



Coordination scientifique et technique
Système d'Information sur la Nature et les Paysages
Observatoire National de la Biodiversité



Titre / Objet	Définition de l'architecture du SINP
Date de saisine	Date de réception au secrétariat de la CST : Historique des échanges : Saisine finale
Type de saisine	finale à un GT ou à une Action (ou Résultats) : résultats d'une action (document, organisation, outil, méthode, doctrine...) avant publication ou mise en œuvre dans le SINP ou l'ONB.

1. Contexte	<p>Le protocole du SINP prévoit que l'accessibilité des données sera facilitée depuis le portail national et les portails régionaux du SINP, qui représentent un outil de travail collaboratif pour les acteurs. L'intégration des données, bases de données ou informations connexes via le SINP soulève un certain nombre d'enjeux, en lien direct avec les systèmes d'information de chaque adhérent et les outils communs du SINP (outil de saisie, IDCNP, CARMEN). Ces enjeux mettent en avant la nécessité de construire un cadre général pour un Système d'Information cohérent du SINP, dans ses composantes organisationnelles, fonctionnelles et technologiques.</p>
2. Questions scientifiques ou techniques à étudier	<p>Question principale :</p> <p>Les résultats du groupe de travail «GT Architecture» apportent-ils une réponse claire sur les attentes et les besoins des acteurs et les principes d'architecture (doc IRSTEA/IFREMER de mai 2011) en particulier sur :</p> <ul style="list-style-type: none"> • La clarté de l'organisation et le rôle des acteurs notamment par rapport aux plate-formes régionales, thématiques et national ? Lesquelles ? Pourquoi ? Selon quels critères ? • L'architecture du SINP est-elle générique ? Indépendante des standards de données adoptés ? • L'architecture est-elle évolutive ? Pouvant fonctionner avec plusieurs technologies d'échange de données; • Est-elle ouverte vers d'autres SI ? (en région, au niveau national comme l'INPN ou le SiE, ou internationale comme le GBIF) ? • Le circuit de qualification de la donnée prévu dans l'architecture du SINP garantit-il la fiabilité scientifique et technique des données ? <p>Sous-questions :</p> <ul style="list-style-type: none"> • La qualité technique des documents finaux est-elle satisfaisante ? si non, quelles sont les améliorations à apporter ? • Quelles recommandations la CST peut-elle émettre en termes de mise en œuvre de l'architecture ? • Quelles sont les fonctions du SINP non traitées dans les livrables d'architecture qui doivent être traitées en priorité ? • Aspect technique et prototype : compte tenu des technologies disponibles actuellement, dans quel mesure les solutions proposées notamment celles présentées dans le prototype sont opérationnelles à court terme ou à moyen terme ?

Rapport final de la CST concernant la saisine sur la définition de l'architecture du SINP.

Rapporteur : Pascal Dupont.

Contributeur : Frédéric Gosselin.

Rapport prenant en compte les discussions et propositions émises au cours de la séance plénière du 29 octobre 2013.

I. Préambule.

L'organisation proposée répond aux attentes des différents acteurs du SINP mais plusieurs points devront être finalisés. Certaines fonctionnalités importantes de l'architecture n'ont pas pu être examinées. Les deux points les plus importants sont la traçabilité de l'utilisation et la valorisation scientifique des données qui n'est pas précisée. Ce point est important car c'est l'un des facteurs de pérennisation de l'adhésion du producteur de données au SINP. Le deuxième point est celui de la qualification des données.

Ces deux points non abordés dans le document sont encore à l'étude par deux groupes de travail. C'est l'une des difficultés fonctionnelles de l'évaluation du projet SINP : un groupe de travail peut être bloqué, car les réponses attendues dépendent des résultats d'autres groupes de travail qui n'ont pas avancé ou ne sont pas encore créés. Il serait important de clairement prioriser et planifier les travaux du SINP pour plus de visibilité.

II. Définition et organisation des plates-formes régionales et thématiques.

Analyse des documents : Document d'architecture ; Relevé des exigences.

L'organisation proposée répond globalement aux attentes des différents acteurs. Plusieurs points importants sont présents :

- L'architecture est structurée en plusieurs niveaux : producteurs de données, plates-formes régionales, plates-formes thématiques et plate-forme nationale. Les plates-formes régionales et thématiques sont les éléments clés de l'architecture. Elles peuvent développer des outils communs. Elles sont responsables de la gestion des données et de leur conformité avec le Protocole du SINP. Le rôle principal de la plate-forme nationale est de mettre à disposition des référentiels techniques et sémantiques afin de faciliter l'interopérabilité.
- Les plates-formes régionales, thématiques et nationales assurent une qualification des données à leur niveau. Elles assurent aussi leur diffusion associée à leur niveau de qualification. Une donnée diffusée ne peut l'être sans les métadonnées attachées à cette donnée.
- Le flux de données et de métadonnées est organisé de la façon suivante :
 - o Du producteur aux plates-formes régionales et thématiques (données non qualifiées).
 - o Des plates-formes régionales et thématiques aux producteurs (données qualifiées aux niveaux régional ou thématique, métadonnées (dont mises à jours))

- Des plates-formes régionales et thématiques à la plate-forme nationale (données qualifiées au niveau régional ou thématique, métadonnées (dont mises à jours))
- De la plate-forme nationale aux plates-formes régionales et thématiques (données qualifiées au niveau national, métadonnées (dont mises à jours))
- Des plates-formes régionales et thématiques et de la plate-forme nationale vers les utilisateurs via, à chaque niveau, une plateforme spécifique de médiation (portail). L'utilisation des données par les utilisateurs est cadrée par un paramétrage des droits d'accès et des droits d'usage.

