



Mandat du groupe de travail

Standard de données du SINP

Version validée



Historique des versions du document

Versio	Date	Auteur	Commentaire
v1	23/11/2011	Laurent Poncet , SPN/MNHN (LP)	version initiale
v1	01/12/2011	Simon Chagnoux, DSI/MNHN (SC)	Version de travail
v1	01/12/2011	Thierry Bourgoïn, Collection/MNHN (TB)	Version de travail
v1	01/12/2011	Régine Vignes, GBIF-Fr (RV)	Version de travail
v2	13/12/2011	Laurent Poncet, SPN/MNHN	version discussion
V3	19/12/2012	Laurent Poncet, SPN/MNHN	Suite réunion de travail 15/12
V4	20/12/2011	Thierry Bourgoïn, Collection/MNHN	Suite réunion de travail 15/12
V5	23/12/2011	TB, RV S. Chagnoux / T. Bourgoïn MNHN	Relecture et modifications
V5	23/12/2011	Laurent Poncet, SPN/MNHN	Envoi ministère
V5.1	16/01/2012	Yannick Lebeau/DGALN/DEB	Relecture et modifications
V6	24/01/2012	LP, SC, TB, RV	Nouvelle version
V7	13/02/2012	Yannick Lebeau/DGALN/DEB	Propositions de modifications
V8	30/03/2012	CST (Laurent Coudercy)	Projet V0 de saisine de la CST
V8	20/04/2012	LP, SC, TB, RV	Prise en compte des remarques de la CST et ajustement du calendrier
V9	02/05/2012	Y. Lebeau/DGALN/DEB	Mise en forme charte graphique SINP
V10			Version finale

Rédacteurs

Laurent Poncet, SPN/MNHN , Régine Vignes, GBIF-Fr (RV) , Thierry Bourgoïn, Collection/MNHN (TB), Simon Chagnoux, DSI/MNHN (SC)

Relecteurs

Yannick Lebeau/DGALN/DEB

Validation

Diffusion

Publiucation sur Naturefrance



Table des matières

Le SINP.....	4
Contexte et enjeux.....	4
Synergies avec d'autres actions du SINP.....	5
Standards existants.....	5
Préalables : précision des besoins des acteurs et périmètre des données.....	6
Objectifs du groupe de travail sur les formats standards de données du SINP.....	6
.....	7
Démarche proposée et livrables associés.....	7
Phase 1 : recensement de l'existant	7
Phase 2 : recensement des besoins et des critères.....	7
Phase 3 : rédaction d'un document de cadrage méthodologique sur le format standard de données.....	7
Phase 4 : définition des spécifications techniques.....	8
Phase 5 : Evaluation et veille.....	8
Organisation, composition et pilotage.....	9
1er cercle.....	9
2ème cercle.....	9
3ème cercle	10
Calendrier prévisionnel.....	11
Calendrier.....	11
Validation des travaux du groupe	11
Suites du groupe de travail.....	12



Le SINP

Le système d'information sur la nature et les paysages (SINP) est défini dans la circulaire du 11 juin 2007. Le SINP est une organisation collaborative favorisant une synergie entre les acteurs pour la production, la gestion, le traitement, la valorisation et la diffusion des données sur la nature et les paysages. Pour faciliter la participation et l'interaction de tous les acteurs, aussi bien nationaux que locaux, l'organisation repose sur un niveau national et un niveau régional. Le SINP facilite la mise en relation d'informations tout en offrant un cadre méthodologique de référence.

Ce système d'information implique donc les services de l'État, notamment les DREAL, DEAL et DRIEE et les établissements publics de l'État, mais aussi les organismes de recherche, les collectivités et les autres producteurs de données publics ou privés. Il concerne à la fois la métropole et les départements d'outre-mer qui ont les mêmes objectifs de connaissance sur la nature et les paysages.

La gouvernance du SINP repose sur deux instances complémentaires :

- pour la conduite générale du dispositif, un comité de pilotage regroupe les producteurs de données et les principaux acteurs du SINP. Il est animé par la direction général de l'Aménagement, du Logement et de la Nature (DGALN) avec l'appui d'une équipe projet ;
- au plan scientifique, une coordination scientifique et technique (CST) garantit l'activité de production scientifique et technique à la fois de l'Observatoire National de la Biodiversité (ONB) et du SINP, avec une configuration adaptée. Elle est co-animée par le Service du Patrimoine Naturel (SPN) du Muséum national d'Histoire naturelle (MNHN) et le Laboratoire dynamiques sociales et recomposition des espaces (LADYSS) du CNRS. Cette coordination est composée de scientifiques représentant tout à la fois les acteurs du domaine et les champs de compétence couverts par le SINP et l'ONB.

