



SINP

Système d'Information
de l'Inventaire du Patrimoine naturel



Comité technique Système d'information de l'inventaire du patrimoine naturel (SINP)

Jeudi 20 octobre 2022

Réunion en présentiel et visioconférence

Présent.e.s :

Salle :

PatriNat : Solène Robert, Noëlie Maurel, Julien Touroult, Judith Panijel
Ministère de la Transition écologique (direction de l'eau et de la biodiversité) : Nathalie Poulet
Observatoire de la faune sauvage de Nouvelle-Aquitaine (FAUNA) : Paul Fromage
Conservatoire botanique national (CBN) du Bassin Parisien (représentant le réseau des CBN) : Silvère Camponovo
BRGM : Juliette Stephan-Perrey
Tela Botanica : Christophe Girod
IGN : Xavier Thauvin
Ligue pour la protection des oiseaux (LPO) : Philippe Jourde

Visioconférence :

Agence régionale de la biodiversité (ARB) de Bourgogne-Franche-Comté : Franck Grossiord
Conservatoire du littoral : Pascal Cavallin
Région Ile-de-France : Antoine Charpentier
Société française pour l'étude et la protection des mammifères (SFEPM) : Romain Baghi
Fédération nationale de chasse (FNC) : Gwendal Trouillet
Direction de l'innovation numérique et des systèmes d'information (DINSI) du MNHN : Francis Clément
Réserves naturelles de France (RNF) : Antoine Schlegel
Ifremer : Philippe Gouletquer

Propos introductifs

Julien Touroult (PatriNat) et Nathalie Poulet (ministère de la transition écologique) introduisent la réunion en rappelant l'enjeu de mieux lier les aspects techniques du SINP (outils, services, etc.) avec les autres actions portées par les autres comités du dispositif, et en particulier avec le comité des inventaires. Le comité technique doit par exemple permettre de répondre de manière opérationnelle aux enjeux identifiés en termes d'apport de nouvelles données et/ou de nouvelles thématiques couvertes.

Les quatre grands axes stratégiques validés par le copil de juin 2022 sont rappelés et seront détaillés en point 2 de l'ordre du jour. Il est important de mettre l'accent sur :

- l'ouverture des données, la fluidification des échanges (faciliter l'accès aux données)
- la mutualisation des services et outils développés
- la qualité : les travaux autour des métadonnées, des standards et des référentiels sont ainsi à conforter

Le comité technique dans la nouvelle gouvernance SINP

Cf. support de présentation

La genèse du SINP / L'atteinte d'une maturité

Et maintenant, le schéma métier du SINP

Quelles évolutions ? Quelles constantes ? Quel chemin emprunté ?

Redéfinition du périmètre / Une nouvelle gouvernance

Les axes stratégiques du SINP

Cf. support de présentation

4 grands axes :

- ➔ les données ouvertes (open data)
- ➔ la cohérence du SINP à tous les échelons territoriaux et pour l'ensemble des programmes et dispositifs contributeurs
- ➔ l'enrichissement mutuel et ciblé du SINP par les différents programmes et dispositifs contributeurs
- ➔ la qualité des données

Discussion :

La qualité des données est un enjeu majeur du dispositif, au même titre que la diffusion rapide des données. Il y a donc un enjeu à ce que les procédures de validation ne freinent pas la disponibilité des informations.

On sait en effet que la validation scientifique des données demande un investissement en temps humain important. La conciliation des deux enjeux passe donc notamment par une recherche d'automatisation des procédures via une expertise amont consistant à mettre en place des bases de connaissances, couches de références, etc.

Il est également important d'outiller les validateurs et de travailler la mutualisation des services de validation.

Un travail prospectif est en cours sur la possibilité de proposer aux acteurs un service générique de validation (avant dépôt des données). L'IGN indique également avoir prévu de mettre à disposition une API de validation dans le cadre du programme Géoplateforme (<https://github.com/IGNF/validator>).

Les outils et services mutualisés pourraient permettre de réduire le temps entre la production d'une donnée (observation sur le terrain) et sa diffusion sur les plateformes du SINP.

Il est rappelé que le SINP doit diffuser les données quel que soit leur niveau de validité. Il s'agit cependant de bien orienter/construire les outils de diffusion pour qu'il n'y ait aucune ambiguïté pour l'utilisateur sur la validation scientifique des données qu'il consulte.

Il n'est donc pas possible de bloquer la mise en diffusion des données parce qu'elles n'auraient pas été validées. Cela risquerait par ailleurs de bloquer la mise à disposition de données pour des questions de disponibilité d'experts, alors même que nous savons qu'il y a des groupes taxonomiques orphelins ou des territoires mal couverts sur certains groupes. Une donnée doit pouvoir circuler avec un niveau « non évalué ».

La question de la cohérence de la démarche technique du SINP, notamment en termes de standards, de procédure de validation, de référentiels, avec le niveau européen/international est posée.

