
Titre	Guide d'accompagnement pour la transmission des données d'observation sur les espèces dans le cadre de programmes nationaux (inventaire, atlas, réseau de surveillance, observatoire, etc.).
Créateur	UMS PatriNat
Version	1
Date	24 Mars 2020

Introduction

Les données d'observation (inventaires, suivis, dénombrements...) sur les espèces sont partagées et diffusées dans le cadre du Système d'Information sur les inventaires du patrimoine naturel (SINP).

Le présent document porte sur les données d'occurrences d'espèces.

Une donnée d'occurrence est l'observation d'un taxon (espèce, sous espèce, ...), à une localisation, une date, avec ses observateurs, par des méthodes d'observation directes (de visu) ou indirectes (empreintes, fèces, nid, ...). Ceci recouvre les données issues d'inventaire, avec ou sans protocole, de même que les suivis temporels et toute autre étude comportant le relevé de la présence ou de l'absence d'une espèce ou d'une communauté d'espèce.

L'occurrence se distingue de la synthèse par son caractère unitaire. Dans le cas de l'occurrence, une observation se traduit en une donnée là où la synthèse consiste en une agrégation de données et où une donnée est donc issue de plusieurs observations (exemple : la liste des espèces observées par commune, etc.).

Tout organisme ou personne peut contribuer au SINP en transmettant ses données d'observation.

Attention : les données d'étude d'impact et autres études réglementaires soumises à la procédure de dépôt légal de données brutes de biodiversité, en application des articles L411-1 A et L 122- 1 du code de l'environnement ne sont pas concernées par cette note. Ces données font en effet l'objet d'une procédure de dépôt spécifique. Cette obligation de dépôt concerne à la fois les études d'évaluation préalable à celles de suivi des impacts, réalisées dans le cadre de l'élaboration des projets d'aménagement ou des plans, schémas, programmes et autres documents de planification.

http://www.naturefrance.fr/sites/default/files/fichiers/ressources/pdf/18158_depobio_brochure_web.pdf

Les données déposées via ce service seront prises en charge dans un second temps par le SINP.

Le SINP

Le SINP est un dispositif partenarial entre le ministère chargé de l'environnement, les associations, les collectivités territoriales, les établissements publics et opérateurs, les services de l'État, etc. Il vise à favoriser une synergie entre ces acteurs pour la production, la gestion, le traitement, la valorisation et la diffusion des données géolocalisées relatives à la biodiversité.