La CST s'assure de la qualité scientifique et technique des différentes actions du SINP et de l'ONB. Des groupes de travail peuvent être créés pour mener à bien certaines de ces actions. Dans ce cas, la CST formule un avis sur les mandats des groupes de travail et sur les résultats intermédiaires et finaux.

Contexte et enjeux

Une multitude de données sur la nature et les paysages est produite chaque année par de nombreux acteurs sur les territoires, selon des méthodes et des formats différents, ce qui rend complexes les échanges de données et la nécessaire production de synthèses à différentes échelles.

L'audit réalisé en mai 2009 par le Cemagref auprès des DIREN-DREAL et des établissements publics membres du SINP a notamment mis en évidence une attente forte



des acteurs en termes de mise en place de méthodes et outils pour le recueil et l'échange d'informations.

Synergies avec d'autres actions du SINP

Des actions déjà en place dans le cadre du SINP permettent de traiter une partie des questions correspondantes : référentiels taxonomiques, métadonnées, outil commun de saisie et de gestion de données d'observations naturalistes et architecture du SINP. Faciliter les échanges de données suppose, en plus en amont de ces actions, de définir un "vocabulaire" commun pour le partage de l'information. Cela passe par un travail sémantique puis plus technique de définition de standards de données (interopérabilité sémantique et syntaxique). La réalisation de ce travail constitue un enjeu primordial pour la mise en place opérationnel du SINP (il n'intègre pas les standards potentiellement nécessaires dans le cadre de l'ONB).

Les protocoles d'échanges informatiques (mail, ftp, http, digir/http, biocase/http, tapir/http, ...) sont hors du champ de l'étude. Ce travail pourra rentrer dans le groupe de travail "Architecture » du SINP

Standards existants

Un certain nombre de données traitées dans le cadre du SINP sont également concernées par la directive Inspire et la convention d'Aarhus qui leur confèrent un statut particulier en obligeant l'autorité publique à leur communication la plus large possible. Des spécifications de données ont ainsi été définies (ou sont en cours de validation) dans le cadre de la directive Inspire pour les données relatives à la conservation de la nature : sites protégés (annexe I), régions biogéographiques, habitats et biotopes et répartition des espèces (annexe III).

Par ailleurs, il existe des standards reconnus et utilisés au niveau international (par le Taxonomic Data Working Group ou TDWG) pour le partage de données sur la biodiversité (ex. Darwin, ABCD,), mais plus spécifiques aux données de collection. D'autres standards internationaux permettent de normaliser certaines informations communément utilisées telles que les dates (normes ISO/ANSI). Il existe également des initiatives innovantes telles que Biological Science Collections Tracker (BiSciCol)...

Au niveau national, un certain nombre de structures mènent des réflexions et implémentent des profils ou formats standards de données (GBIF-Fr, Conservatoires botaniques nationaux, MNHN, BRGM, SINP régionaux, INPN, SI-eau ...).

Par ailleurs, des travaux similaires de standardisation devraient être menés dans le cadre du programme de cartographie nationale des habitats. En parallèle, des réflexions sont en cours sur les données paysages.

Les travaux du groupe de travail SINP relatif aux standards de données devra ainsi respecter des exigences européennes (Directive INSPIRE, notamment le thème installations de suivi environnemental – annexe III) et tenir compte des expériences existantes pour s'intégrer de manière cohérente dans les démarches en cours.



Préalables : précision des besoins des acteurs et périmètre des données

Le protocole SINP et la circulaire associée définissent un périmètre général pour les données du SINP. Ce périmètre a été rediscuté lors du comité de pilotage SINP du 1^{er} juillet 2011 et devra être validé avant le lancement du groupe de travail "formats standards de données du SINP".

L'action propose de définir un format d'échange de données élémentaires au niveau national, défini en aval de la production pour échanger des données existantes dans diverses bases de données des producteurs et produites selon des protocoles différents.

Une attention particulière sera portée à la gestion des échelles et de leur implication dans le standard de données (local, régional, national, international): subdivision en niveaux hiérarchiques comme pour le standard géographique proposé par le TDWG (*World Geographical Scheme for Recording Plant distribution, edit. 2*) ou tout autre système.