Les travaux des différents groupes de travail et groupes projet du SINP (que ce soit celui sur la validation scientifique, celui sur la sensibilité des données, celui sur les standards) ont intégré dans la réflexion une analyse des travaux existants au niveau international. A titre d'exemple, le GT Validation s'est notamment basé sur les travaux d'Arthur Chapman (<https://www.gbif.org/document/80509/principles-of-data-quality>) ; le GT Standards étudie la compatibilité avec le darwin core (DwC) et l'EML (ecological metadata language) afin de rapprocher le standard français de ces standards lorsque cela est possible. Pour autant, il y a parfois des spécificités SINP qui ne permettent pas de reprendre directement les travaux internationaux. Dans ce cas, on recherche l'interopérabilité et on travaille les mises en correspondance (*mappings*).

La qualification d'un jeu de données (pour orienter sur les usages possibles) est à distinguer de la validation scientifique des données. Ce sont deux axes à travailler dans le cadre de la feuille de route à 3 ans.

Les projets en cours

Cf. support de présentation

Focus sur le projet GINCO :

Retour sur l'historique du projet et présentation de la nouvelle gouvernance.

Organisation du projet GINCO avec le projet GeoNature.

Récapitulatif des travaux 2022 et des axes de développement pour 2023.

Discussion :

Le programme Geoplateforme de l'IGN prévoit une partie des fonctionnalités attendus pour GINCO. La question de l'interfaçage des outils est posée. Plus globalement, cela pose la question de mieux mutualiser les différents outils existants lorsque cela est possible. Il pourrait être envisagé de labelliser des modules/briques techniques existantes et de proposer dans le cadre SINP un ensemble d'outils, de modules, etc. Plusieurs outils sont ainsi possibles, il faut donc travailler l'articulation de ces outils et services entre eux.

Ce travail nécessite une réflexion *ad hoc* et une présentation/connaissance plus fine des différents outils existants, de leur organisation et des calendriers. Il pourrait être envisagé de mettre en place un groupe de travail du SINP spécifique à ce sujet de la mutualisation des outils (le proposer comme une action de ce comité dans la feuille de route).

Au-delà de ce travail global, une discussion entre l'IGN et PatriNat concernant la complémentarité des outils de diffusion proposées pourrait être organisée, avec l'objectif de voir comment articuler ces services à moyen terme. A ce jour 4 régions utilisent une instance GeoNature hébergée au MNHN. Les développements réalisés dans le cadre du projet GINCO sont reversées à la communauté GeoNature et mis à la disposition de tous. Ils sont donc notamment utilisés par les régions qui utilisent GeoNature (avec une instance propre, non hébergée au MNHN).

On note également que de nombreux producteurs installent leur propre instance de GeoNature. Cela induit parfois un coût important pour l'installation ainsi que pour la gestion des montées de version, les installations de nouveaux modules et la maintenance des référentiels.

S'il est important de faciliter le déploiement de ces outils pour qu'il soit moins compliqué de les faire évoluer, il est aussi important de s'interroger sur le coût humain et énergétique lié à la multiplication des instances au sein d'un réseau régional. Il serait intéressant de réfléchir à la mutualisation des instances en incitant les producteurs à utiliser les instances régionales mises en place au sein des plateformes.

Panorama des projets en cours

Standards pour l'échange des données :

Standard pour les données sur les habitats publié, s'accompagne de gabarits -> en ligne sur l'INPN
GT en cours sur les standards espèces et métadonnées. Recherche d'interopérabilité avec DWC et EML.
Pour les évolutions de standard : compromis à trouver entre la prise en compte des évolutions et la stabilité au regard des impacts des changements sur les outils en place.
Déclinaison dans les outils (GINCo notamment) à prévoir en 2023.

GT métadonnées « structuration et accompagnement » :

Document de synthèse avec priorisation d'ores et déjà disponible
Déclinaison en supports « vulgarisés » en cours
Organisation d'un séminaire en 2023, couplé aux journées annuelles du SINP

GT Architecture – dossier « mises à jour des données » :

Publication d'une note de synthèse des besoins avec 3 scénarios proposés : plancher, augmenté, prospectif
À venir : test sur un prototype de service numérique (approche preuve de concept)

CAMPanule – référentiel des protocoles et méthodes :

Diffusion d'une première version du catalogue → en ligne sur l'INPN
(<https://inpn.mnhn.fr/programme/campanule>)
Référentiel à compléter au fur et à mesure des retours/besoins

Référentiel des organismes :

7 100 organismes référencés
Bancairisation et diffusion des logos des organismes prévues dans le webservice. Ce référentiel permettra de mieux valoriser les acteurs sur les différents outils, notamment sur le portail de l'INPN et son service OpenObs.

Outils de diffusion :

Espèces : OpenObs → nouvelle architecture technique installée. Mutualisation en cours avec le réseau des conservatoires botaniques nationaux (cadre SI Flore).
Habitats : perspective d'un visualiseur généraliste des données habitats pour fin 2023

Discussion :

La valorisation des réutilisations de données est abordée. Faut-il prévoir des métadonnées sur le produit d'une réutilisation ? Comment faire le lien entre un produit issu de diverses sources (jeux de données) et les données « sources » utilisées ?