Les acteurs du SINP

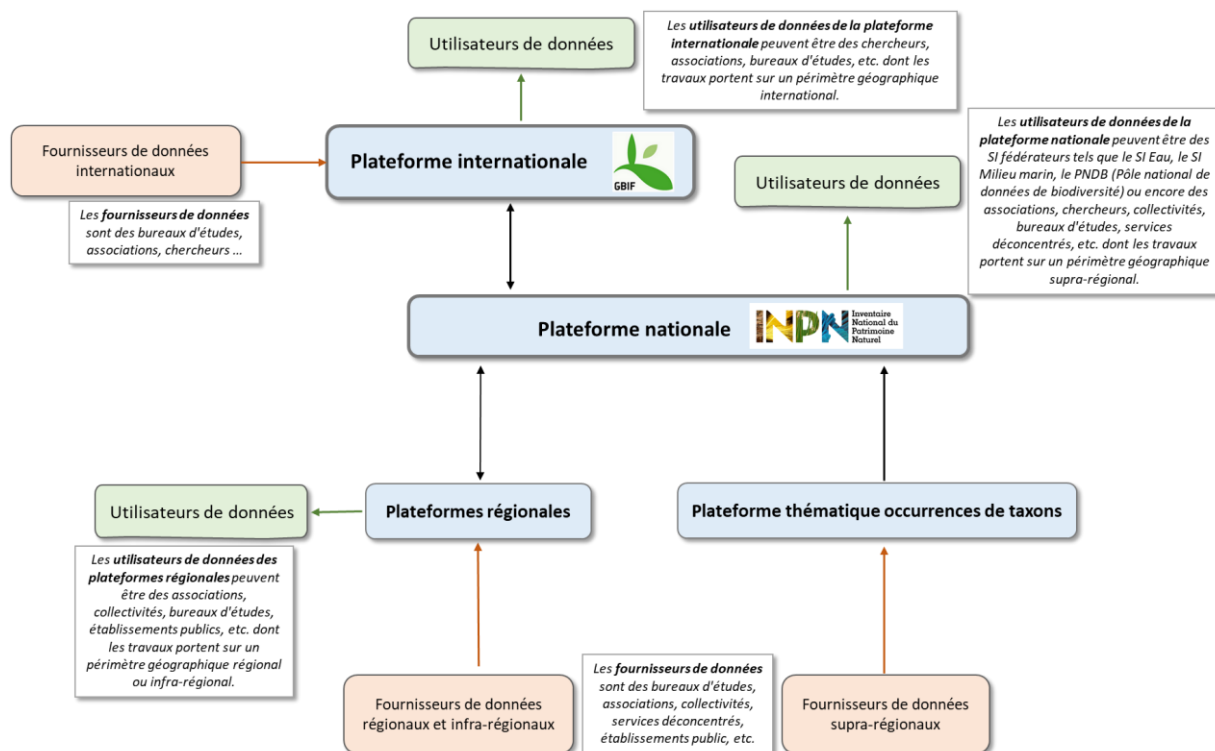
Le SINP privilégie une organisation en réseaux. Il repose sur des producteurs de données, des plateformes régionales, une plateforme thématique et une plateforme nationale¹.

L'animation du SINP au niveau régional est confiée aux **directions régionales de l'environnement, de l'aménagement et du logement en métropole (DREAL/DRIEE pour l'Île-de-France)** ou aux **directions en outre-mer chargées de la nature**, en coordination avec la région si cette dernière le souhaite et l'agence régionale de la biodiversité (ARB) lorsqu'elle existe, et avec l'appui des structures et réseaux de producteurs de données.

L'UMS PatriNat (OFB/CNRS/MNHN) est responsable de la plateforme thématique (d'envergure nationale) sur les données d'occurrence de taxon. PatriNat est également responsable de la plateforme nationale du SINP, en charge des liens avec le niveau régional et le niveau international (GBIF notamment).

Cette organisation permet de constituer un système global proposant une continuité entre les échelles locale et internationale. Elle permet également d'intégrer dans le SINP les données d'observation sur la France, diffusées au niveau du GBIF par des acteurs étrangers.

¹ Une plateforme est un dispositif habilité qui assure l'animation des réseaux d'acteurs et repose sur des outils permettant le partage des données selon les principes du protocole SINP ».



Comment contribuer, vers qui se tourner ?

On distinguera trois niveaux de production : le niveau régional, le niveau national et le niveau international.

POUR LES PRODUCTEURS DITS « RÉGIONAUX »

Sont considérés comme producteurs régionaux, les personnes ou structures dont l'aire géographique de prospection se limite essentiellement à une région administrative.

Les plateformes régionales habilitées constituent les points d'entrée à privilégier.

Pour plus de précision sur les plateformes régionales, vous pouvez consulter le site [Nature France](#) ou prendre contact avec le service environnement de la DREAL/DEAL/DRIEE.

Pour connaître la liste des régions échangeant leurs données avec la plateforme nationale (INPN), consultez la page suivante : <https://inpn.mnhn.fr/programme/donnees-observations-especes/contribuer>

En cas de doute sur l'interlocuteur à contacter, vous pouvez demander assistance à l'adresse suivante : sinp@mnhn.fr

POUR LES PRODUCTEURS DITS « NATIONAUX »

Sont considérés comme producteurs nationaux, les personnes ou structures dont l'aire géographique de prospection se situe sur plusieurs régions françaises.

