statutObservation

Localisation

Critères Sélectionner...

Résultats Sélectionner...

Regroupements

Critères Sélectionner...

Résultats Sélectionner...

Standardisation

Critères Sélectionner...

Résultats Sélectionner...

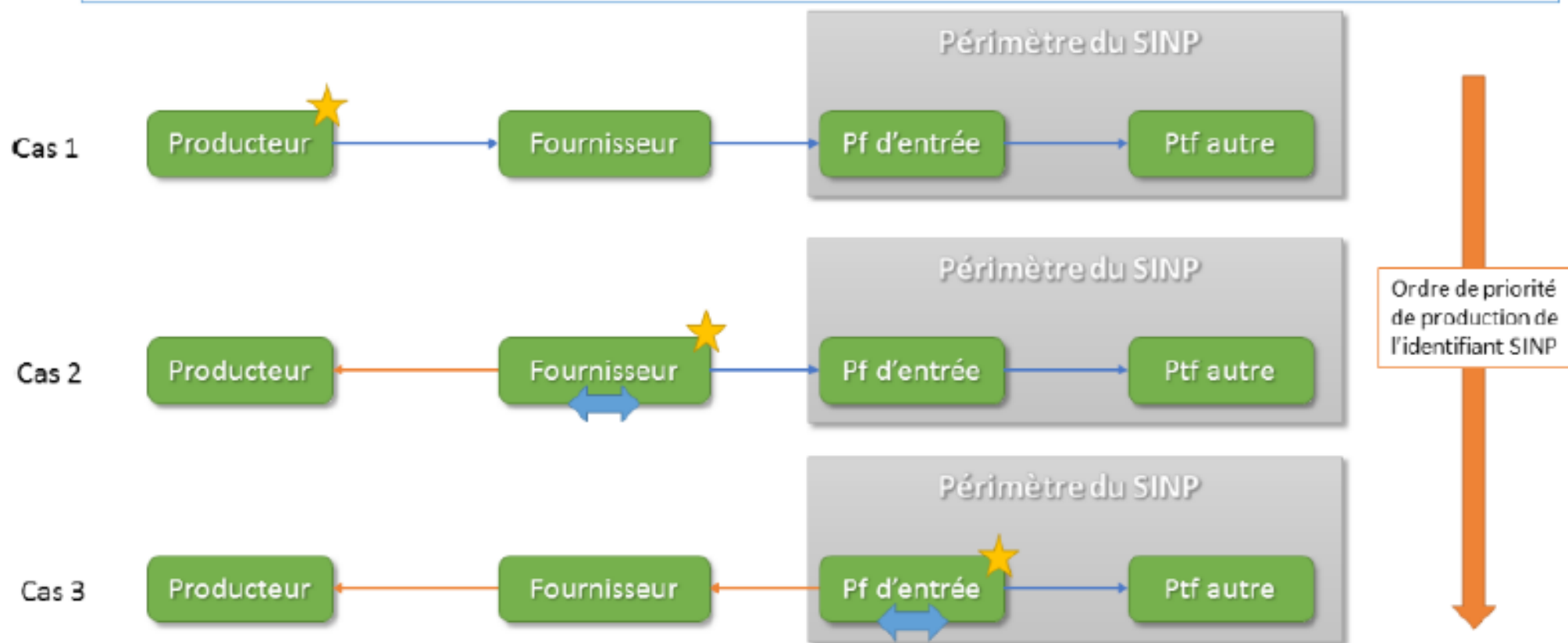
Carte

Résultats

	nomCite	observateurIde...	observ...	statutObserv...	dSPublique	identifiantPerma...	jddMetadonnee...	organismeGe	
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	2c1222b2-6680-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	7df55033-99f2-...	7289F131-A140...	SMCG
		Bufo bufo / spin...	EXERTIER Coline	Ecologis...	Présent (Pr)	Privée (Pr)	b1565520-73ce-...	7289F131-A140...	Ecologistes d
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	38063a20-d63a-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	9f4a5163-48e3-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	1d2b309d-c8d0-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	931f87ee-c586-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	5672412a-9e73-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	eb3745c0-eca9-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	41138ff4-1b98-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	bc7e6a8f-52c0-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	721e4ed4-720b-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	524e65b8-109c-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	4434979a-25d7-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	5ac80168-235c-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	cd9cf1ea-bfbd-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	5d039841-964f-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	17796e36-2f7b-...	7289F131-A140...	SMCG
		Emys orbicularis	Bosson Tom	SMCG	Présent (Pr)	Publique (Pu)	34528e57-8749-...	7289F131-A140...	SMCG
		Emys orbicularis	Petit Jérémiah	SMCG	Présent (Pr)	Publique (Pu)	9ef269fd-b5d4-...	7289F131-A140...	SMCG

