

## Comité de pilotage GBIF France

### Compte-rendu de la réunion du 8 décembre 2011

#### Présents :

Nom	Institution	Courriel
Didier Hoffschir	MESR	didier.hoffschir@recherche.gouv.fr
Yannick Lebeau	MEDDTL	Yannick.Lebeau@developpement-durable.gouv.fr
Marcel Jouve	MAEE	Marcel.JOUVE@diplomatie.gouv.f
Michel Guiraud	MNHN	mguiraud@mnhn.fr
Catherine Borremans	IFREMER	Catherine.Borremans@ifremer.fr
Jean-Philippe Siblet	SPN, MNHN	siblet@mnhn.fr
Philippe Feldmann	CIRAD	philippe.feldmann@cirad.fr
Maité Delmas	DREI, MNHN	delmas@mnhn.fr
Jean-François Silvain	IRD	Jean-Francois.Silvain@legs.cnrs-gif.fr
Régis Hocdé	SOES, IRD	regis.hocde@ird.fr
Charlotte Leroy	MNHN	charlotte@mnhn.fr
Régine Vignes	UPMC, GBIF	regine.vignes_lebbe@upmc.fr
Eric Chenin	IRD, GBIF	Eric.Chenin@ird.fr
Anne-Sophie Archambeau	GBIF	archambeau@gbif.fr

Didier Hoffschir ouvre la séance en indiquant l'objectif de faire un bilan après 5 ans d'existence du point nodal et de voir ce que l'on va faire à l'avenir.

**Rappel sur le GBIF :** il s'agit d'un réseau international avec des points nodaux nationaux et thématiques. Le mandat d'un point nodal national est de recenser les sources de données primaires, aider les fournisseurs à valoriser leurs données, organiser des formations, avec notamment un appui au Sud, informer sur le GBIF, promouvoir et faciliter l'utilisation des données, collaborer avec les autres dispositifs nationaux et internationaux, contribuer à des développements informatiques.

L'équipe du point nodal français est une petite équipe coordonnée par Eric Chenin (IRD) et Régine Vignes (UPMC). Elle est désormais rattachée à la direction des collections du MNHN et sa mission intégrée dans le contrat d'établissement du MNHN.

Didier Hoffschir soulève cependant la question de l'opportunité d'une future structure dédiée au GBIF : quel type de structure ? créer une UMS MNHN-IRD-UPMC ? A discuter avec les institutions impliquées.

Les 3 ministères représentés et les autres membres du Copil présents considèrent que le GBIF est utile et doit être soutenu.

#### **Présentation du bilan par E. Chenin (document de bilan d'activité du point nodal diffusé précédemment aux membres du Copil) :**

E. Chenin insiste sur la diversité des données sur la biodiversité, ainsi que des utilisateurs et

producteurs de données. Une analyse des usages a été réalisée dans le cadre de ALA (Atlas of Living Australia).

#### Discussion sur la validation des données :

M. Jouve pose la question de la validation des données, notamment dans la perspective de l'IPBES ?

L'IPBES a besoin du GBIF et le GBIF doit donc imposer ses standards.

E. Chenin : c'est le détenteur des données qui valide. Le secrétariat du GBIF demande aux points nodaux de cautionner chaque nouvelle base de données. Les points nodaux belges et anglais développent des outils de validation. Ces validations « automatiques » concernent la validation des noms taxonomiques (par rapport aux référentiels adoptés par le GBIF) et les coordonnées géographiques.

J.-F. Silvain : pour certains groupes taxonomiques, il existe un gros problème de référentiel non mis à jour. Importance de l'actualisation de ces référentiels.

J.-P. Sibley : La validation est très importante et nécessaire mais il faut également avoir un grand nombre de données disponibles.

R. Hocdé : importance de l'adaptation des données à leur utilisation (« fitness for use »).

D. Hoffschir : Importance de bien former les collecteurs.

GBIF : concernant les données sur le territoire français, le SPN effectue ce travail de validation dans le cadre de l'INPN. Un travail est en cours entre le GBIF France et le SPN sur les flux de données pour permettre au GBIF de donner un accès aux données de l'INPN même avant validation par le SPN en ajoutant aux données un indicateur de niveau de validation. Un document à destination des fournisseurs de données est en cours de rédaction pour préciser ce fonctionnement.

Forte discussion sur la qualité des données et leur complétude temporelle et spatiale. L'utilité du portail GBIF pour les recherches et la gestion des données est bien mise en évidence.

#### Connexion de données au portail GBIF :

La France est le 5ème pays contributeur avec plus de 14 millions de données. (A.-S. Archambeau : en fait la France en fournit plus que cela via les programmes thématiques internationaux tels que OBIS, Fishbase etc.).

Au niveau international, un total de 312 millions d'occurrences d'espèces est accessible en décembre 2011, fourni par 342 institutions comprenant 8548 bases de données provenant des 57 pays et 47 autres membres participants.

Présence de fournisseurs du sud ; une dizaine de pays africains sont membres (mais fournissent peu). Importance du Mexique et du Costa Rica.

D. Hoffschir : Créer autour du GBIF un réseau « francophile » de données.

### Quelles sont les utilisations en France du portail GBIF ?

E. Chenin présente les statistiques d'utilisation du portail du GBIF : plus de 37000 connexions au portail par an en provenance d'ordinateurs situés en France ont été enregistrées, ce qui fait de la France le 5<sup>ème</sup> pays consultant le portail GBIF. Le portail est donc déjà très utilisé en France, mais un travail complémentaire d'analyse des utilisateurs serait à faire pour préciser d'avantage les utilisations.