Il apparaît important de bien citer les sources utilisées dans toutes nouvelles productions. L'utilisation des DOI (à l'image de ce que fait le GBIF) est également préconisée. Cela permet notamment de connaître les différents usages faits des données diffusées dans le SINP. C'est un point important pour les producteurs de données et cela aiderait à maintenir la dynamique de partage, la motivation du réseau.

Il semble important de prévoir de faire un lien entre des métadonnées « de synthèses » et les différentes métadonnées « sources ». Cela pose quand même la question de l'évolution de ces données sources

(corrections, ajouts de nouvelles données dans les jeux ...). Garder une trace des jeux historiques (sans modification) serait très coûteux en terme de stockage.

La question de la citation des sources pose question dès lors que ces sources sont mal renseignées et/ou mal normalisées. Il est donc important de bien renseigner les métadonnées, et de poursuivre les travaux d'accompagnement pour mieux normaliser les noms des observateurs. La question d'un référentiel des observateurs est posée. À ce stade, ce travail paraît très complexe. Il peut se mener à une échelle plutôt locale mais est difficilement applicable à une échelle plus large (nationale ou internationale).

Concernant les outils de diffusion, l'IGN propose un service de tuiles vectorielles. Cela pourrait potentiellement alimenter les outils de diffusion de données, ce qui irait dans le sens d'une mutualisation de services. À évaluer. API Entrepôt bêta mis en œuvre dans l'expérimentation Géotuileur (<https://www.ign.fr/geoplateforme/geotuileur/>) préfigurant la future API qui sera déployée sur la Géoplateforme.

OpenObs n'est pas prévu pour diffuser des données habitats même s'il permet d'afficher des couches sur les interfaces cartographiques. Le visualiseur de données sur les habitats devra aussi permettre de requêter sur les attributs des données sur les habitats. Il y a bien sûr un lien à faire entre les données espèces et les données habitats et donc au travers des différents outils proposés.

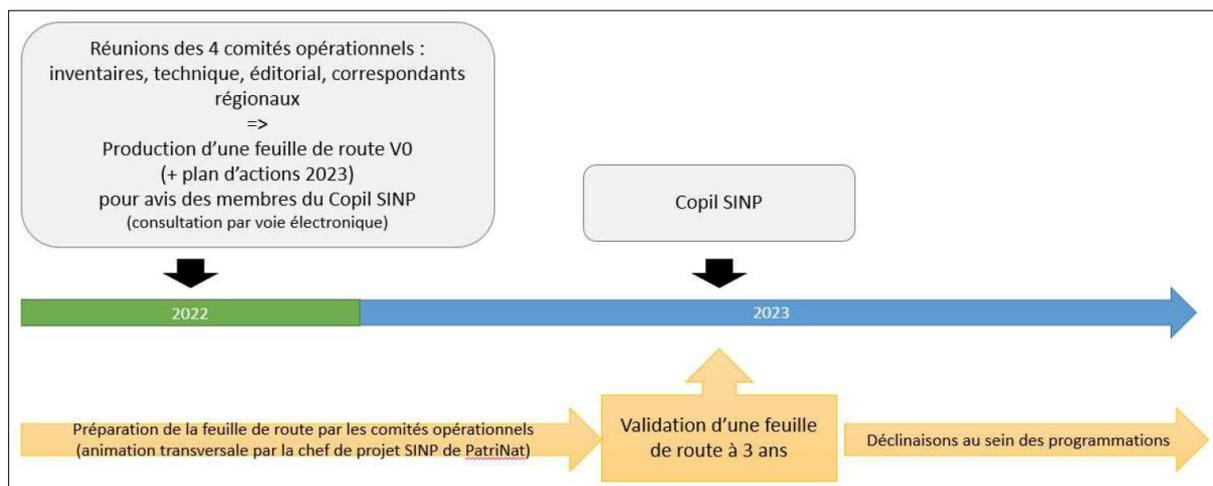
Vers une feuille de route à 3 ans

Cf. support de présentation

Présentation de la V0 de la feuille de route.

Objectif de produire une feuille de route consolidée soumise à la validation lors du prochain copil SINP.

Proposition d'ajouter un axe dans la feuille de route sur la mise à niveau de toutes les plateformes. Il y a un risque de grand écart entre les régions, la cohérence d'ensemble en pâtit. Priorité de mise en place d'un accompagnement spécifique pour les plateformes qui le nécessitent.



D'ici à fin 2022 : validation par mail des principales actions/objectifs de la feuille de route/recueil des propositions d'ajout, de modification, ... (document de travail commun)

Printemps 2023 : 2e réunion du comité technique – consolidation de la feuille de route en séance

Juin 2023 : présentation de la feuille de route en copil pour validation

Conclusion

Le défi initial du SINP consistant à mutualiser et échanger les données de manière normalisée est en passe d'être gagné. Il est maintenant nécessaire de travailler sur la qualité des données, leur mise à disposition rapide et précise, et de lancer le chantier des autres thématiques du SINP. Cela doit être au cœur de la feuille de route collective.