Identifiant unique de la donnée

Chaîne de responsabilité de production de l'identifiant SINP



Légende

Système de gestion de données

Production de l'id SINP

Interface de correspondance entre id origine et id SINP

L'identifiant est partagé avec la donnée

L'identifiant est propagé vers le producteur

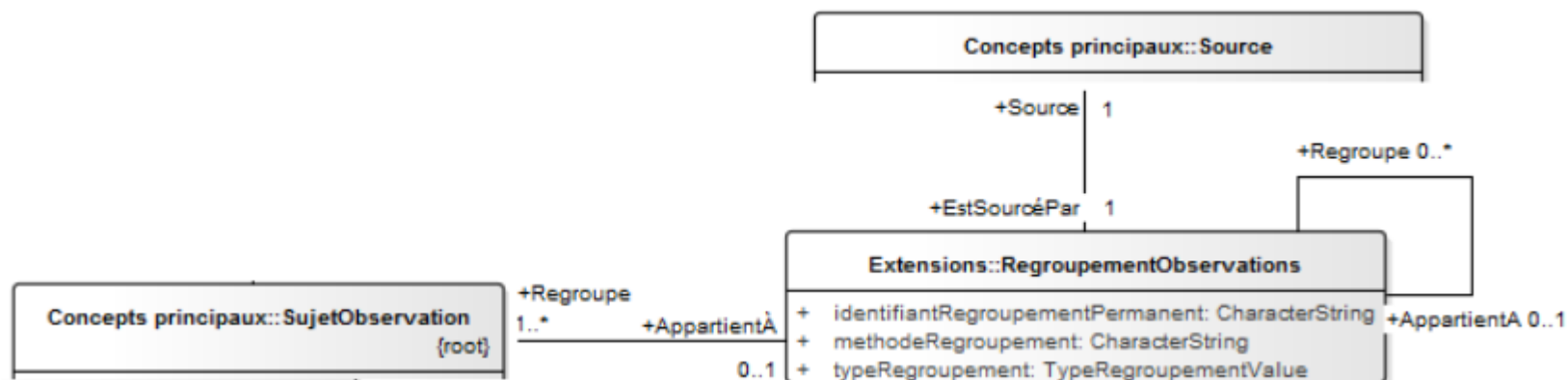
Tiré de : http://www.naturefrance.fr/sites/default/files/fichiers/ressources/pdf/note_identifiant_sinp_1.pdf

Travail à réaliser sur les métadonnées



- Cadre d'Aquisition ← Jeux de données
- Rassembler les données par cadre d'acquisition au sein desquels on distingue des jeux de données
- Comment «rassembler» les données en CA et JDD pour être efficient mais sans y passer un temps trop important ?
- DUPONT P. & TOUROULT J. 2018. Guide pour la constitution des jeux de données du SINP et le renseignement des métadonnées associées. Rapport UMS PatriNat (AFB, MNHN, CNRS), 41 pp.

Problématique du regroupement de données



Le regroupement des données est un concept prévu par le standard d'occurrence du taxon comme une extension

Objectifs, exemples :

- Rendre sensibles toutes des observations d'un relevé (ex. : floristique) où l'une des observations est sensible
- Regrouper des informations d'observations d'occurrences de taxon par strate dans un relevé phytosociologique
- Un papillon peut être groupé avec sa plante hôte
- Un parasite peut être groupé avec son hôte

Problématique du regroupement de données

Ne pas confondre le regroupement de données avec le rassemblement de données dans un jeu de données

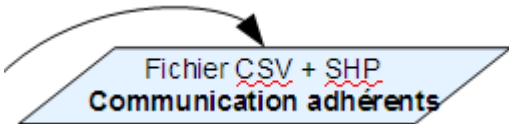
JDD_1 et JDD_2 peuvent éventuellement appartenir au même Cadre d'Acquisition

identifiantunique	identifiantRegroupement	identifiantJDD	Quoi	Qui	Où	Quand	Comment
UUID_OBS1		UUID_JDD_1	taxonX	Auteur1	Point_6	Date_1	Observé
UUID_OBS2		UUID_JDD_1	TaxonY	Auteur1	Point_7	Date_2	Observé
UUID_OBS3		UUID_JDD_1	TaxonY	Auteur2	Point_8	Date_3	Observé
UUID_OBS4		UUID_JDD_1	TaxonY	Auteur3	Point_9	Date_4	Observé
UUID_OBS5		UUID_JDD_1	TaxonY	Auteur4	Point_10	Date_5	Observé
UUID_OBS6		UUID_JDD_1	TaxonY	Auteur2	Point_11	Date_3	Observé
UUID_OBS7	UUID_R1	UUID_JDD_1	taxonX	Auteur1	Point_1	Date_1	Observé
UUID_OBS8	UUID_R1	UUID_JDD_1	TaxonY	Auteur1	Point_1	Date_1	Observé
UUID_OBS9	UUID_R1	UUID_JDD_1	TaxonZ	Auteur1	Point_1	Date_1	Observé
UUID_OBS10		UUID_JDD_2	TaxonH	Auteur4	Poly_1	Date_2	Observé
UUID_OBS11		UUID_JDD_2	TaxonH	Auteur4	Poly_2	Date_3	Observé
UUID_OBS12		UUID_JDD_2	TaxonH	Auteur4	Poly_3	Date_4	Observé
UUID_OBS13	UUID_R2	UUID_JDD_2	TaxonZ	Auteur1	Point_14	Date_5	Observé
UUID_OBS14	UUID_R2	UUID_JDD_2	TaxonY	Auteur1	Point_14	Date_5	Observé
UUID_OBS15	UUID_R2	UUID_JDD_2	taxonX	Auteur1	Point_14	Date_5	Observé
UUID_OBS16	UUID_R2	UUID_JDD_2	TaxonV	Auteur1	Point_14	Date_5	Observé
UUID_OBS17	UUID_R2	UUID_JDD_2	TaxonU	Auteur1	Point_14	Date_5	Observé