La plateforme thématique portée par PatriNat est privilégiée comme point d'entrée dans le SINP.

Contact : sinp-data@mnhn.fr

POUR LES PRODUCTEURS DITS « INTERNATIONAUX »

Sont considérés comme producteurs internationaux, les personnes ou structures détenant ou produisant de la donnée sur une zone géographique plus large que la France (plusieurs pays).

Le GBIF France – point nodal français du réseau GBIF, est privilégié comme point d’entrée.

Contact : connexion@gbif.fr

Quelles informations faut-il transmettre?

Selon que la transmission des données se fait au niveau régional, national ou international, le standard (ou format) pour fournir des données peut varier, mais l’interopérabilité est assurée par une mise en correspondance des différents standards au sein du dispositif.

Dans tous les cas, il sera demandé de transmettre *a minima* :

- ➔ la date de l’observation ;
- ➔ le nom scientifique du taxon ;
- ➔ la localisation la plus précise disponible², c’est à dire si possible l’emplacement exact de l’observation (coordonnées GPS) ou à défaut la commune où a eu lieu l’observation;
- ➔ le nom de l’observateur ;

et si possible toute autre information que vous jugerez pertinente, en particulier celles habituellement utilisées pour le groupe taxonomique concerné (exemple : statut de reproduction, stade de vie, etc.) ainsi que les informations requises par le protocole de collecte d’informations que vous suivez.

Des métadonnées³ (informations décrivant le contexte d’acquisition des données – le pourquoi, pour qui et comment les données ont été produites) devront également être transmises.

Il sera demandé de transmettre *a minima* :

- ➔ le nom (libellé) du jeu de données/du cadre d’acquisition (compréhensible pour les futurs utilisateurs) ;
- ➔ une description du jeu de données/du cadre d’acquisition. La description doit être la plus précise possible notamment en ce qui concerne les dispositifs opérationnels mis en place pour l’acquisition des données et en ce qui concerne les finalités recherchées (un programme, un projet, une étude pour quoi faire ?). Dans la description, peuvent aussi apparaître des informations sur l’historique du programme ou du projet. Peuvent également être associé un ou des liens URL ;
- ➔ le protocole utilisé pour la collecte des informations (plan d’échantillonnage, fréquence, technique utilisée) ;
- ➔ les acteurs impliqués : producteurs, maître d’œuvre, maître d’ouvrage, financeur, ...

Selon le contexte de vos observations (et le programme d’acquisition), une organisation particulière/spécifique pour l’organisation des métadonnées peut avoir été définie. Il conviendra donc de se tourner vers le commanditaire ou porteur de projet pour connaître la structuration à respecter et la manière de renseigner les informations attendues.

² Pour vous localiser, vous pouvez utiliser votre GPS directement sur le terrain ou par exemple utiliser le site de l’IGN : geoportail.gouv.fr -> clic droit sur la carte>Adresse/coordonnées du lieu

³ Voir la rubrique « Que sont les métadonnées » ci-après.

De manière générale, un guide de préconisation national a été produit dans le cadre du SINP : <https://inpn.mnhn.fr/docs-web/docs/download/244790>

Que sont les métadonnées ?

Les métadonnées décrivent des ensembles de données ayant des caractéristiques communes.

On distingue deux niveaux de métadonnées pour décrire les données d'observation :

- **Le cadre d'acquisition**

Il décrit le contexte stratégique et opérationnel d'un programme ou d'un projet qui conduit à la production d'un ou de plusieurs jeux de données.

Les métadonnées du cadre d'acquisition renseignent notamment l'objectif du programme d'acquisition, les moyens mis en œuvre, le contexte d'acquisition et sur les acteurs impliqués (maître d'œuvre, maître d'ouvrage).

- **Le jeu de données**

C'est un regroupement de données qui se veut homogène, c'est-à-dire qu'un jeu de données contiendra des données ayant des caractéristiques communes, notamment le protocole d'acquisition utilisé (par exemple un même plan d'échantillonnage, les mêmes paramètres recueillis, etc.).