L'objectif du GT est de définir par grandes thématiques du SINP un format national (tel que prévu dans le projet de protocole en tant que données élémentaires d'échange ou DEE) compatible avec les standards existants au niveau international.

Par rapport au périmètre du SINP, cette action traitera par ordre de priorité:

- les données espèces,
- les données espaces,
- les données habitats,

Par la suite, les thématiques suivantes pourront être abordées:

- les données de géologie,
- les données décrivant les écosystèmes,
- les données paysages,

Objectifs du groupe de travail sur les formats standards de données du SINP

L'objectif est de proposer une description commune des données pour permettre à terme l'interopérabilité des systèmes d'information et ainsi faciliter les échanges d'informations entre les différents projets et acteurs du SINP. Il s'agit donc de :

- recenser les formats standards existants dans le domaine des données sur la nature et le paysage et en faire une analyse (en abordant les différentes



thématiques existantes);

- pour la thématique «biodiversité», et en particulier les données d'observation, définir un standard minimal (ou cœur commun) en tenant compte de l'existant ;
- spécifier les structures et représentations des données retenus pour les échanges numériques (spécifications techniques) ;
- de porter à connaissance de l'ensemble des acteurs du SINP ces formats standards.

Démarche proposée et livrables associés

Phase 1 : recensement de l'existant

Cette phase préalable vise à faire un premier état de l'art des formats standards les plus avancés existants au niveau français (régional, national) et de rechercher les exemples les plus structurants au niveau international.

Cette synthèse sera présentée sous forme de fiches d'expérience et d'un tableau récapitulatif.

Phase 2 : recensement des besoins et des critères

Cette phase 2 est centrale parce qu'elle permettra à terme de fixer les standards du SINP. La phase 2 comportera 5 étapes :

- 1) Recensement des besoins et des attentes à partir des travaux conduits dans le cadre du GT Architecture du SINP et à partir des principes énoncés dans le nouveau protocole SINP : Pratiques régionale, nationale et internationale, Quelles informations sont recherchées? Quels flux ?
- 2) Détermination des critères de choix entre les standards existants. Traitement du problème d'échelle entre les logiques régionale, nationale et internationale.
- 3) Analyse des manques: adéquation entre le recensement des standards existants et le recensement des besoins. Dans cette phase, les standards identifiés dans la phase 1, seront examinés pour déterminer les avantages et inconvénients scientifiques et de mise en œuvre.
- 4) Proposition de hiérarchisation du traitement des manques identifiés.

Cette phase 2 s'appuiera sur des sous-groupes de travail SINP.



Phase 3 : rédaction d'un document de cadrage méthodologique sur le format standard de données

Réalisation d'un document de cadrage méthodologique : cette phase aboutira à un cadrage méthodologique avec les grands axes du format de données SINP, en distinguant les attributs obligatoires (correspondants aux utilisations nationales des données élémentaires d'échanges listées dans le projet de protocole SINP) et les attributs optionnels pour couvrir le périmètre du SINP. Cette phase permettra de produire un document de référence pour les acteurs du SINP.

Il comprendra :

- la liste et la sémantique des attributs obligatoires (y compris le choix des vocabulaires contrôlés)
- la liste et la sémantique des attributs optionnels
- les principes de la définition de l'identifiant national

Ce document sera soumis à la CST. Il est important à ce stade que l'analyse des manques, les choix opérés, l'identifiant national et la liste des attributs obligatoires ou optionnels fassent l'objet d'un avis scientifique et technique.

Phase 4 : définition des spécifications techniques

A recueil des avis de la CST, cette phase permettra de traduire les travaux d'ordre sémantique sous forme de formats exploitables par des applications informatiques :

- **rédaction des spécifications techniques** (type Darwin Core): livrables de thésaurus ainsi que la définition de l'identifiant national unique des données élémentaires d'échange. La logique de cette identifiant unique ayant été défini par le GT architecture
- **format d'échange** : définition des attributs du format standard (nommage des champs, syntaxe, règles de gestion, contrôles, règles de codage informatique, norme de langage XML, GML; format de fichier, etc..). Cette phase ne concerne pas les protocoles informatiques d'échange (FTP, flux ou autres) qui restent dans le champ du GT architecture SINP.