Merci de votre attention.





La DREAL organise la répartition des dossiers extractions entre chaque Co-administrateur.

Traitement des dossiers d'extraction :

- Vérification de la complétude de la demande
- Saisie en base des éléments du formulaire de demande (dont import du périmètre d'extraction)
- Module de génération de la synthèse des données disponibles avec calculs :
 - nombre de communes intersectées
 - répartition des obs. par groupe taxo / sensible - non sensible / Statuts
 - discrimination lots avec/sans tour de table
- avec tour de table :
 - Demande sur plus de 50 communes, demande pour étude ou publication scientifique
 - Demande émanant d'un NON adhérent
 - Données sensibles
- si non : sans tour de table
- module d'extraction des lots :
 - Exploite les informations saisies et celles calculées par le module de génération de la synthèse des données disponibles pour exporter les lots de données correspondants.

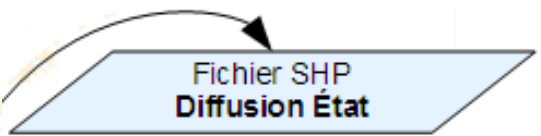
Répartition de l'envoi des lots :

- La DREAL maintient un fichier en ligne où sont listées les demandes d'extraction et les lots correspondants avec désignation du co-administrateur responsable de l'envoi du lot :
 - Les lots avec tour de table sont gérés par la DREAL
 - Autres lots répartis sur les autres co-administrateurs.

Le responsable de l'envoi : vérifie que le contenu du lot correspond à la demande, informe les pôles de l'envoi, procède à l'envoi du lot au destinataire, enregistre la date de l'envoi.

cf. logigramme





Fichier SHP
Diffusion État

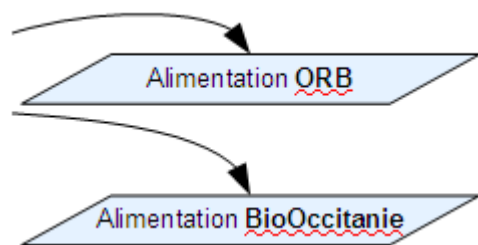
La DREAL se charge de la diffusion des DEE aux agents chargés de missions environnementales

Identifier les agents :

- mise en place d'un processus d'identification des agents de l'État accédants aux DEE via les chefs d'unité bio. Identification et signature d'un acte d'engagement.

Organiser la communication :

- La DREAL extrait les DEE de la BD Biodiv-Oc:
 - pour la DREAL : export au format shapefile des données par groupes taxinomiques, éventuellement subdivisés géographiquement pour les groupes les plus importants (gestion de la taille des shapefiles). Ces shapefiles sont mis en visualisation et téléchargement sur Cartonat-Picto dont l'accès est identifié pour chaque agent (login-MdP). Mise à jour par remplacement global des shapefiles tous les six mois.
 - pour les DDT(M) export au format shapefile par groupes taxinomiques des données du département augmenté d'un buffer de 5 km. Mise à jour par remplacement global des shapefiles tous les six mois. Shapefiles mis en téléchargement sur le serveur DREAL.
 - pour l'AFB et ONCFS (OFB) : pour la direction régionale l'accès est organisé via la communication des DEE au format à définir selon leurs outils d'utilisation. Mise à jour incrémentielle tous les ? 6 ou 12 mois ?



La DREAL se charge de la mise à disposition des données à l'ORB et à BioOccitanie

Définir avec le CoTech le type de mise à disposition qui sera mis en place : adaptation à l'utilisation des données

- Fournir dans le même format que pour la **communication** (Fichiers types F1, F2, F3) ?
- Fournir uniquement dans un format fichiers type F1 de la V1 communication aux adhérents ?
- Fournir uniquement des données floutées *a minima* à la commune pour les données non sensibles et floutage du référentiel pour les données sensibles (= **Diffusion** au Sens Protocole et Charte) ?