Les métadonnées du jeu de données renseigneront notamment le processus de collecte, les acteurs impliqués (producteurs, fournisseurs) et les modalités de diffusion⁴ de cet ensemble de données.

Pour les transmissions internationales : Le GBIF ne gère pas le niveau « cadre d'acquisition ». Seul le niveau « jeu de données » est demandé.

Les métadonnées, pourquoi faire ?

Les métadonnées permettent de disposer d'informations sur le niveau de cohérence des données entre elles mais aussi de mettre en avant la ou les personnes ou organismes ayant contribué à l'acquisition des données.

Elles permettent une meilleure réutilisation des données, cohérente avec l'usage recherché. Elles sont en effet essentielles pour bien percevoir le niveau d'homogénéité/d'hétérogénéité des données et déterminer si celles-ci sont des données protocolées ou dites « opportunistes », ces informations étant indispensables pour réaliser des analyses statistiques et des modélisations et interpréter scientifiquement les résultats.

La diffusion des métadonnées sous la forme d'un catalogue⁵ de métadonnées permet ainsi de faciliter la recherche des utilisateurs pour identifier les sources les plus pertinentes pour leurs besoins.

En outre, grâce à des métadonnées complètes, à jour et structurées, il est possible de publier un [Data Paper](#) pour valoriser une base de données issue d'un projet, ainsi que les personnels scientifiques et techniques liés à sa maintenance et à sa gestion.

⁴ Voir rubrique « Comment les données seront-elles diffusées ? »

⁵ <https://inpn.mnhn.fr/accueil/recherche-de-donnees/metadonnees>

Comment transmettre les observations ?

Pour la transmission des données *via* le niveau régional : chaque région propose une méthode d'alimentation qui lui est propre (standard, outils, etc.) et interopérable avec le standard national. Il conviendra donc de prendre directement contact avec la plateforme régionale pour préciser le format attendu.

Pour en savoir plus et transmettre vos données, vous pouvez consulter le site [Nature France](#) ou prendre contact avec le service environnement de la DREAL/DEAL/DRIEE.

En cas de doute sur l'interlocuteur à contacter, vous pouvez demander assistance à l'adresse suivante : sinp@mnhn.fr

Pour la transmission des données au niveau national :

Les données peuvent être transmises *via* :

- 1) l'utilisation de fichiers (ou masques de saisie) standardisés ;
- 2) la transmission de votre base de données, en ayant au préalable identifié la correspondance avec les champs standard du SINP ;
- 3) par saisie directe de vos données dans un outil dédié compatible et interfacé avec le SINP (voir annexe 1).

Des métadonnées⁶ (informations décrivant le contexte d'acquisition des données – le pourquoi, pour qui et comment les données ont été produites) doivent être associées à la transmission des données. Les métadonnées peuvent être transmises :

- 1) en utilisant des fichiers (ou masques de saisie) standardisés ;
- 2) en utilisant l'application nationale de [gestion de métadonnées dédiée](#)⁷ ;
- 3) *via* l'outil de saisie que vous utilisez s'il est interfacé avec le SINP (**voir annexe 1**).

Pour en savoir plus et transmettre vos données et métadonnées: sinp-data@mnhn.fr

Pour la transmission des données à l'international : le standard en vigueur est le [DarwinCore](#) (pour les données) et l'eml⁸ (pour les métadonnées).

Données et métadonnées peuvent être transmises :

- 1) en utilisant des fichiers standardisés ;
- 2) en utilisant l'outil de publication de données « IPT » du GBIF (ipt.gbif.fr) après contact avec l'équipe du GBIF.

Une présentation du processus de publication des données dans le réseau GBIF est accessible sur le site du GBIF France : <http://www.gbif.fr/page/contrib/connecter-des-donnees-au-gbif>

Pour en savoir plus et transmettre vos données et métadonnées : connexion@gbif.fr

⁶ Voir la rubrique « Que sont les métadonnées » ci-après.