Phase 5 : Evaluation et veille

En parallèle de la phase 4, il sera nécessaire de travailler sur tous les manques devant



être traités par le SINP. Un groupe de veille devra être mis en place afin d'évaluer les manques du format adopté, de suivre le développement des standards internationaux (GBIF, INSPIRE, etc..) et veiller à leur bonne interopérabilité avec le standard français de données élémentaires d'échanges.

La méthodologie exposée ci-dessus pour la définition d'un format de données élémentaires d'échange (DEE) dans le domaine des observations naturalistes sera évaluée, éventuellement amendée et reconduite sur les autres formats thématiques à mettre en œuvre dans le SINP : paysages, géologie, génétique...

Organisation, composition et pilotage

Le groupe est piloté et animé par le MNHN (SPN, DSI, GBIF-Fr). Il comprend des membres organisés en trois cercles, conformément au modèle de gouvernance SINP.

1^{er} cercle

Un premier cercle est composé des membres actifs qui participent aux réunions, travaux et à la rédaction des documents intermédiaires ou finaux. C'est à ce niveau que ce fait le lien avec les autres groupes de travail du SINP. Ce groupe pourra être divisé en sous-groupes thématiques le cas échéant.

- MNHN (SPN/DSI/Collection) / GBIF-Fr - **pilote**
- MEDDTL/DGALN/DEB/PEM4
- MEDDTL/DHUP – Paysage
- ATEN
- BRGM
- CNRS (UMS BBEES)
- FCBN
- IGN
- IFREMER
- ONEMA (SANDRE)
- 2 représentants de plateformes régionales SINP (à définir)

Ce groupe se réunira en formation plénière au minimum 4 fois. En fonction de l'ordre du jour, pourront être conviées toutes autres personnes dont la présence serait jugée nécessaire.

2^{ème} cercle

Le cercle 2 comprend un ensemble de contributeurs. Ces derniers participent aux travaux soit à l'occasion de réunions soit via des échanges sur la plate-forme collaborative. Ce cercle est le porte-parole et le représentant des utilisateurs. Il émet des avis, des



souhaits et en informe le cercle 1. Il est également tenu du retour d'information vers les utilisateurs, le moment venu.

- CBNx
- RNF
- FNE
- PNF
- FCEN
- FPNRF
- Autres représentants de plateformes régionales SINP (à définir)
- ONCFS
- SFEPM
- OPIE
- SHF
- ONF
- ONEMA
- LPO
- AAMP
- Natureparif

3^{ème} cercle

Un troisième cercle de membres informés, qui sont explicitement destinataires de l'état d'avancement des travaux du groupe, sans réaction attendue de leur part.

L'outil de travail collaboratif du SINP servira de support aux échanges entre les participants, avec le niveau d'accès spécifique au cercle auquel ils appartiennent.



Calendrier prévisionnel

A préciser avec les membres du cercle 1 en fonction de la démarche validée,

Calendrier

échéance	phase	étape
Mai-juin- juillet 2012	0	Initialisation du projet, recrutement
août-sept 2012	1	Étude de l'existant, appel d'offre
oct. 2012 / fev 2013	2	Recensement des besoins et des critères
mars/mai 2013	3	rédaction d'un document de cadrage
		Saisine CST sur le document de cadrage
		Avis de la CST
juin 2013	4	Rédaction des spécifications techniques
juillet. 2013	4	Saisine CST sur les livrables du GT
septembre 2013	4	Publication du format
septembre 2013	5	Mise en place de l'évaluation et veille du standard national

Nota :

Après septembre 2013, une organisation continue sera mise en place pour prendre en compte le côté itératif de la révision du standard (nouveau besoin, évolution des standard internationaux...).

Un nouveau cycle de définition de format (paysage, géologie, autres..) pourra être enclenché à partir septembre 2013.

Validation des travaux du groupe

Le mandat du groupe de travail sera transmis à la coordination scientifique et technique (CST) conjointe au SINP et à l'ONB, pour validation.

La CST désignera un de ses membres comme correspondant-référent pour ce groupe de travail et éventuellement un ou plusieurs membres qui participeront aux travaux du groupe.

Le groupe rendra compte de ses travaux à la CST pour validation ou avis.

Suites du groupe de travail

Le travail de l'action "formats standards" pourrait être prolongé par une structure légère de quelques personnes afin de suivre le plan d'accompagnement des acteurs du SINP pour l'implémentation de ces formats.