Des protocoles locaux aux échanges de données nationaux

Pour le gestionnaire d'un territoire, l'organisation des données doit répondre à son rôle de gestionnaire. Les besoins sont donc très hétérogènes et spécifiques, parfois complexes. Dans le même temps, le besoin de mutualisation, de partage et d'agglomération des données entre structures ou à l'échelle nationale augmente. Il devient une nécessité voir une obligation.

Répondre à ces deux besoins très différents, simultanément et à partir des mêmes données, peut parfois paraître contradictoire. Pourtant la mise en place d'une organisation spécifique des données peut permettre de répondre à l'ensemble de ces besoins.

Partons d'un constat simple : les besoins des uns et des autres en matière de production de données sont généralement différents.

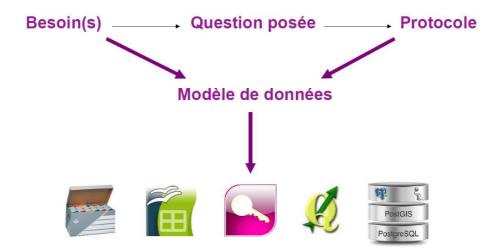
Que ce soit pour la faune, la flore ou d'autres sujets, nous avons tous une très bonne raison de produire et d'organiser nos données comme bon nous semble puisque l'objectif est de répondre à notre **BESOIN**.

Faire l'inventaire des espèces d'un territoire, suivre la reproduction d'une espèce, évaluer l'impact d'une mesure de gestion, ... Ces actions visent à **REPONDRE A UNE QUESTION** et donne généralement lieu à la mise en place d'un **PROTOCOLE**.

Pour organiser ses données, quelque soit ses compétences ou la technologie retenue, chacun va structurer ses données et donc mettre en place un MODELE DE DONNEES.

Ce modèle de données va permettre de générer une BASE DE DONNEES dont le but est de répondre à la question posée.

Quelque soit l'**OUTIL** retenu, selon les besoins, les moyens et les compétences, cet outil doit permettre de SAISIR, STOCKER et UTILISER les données produites.



Base de données

Partant de ce constat, il est donc illusoire de croire qu'il va être possible de créer une méga ou mieux une giga base de données qui va répondre au besoin spécifique de chacun! Pour faire ça, il faudrait que toutes les personnes concernées se mettent d'accord, puissent le financer, que tous les acteurs disposent des compétences nécessaires pour utiliser le même outils de base de données ou qu'une super interface permette à chacun de saisir ses données puis d'y accéder. Les animateurs de réseau et les développeurs capables d'une telle prouesse sont loin d'être identifiés et même s'ils l'étaient, il restera toujours des réticents ou des sceptiques pour ne pas adhérer à la démarche. Et puis, une giga base de données, oui, mais pour quel(s) besoin(s)?

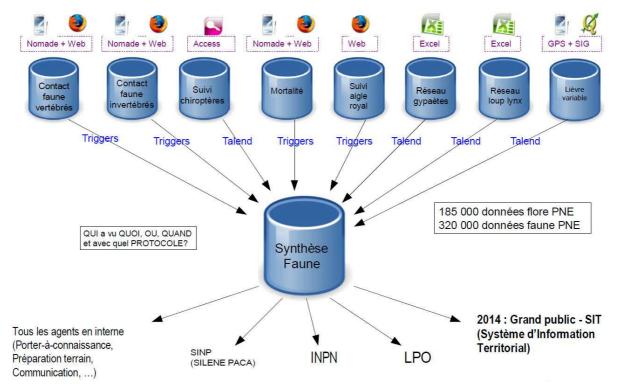
Il est donc préférable de faire ce constat simple :

UN BESOIN = UN PROTOCOLE = UN MODELE DE DONNEES = UN OUTIL.

OUI MAIS...

Oui, oui ! Mais comment on fait pour échanger / agglomérer des données si elles sont hétérogènes ?

Et bien, au Parc national des Ecrins, on fait comme ça :



Les données des différents protocoles sont saisies avec des outils différents. Il peut s'agir d'un simple tableur ou couche SIG pour les besoins simples concernant peu d'utilisateurs comme d'une base Access plus ou moins élaborée ou encore d'une base de données PostGIS accompagnée d'une interface web lorsque les utilisateurs sont nombreux. Certaines données sont même saisies directement sur le terrain grâce aux applications et aux outils nomades. Les données sont stockées par protocole, dans des schémas différents. Chaque schéma possède un modèle de données correspondant strictement au protocole. Il est structuré pour répondre aux besoins spécifiques de ce protocole. On respecte bien ici le principe UN BESOIN = UN PROTOCOLE = UN MODELE DE DONNEES = UN OUTIL.

Grâce aux REFERENTIELS tels que le taxref pour la taxonomie ou encore les référentiels géographiques de l'IGN, les informations communes à tous les protocoles peuvent être regroupées dans un schéma de SYNTHESE. En résumé qui a vu quoi, quand, où et comment (le protocole) ? Ce schéma de synthèse est automatiquement alimenté par des déclencheurs (triggers) au sein de la base de données ou périodiquement grâce à un outil ETL (Extract Transform and Load) tel que Talend Open Studio pour les données saisies avec d'autres outils tels que des bases de données fíchiers (Access) ou des tableurs.

Le schéma de chacun des protocoles répond donc au besoin du protocole et le schéma de synthèse qui regroupe toutes les données produites répond lui aux besoins d'échange et d'analyse des données ainsi qu'au besoin de porter à connaissance. Une vue spécifique est mise en place sur la base de données de synthèse pour chaque organisme partenaire (SINP, LPO, INPN...). Elles leur permettent d'extraire les données en temps réel en totale autonomie.

Début 2014, le SINP a produit un STANDARD des occurrences de taxons décrivant la donnée élémentaire d'échange (DEE) en matière de données Faune et Flore. Ce standard facilite les échanges de données et devrait à terme constituer la norme en matière d'échange de données naturalistes. Il est important de noter que ce standard ne constitue pas une référence sur la bonne manière de structurer ses données ; il précise ce que doit comporter à minima une données faune flore et comment la partager. On y retrouve le qui, quand, quoi, où et le comment (le protocole) de notre synthèse.

Ce qui donne :