⁷ Cela nécessite de vous ouvrir des droits d'accès : prendre contact avec sinp-data@mnhn.fr

⁸ Ecological Metadata Language

Quels sont les contrôles qui seront appliqués sur les données?

Les données seront soumises à un ensemble de **contrôles** (conformité des éléments transmis, cohérence des informations et validation scientifique).

Selon la plateforme d'entrée dans le SINP, les contrôles mis en place peuvent varier peu ou prou.

Pour connaître les contrôles mis en place au niveau de chaque région, il conviendra de se tourner vers les porteurs de plateforme régionale.

Les contrôles mis en place de manière générique par le niveau national sont publiés sur le site de l'INPN : <https://inpn.mnhn.fr/programme/donnees-observations-especes/references/validation>

Les contrôles mis en place au niveau international (GBIF) sont les suivants :

- Avant publication et indexation sur le site GBIF.org :
 - o Mise en conformité ou vérification de la conformité des données avec le standard d'échange Darwin Core
 - o Vérification de l'absence de doublon dans les identifiants de données
 - o Formatage des dates et des coordonnées géographiques pour correspondre aux normes en vigueur dans le réseau GBIF
 - o Vérification et validation des métadonnées
- Après publication et indexation sur le site GBIF.org :
 - o Détection des éventuelles erreurs de géoréférencement
 - o Validation des noms scientifiques sur la base de la liste taxonomique de référence du GBIF (construite à partir de listes de référence nationales (incluant TAXREF) et thématiques)
 - o Détection d'autres erreurs fréquentes : dates mal formatées ou absentes, valeurs mal formatées dans d'autres champs (altitude, profondeur, nombre de spécimens), url d'images invalides, etc.

Comment les données seront-elles diffusées ?

La réglementation fixe les règles de diffusion des données (convention d'Aarhus, Directive INSPIRE, Loi sur la reconquête de la Biodiversité et des Paysages, Loi République numérique et Code des relations entre le public et l'administration).

Les données du SINP sont publiques, diffusées gratuitement, en Open data par défaut.

Aux niveaux régional et national :

Le SINP propose une licence dite ouverte pour la diffusion des données : https://inpn.mnhn.fr/docs/demandes_sinp/Licence_ouverte_SINP.pdf

Les données d'**origine privée** (sans financement public pour leur acquisition) peuvent cependant être floutées géographiquement par la plateforme (à la demande du producteur) pour la diffusion en Open data.

L'INPN propose une charte aux producteurs de données permettant d'indiquer cette éventuelle limitation de diffusion⁹: <https://inpn.mnhn.fr/docs/charte-donnees-INPN.pdf>

Les plateformes régionales proposent également des chartes déclinant ces mêmes modalités.

Les données dites sensibles constituent une exception à ces modalités de diffusion. Un niveau de sensibilité sera affecté aux données transmises (ce niveau est établi sur la base d'un référentiel national). Ce niveau de sensibilité détermine le niveau de floutage géographique des données diffusées. Il est calculé par la plateforme à laquelle sont transmises les données.

Toutes les données précises du SINP, quelle que soit leur origine, sensibles ou non, peuvent être communiquées (selon la résolution la plus précise disponible dans le SINP) sur demande motivée. Les données sont alors « communiquées » sous licence dite fermée autorisant leur usage dans le cadre d'un projet défini mais interdisant la rediffusion publique ou à des tiers de ces données: https://inpn.mnhn.fr/docs/demandes_sinp/Licence_ferme_SINP.pdf

Attention: dans le cas où le producteur ne souhaiterait pas que ses données soient accessibles dans le SINP au plus précis (même dans le cas d'une diffusion sous licence fermée), alors les données devront être transmises préalablement floutées géographiquement à la plateforme SINP. Ce floutage préalable n'est possible que lorsque les données sont d'origine privée, c'est-à-dire qu'elles n'ont pas été financées dans le cadre d'un programme public ou acquises par un établissement sur mission de service public. Ce floutage des données n'est cependant pas recommandé dans le sens où cela signifie que les données précises ne pourront donc pas être prises en compte dans le cadre des politiques publiques de protection et de conservation de la nature, ni dans les travaux de recherche.