La notion de plate-forme thématique à besoin d'une définition plus précise. Il y a une confusion entre la constitution éventuelle d'une plate-forme thématique et le fait d'être producteur national ou en compétence pour expertiser et valoriser les données. C'est de cas notamment de structures nationales organisées en réseau comme Réserve Naturelle de France. Les incompréhensions sémantiques sont possibles entre plates-formes thématiques nationales, plates-formes thématiques régionales et plates-formes régionales généralistes. Le principe proposé pour la définition d'une plate-forme thématique est que le thème associé (en priorité ceux d'INSPIRE), correspond à un standard spécifique de données. Par ailleurs, il faut éviter de multiplier le nombre de plates-formes pour assurer une meilleure visibilité de l'architecture. De plus une architecture trop complexe aurait sans doute des répercussions sur sa performance concernant la gestion des flux de données.

III. Rôle des acteurs aux plates-formes régionales, thématiques et nationales.

Analyse des documents : Document d'architecture ; Relevé des exigences.

Le document principal fournit peu d'informations, simplement un organigramme simplifié du rôle des acteurs. Pour mieux présenter le rôle des différents acteurs du SINP vis-à-vis du cheminement de la donnée, une présentation des parcours possibles d'une donnée en fonction de divers scénarii (données sources d'un producteur national ; donnée source d'un producteur régional ; ...) apporterait des éclaircissements.

IV. Aspect générique de l'architecture, indépendance vis-à-vis des standards de données.

Analyse du document : Document d'architecture.

Dans la mesure où le groupe de travail « Standard de données » travaille toujours sur le sujet et que nous n'avons pas d'éléments sur les exigences pouvant être des contraintes pour l'architecture, il est difficile de se prononcer. Cependant, aux vues des caractéristiques techniques de l'architecture présentée celle-ci paraît générique.

V. Aspect évolutif de l'architecture et possibilité de fonctionnement avec plusieurs technologies d'échange de données.

Analyse du document : Document d'architecture.

Le choix est porté sur une plate-forme de médiation pour les flux de données qui peut être implémentée sur la base de langage XML ou JMS facilitant l'interopérabilité. L'architecture globale n'est pas liée à un choix technique unique mais à un système réparti qui repose sur les infrastructures propres de chacun des partenaires. Les partenaires sont libres des choix techniques au niveau de leur propre architecture mais ceux-ci devront développer des interfaces afin de d'assurer une traçabilité des données et des métadonnées dans le cadre des flux. Il faut s'assurer que les choix techniques de certains partenaires ne soient pas liés à une contrainte évolutive trop forte pour l'évolution de l'architecture globale.

VI. Ouverture vers d'autres Systèmes d'information.

Analyse des documents : Document d'architecture ; Relevé des exigences ; Prototype SINP.

Le prototype présenté semble montrer que l'architecture proposée est ouverte vers d'autres systèmes d'information.

VII. Qualification et fiabilité scientifique et technique des données.

Analyse du document : Document d'architecture.

La qualification des données est très peu abordée. Les processus sont à préciser. Du point de vue sémantique, le terme de qualification et ce qu'il contient (validation, précision géographique) doit être mieux défini. Le circuit de la qualification de la donnée ne va pas garantir la fiabilité scientifique et technique comme cela semble implicite dans le document. La fiabilité de la donnée va dépendre des processus endogène et/ou exogène associés à cette qualification.

VIII. Qualité techniques des documents finaux.

Analyse des documents : Document d'architecture ; Prototype SINP ; Etat de l'art.

Les documents étudiés associés à la saisine sont globalement de bonne qualité, particulièrement l'état de l'art qui couvre ce qui a été fait à l'échelle européenne. Le document présentant l'architecture pose problème pour sa diffusion en l'état car il est très difficile d'accès pour un public non averti. Une refonte est prévue afin d'améliorer le contenu dans cette voie.

IX. Opérationnalité du prototype à cours et moyen terme.

Analyse des documents : Prototype SINP.

Dans le prototype présenté, une problématique technique liée à la volumétrie des données requêtes est posée. Ceci devra être résolu techniquement si possible car les requêtes des différents acteurs du SINP concerneront de gros volumes de données.

X. Fonctions du SINP non traitées.

Analyse du document : Document d'architecture.

Il paraît indispensable de développer :

- F05 : Valoriser le producteur (hautement prioritaire. Ne pas mettre en œuvre l'architecture si cette fonction n'est pas développée en particulier la traçabilité des flux de données et de leur utilisation).
- F03 : Créer des normes d'interopérabilité (prioritaire).
- F06 : Qualifier les données (prioritaire).
- F02 : Partager l'information. La gestion du partage de la connaissance scientifique sur les espèces n'est pas abordée. Ceci pourrait être réalisé au niveau de la plate-forme nationale en liaison avec une base de données bibliographique. De plus, un focus plus important pourra être donné sur la valorisation scientifique des données du SINP.

XI. Recommandation de la CST.

L'intégration des autres structures nationales, régionales, interrégionales, infrarégionales (parc régionaux, département) au sein de l'architecture du SINP se pose. En effet, les données du SINP vont alimenter les indicateurs de l'ensemble des observatoires territoriaux. La réalisation d'une phase de test des flux de données entre les niveaux national et régional, est préconisée avant la mise en œuvre de l'architecture. Les tests révéleront l'existence éventuelle de blocages entre les différents niveaux. Ils permettront ainsi une amélioration de l'architecture et de sa cohérence.

Pour les régions n'ayant pas encore mis en œuvre la déclinaison du SINP, il est important d'envisager les outils pour favoriser les échanges et les appuis techniques afin que l'architecture soit opérationnelle à court terme à l'échelle de l'ensemble du territoire national.