Au niveau international (GBIF) :

Les données seront diffusées librement et associées à des licences d'exploitation définies au moment du versement. Un choix doit être fait au niveau institutionnel parmi les trois licences suivantes :

<http://www.gbif.fr/page/contrib/les-licences-de-donnees-dans-le-reseau-gbif>

Toutes les données transmises par PatriNat au GBIF (floutées selon les règles de sensibilité et selon les autorisations des producteurs) sont diffusées sous la licence CC BY NC¹⁰.

Que sont les données sensibles ?

Les données sensibles constituent un cas particulier de restriction de la diffusion.

Les impacts d'aménagements sur la biodiversité résultent souvent d'une méconnaissance de la localisation du patrimoine naturel. Il est ainsi considéré d'intérêt général de communiquer sur les sites abritant des éléments remarquables et, plus généralement, de diffuser les données de biodiversité pour qu'elles puissent être utilisées dans les dispositifs de conservation, pour la recherche et l'information du public.

Les données sensibles constituent cependant une exception bien encadrée. La notion est définie par le protocole du SINP, en référence à l'article L.124-4 du code de l'environnement : **ce sont des données qui**

⁹ Ces informations sont à renseigner dans l'onglet « diffusion » de l'application métadonnées ou dans les fichiers de métadonnées transmis.

¹⁰ Licence sous laquelle les données sont mises à disposition pour toute utilisation sous condition de citation de la source des données et d'utilisation non commerciale.

ne doivent pas être largement diffusées pour éviter de porter atteinte aux éléments qu'elles concernent.

Un groupe de travail partenarial coordonné par le MNHN a proposé en 2014 les critères qui doivent être réunis simultanément pour définir une espèce comme potentiellement sensible:

- **Le risque d'atteinte volontaire.** Par exemple est-ce que l'espèce est sujette à des destructions volontaires ou à des prélèvements pour la consommation, la pharmacopée, la collection, l'horticulture..., ou est-ce que son attrait risque d'attirer de nombreux curieux ou naturalistes, au point de causer un piétinement de la station ou un dérangement de la population ;
- **La sensibilité intrinsèque de l'espèce.** Par exemple l'espèce est-elle menacée au sens de la Liste rouge, ou très localisée dans la région... A l'inverse, des prélèvements d'une espèce répandue dont les populations sont abondantes ne doivent pas être un motif de restriction de diffusion ;
- **L'effet probable de la diffusion de l'information.** Est-ce que la disponibilité de l'information dans le SINP augmenterait le risque d'impact sur l'espèce. Notamment dans le cas où on trouve déjà des informations dans la bibliographie et sur internet.

À ces critères sont associés des aspects de contexte s'appliquant à la donnée. Il s'agit de l'ancienneté de l'observation et du statut biologique associé. Par exemple, l'endroit précis de nidification de la Cigogne noire est considéré comme une donnée sensible, alors qu'une observation en migration n'est pas sensible.

Le floutage géographique n'est pas effectué par le producteur des données mais par le porteur du SINP pour les besoins de diffusion, sur la base du référentiel national. S'agissant d'espèces rares et/ou menacées, il est en effet primordial que l'information précise puisse servir pour éviter les impacts sans pour autant entraîner un risque accru sur les espèces.

Le référentiel national des données sensibles est une consolidation des référentiels de sensibilité établis par les acteurs du SINP dans chaque région et soumis à la validation du Conseil scientifique régional du patrimoine naturel (CSRPN).

Le référentiel de sensibilité est consultable en ligne : <https://inpn.mnhn.fr/programme/donnees-observations-especes/references/sensibilite>

Je transmets mes données à des partenaires, comment puis-je savoir si mes données contribuent au SINP ?

Vous pouvez consulter le catalogue national en ligne : <https://inpn.mnhn.fr/accueil/recherche-de-donnees/metadonnees>

Vous pouvez consulter le requêteur national OpenObs : <https://openobs.mnhn.fr/>

Vous pouvez demander directement aux partenaires à qui vous transmettez vos données.

Vous pouvez demander aux porteurs de plateforme.

Je transmets mes données à des partenaires, comment puis-je savoir si mes données contribuent au GBIF ?

Vous pouvez consulter la liste des fournisseurs de données et des jeux de données diffusés par la France dans le GBIF : https://www.gbif.org/dataset/search?publishing_country=FR et <https://www.gbif.org/publisher/search?country=FR>

Vous pouvez demander directement aux partenaires à qui vous transmettez vos données.

Vous pouvez demander à l'équipe du GBIF France : gbif@gbif.fr

À noter que toutes les données transmises au niveau régional ou national seront transmises au GBIF par la plateforme nationale (INPN) selon les modalités fixées par les producteurs *via* la charte INPN : <https://inpn.mnhn.fr/docs/charte-donnees-INPN.pdf>

Pour aller plus loin

Consulter les guides d'accompagnement pour la collecte, la saisie et la standardisation des données :

<https://inpn.mnhn.fr/programme/donnees-observations-especes/references/qualite>

Consulter les données déjà disponibles sur la plateforme nationale :

<https://openobs.mnhn.fr/>

<https://inpn.mnhn.fr/programme/donnees-observations-especes/decouvrir>

En savoir plus sur le SINP :

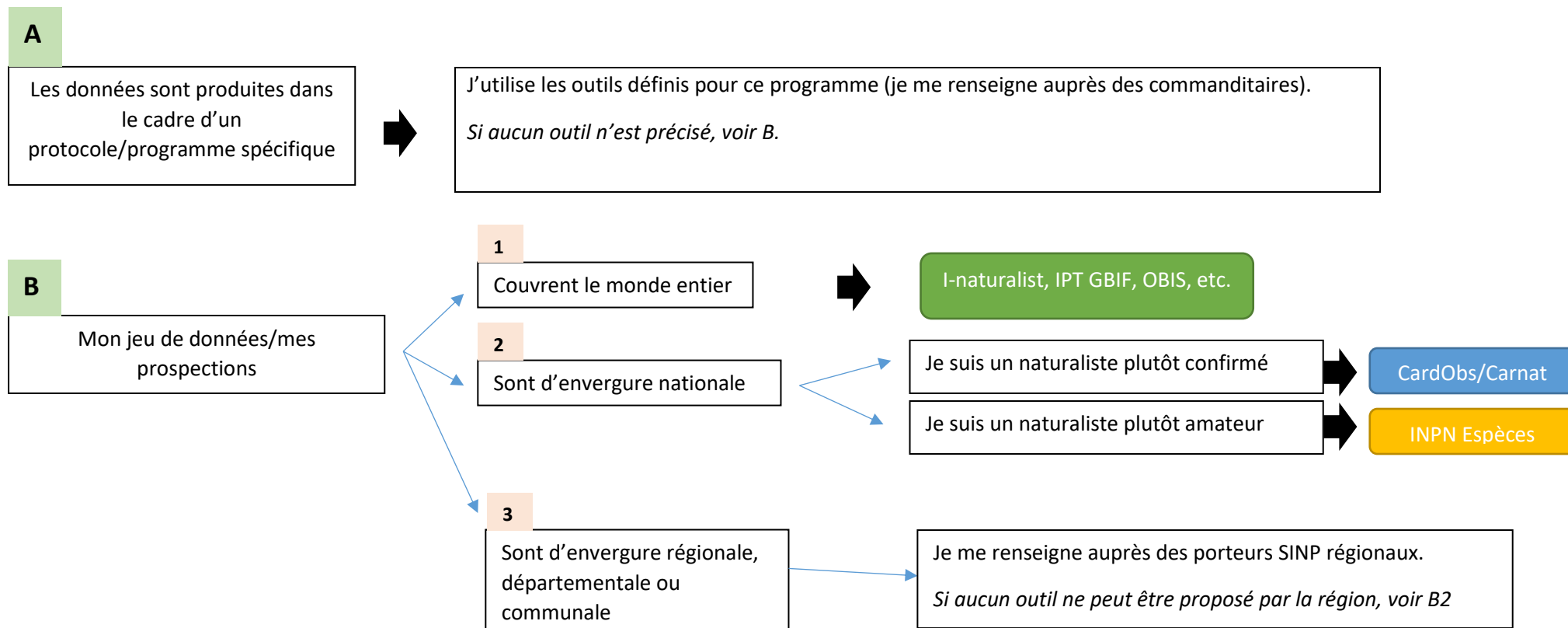
<http://www.naturefrance.fr/sinp/presentation-du-sinp>

<https://inpn.mnhn.fr/programme/donnees-observations-especes/presentation>

En savoir plus sur le GBIF : www.gbif.fr, www.gbif.org

Annexe 1

Aide pour le choix d'un outil de saisie pour partager les données¹¹



¹¹ À noter que vous pouvez utiliser des outils différents selon vos programmes/projets. Il faudra juste veiller à **ne pas saisir une même donnée sur plusieurs outils** afin d'éviter la génération de doublons dans le système.

Annexe 2

Liste (non exhaustive) d'outils de saisie gratuits proposés au niveau français pour la saisie et le partage des données naturalistes sur les espèces dans le cadre du SINP

Ne sont listés ici que des outils assurant le partage des données saisies selon les modalités définies pour le SINP et dont les standards sont compatibles avec le SINP. Ne sont pas listés des outils de portée régionale ou spécifiques à un groupe taxonomique.

Outil	Description	Emprise taxonomique	Emprise géographique
CarNat CardObs https://cardobs.mnhn.fr	Outil sur mobile associé à CardObs, pour naturaliste « confirmé ». Possibilité de contribuer à plusieurs sur un même jeu de données avec les paramètres de protocole associés. La détermination est directement proposée par l'observateur.	Tous taxons/espèces/groupes taxonomiques	France entière (interfacée avec les référentiels français). Possibilité de bancariser également les données hors France.
Oison https://oison.afbiodiversite.fr	Outil mobile destiné aux personnels OFB. La détermination est directement proposée par l'observateur.	Tous taxons/espèces/groupes taxonomiques	France
GeoNature https://geonature.fr/	Outil développé par les parcs nationaux de France. La détermination est directement proposée par l'observateur.	Tous taxons/espèces/groupes taxonomiques Permet également la saisie de données sur les habitats (module dédié)	France
INPN-Espèces https://inpn.mnhn.fr/informations/inpn-espèces	Outil fondé sur la photographie, recommandé pour les naturalistes débutants et pour les taxons qu'un expert maîtrise moins (exemple pour qu'un lépidoptériste puisse transmettre ses observations de coléoptères). La détermination est effectuée ou validée par un ou plusieurs experts de l'INPN.	Tous taxons/espèces/groupes taxonomiques	France
i-Naturalist https://www.inaturalist.org	Sur le même principe qu'INPN-espèce, mais permettant de saisir des photographies d'espèces observées dans le monde entier, France comprise. La détermination est validée par une majorité des deux-tiers d'utilisateurs d'iNaturalist ayant exprimé leur avis sur la photographie soumise. Fort aspect collaboratif.	Tous taxons/espèces/groupes taxonomiques	Mondiale

<p>Observatoires Vigie Nature</p> <p>http://www.vigienature.fr</p>	<p>Outil propre à chacun des programmes. Données standardisées contraintes par des protocoles définissant le mode et la stratégie de collecte.</p>	<p>Emprise et résolution taxonomique déterminée en fonction des programmes</p>	<p>France</p>
--	--	--	---------------