R. Vignes mentionne une étude réalisée sur la Nouvelle Calédonie : « *Etude de la répartition spatiale des données GBIF en Nouvelle-Calédonie* ». Cette étude réalisée par Gaspard Charrier (sous la direction d'E. Guilbert et P. Grandcolas au MNHN) montre l'intérêt et l'utilité du portail GBIF pour accéder aux données détenues par des institutions diverses sur une même zone géographique, et analyse les biais que l'on peut rencontrer sur la répartition des données (regard critique éclairant l'interprétation des données après une analyse scientifique).

### Indexation :

Il est rappelé qu'une décentralisation de l'indexation des données par le secrétariat du GBIF vers les points nodaux nationaux a été décidée il y a deux ans. La France travaille sur un outil d'indexation des données françaises. C'est un projet pilote du « GBIF lab » qui pourra ensuite être utilisé par d'autres points nodaux.

### Métadonnées :

L'interface de saisie des métadonnées a été refondue.

Au niveau français, le SINP s'occupe des métadonnées observations et le point nodal GBIF France des métadonnées collections.

Concernant les interactions avec le SINP, le GBIF France participe aux groupes de travail du MEDDTL : groupe « métadonnées », mais aussi « architecture des données », « aspects juridiques », « standards des données ».

M. Guiraud : Existe-t-il un lien avec l'OCIM (Office de Coopération Inter-Musées) ?

A.-S. Archambeau contactera l'OCIM pour discuter avec eux afin d'éviter les duplicats.

Les participants ont longuement discuté de la question des conséquences de l'application du protocole de Nagoya sur l'accès aux ressources et le recueil de données dans les régions d'outre-mer françaises et des pays étrangers.

R. Hocdé : Les données acquises par le GOPS sont d'abord gérées localement puis transmises sur la métropole.

J.-F. Silvain : Est-ce que le GBIF ne peut pas constituer un bon exemple pour illustrer les effets en retour de l'application du protocole de Nagoya dans le domaine de la recherche (le GBIF gère des données « universalisées », c'est son but et sa justification ; est-ce que cela sera encore possible à l'avenir) ?

Le MESR rappelle son souci de voir être mise en place une procédure aussi simple que possible.

D. Hoffschir : Il faut retravailler avec les juristes.

P. Feldmann : voir le document sur les lignes directrices du transfert des données génétiques mis en ligne par CIRAD-IRD-INRA.

### Appui au Sud :

Le projet SEP – CEPDEC (Sud expert plante et GBIF) s’est terminé en 2011. Il subsiste des actions ponctuelles de mentorat avec le Togo et la RCA (mentorat également sud-sud avec le Cameroun). Le projet SEP2 (suite de Sud Expert Plantes) est actuellement en cours de montage par le MAEE avec le FFEM, l'AFD et l'AIRD.

### Projet “OpenUp !” Opening up the Natural History Heritage for Europeana (open-up.eu) :

Ce projet Européen a pour but de contribuer au site web culturel européen EUROPEANA (<http://www.europeana.eu>) en mettant à disposition des ressources multimédia sur la biodiversité via les infrastructures et les standards GBIF.

Le GBIF France (avec pour institution légale le MNHN) fait partie du projet en tant que partenaire participant à l'aide technique (standards, métadonnées, infrastructure) et à l'identification des institutions possédant ce genre de ressources en France (sonothèques, banques d'images...). Il est également membre du groupe ODG (Outreach and Dissemination Group).

Le projet a démarré en Mars 2011 pour une durée de 3 ans et le consortium est constitué de 23 institutions scientifiques provenant de 12 pays européens.

### Governing Board 18 :

Il y a pour l'année 2012 une forte réduction des contributions de plusieurs pays, dont le Japon qui était jusque-là le plus gros contributeur. Le Canada s’est retiré des pays votants pour 2012, il reste présent en tant que pays associé. Le budget 2012 du GBIF a été réduit ce fait de 4,5 M€ à 3,6 M€, ce qui implique une révision des priorités, actuellement en cours de discussion.

La Chine et le Brésil ne sont pas encore partenaires du GBIF.

Claude-Anne Gauthier a été élue vice-présidente du Governing Board : à ce titre, elle est aussi vice-présidente de l’Executive Committee et membre du Science Committee ;

Jean Ganglo (Bénin) a été élu vice-président du Science Committee. Il s’est impliqué dans le GBIF grâce à SEP, et sa candidature a été fortement soutenue par l'Afrique du Sud ; c'est le premier africain, hors Afrique du Sud, dans un comité du GBIF.

### **Priorités pour 2012 :**

Analyse du contenu et des usages, articulation avec les autres structures (IPBES et GEOBON).

Analyse des besoins : Quelles stratégies d’acquisition et de connexion des données ? Quelles utilisations (« data use cases ») ? Quelles granularités d’analyse ?

Le secrétariat a lancé une étude sur les publications mentionnant et/ou utilisant le GBIF.

J.-P. Sibley : Peut-on, faut-il, faire des choix géographiques et taxonomiques ? Importance de la question de la complétude des données. Quelles sont les données potentiellement manquantes ?

Comment évaluer l’utilisation du GBIF par la communauté française ? A quelles questions le GBIF peut-il répondre ? Quelles aides apporte ce portail pour la gestion du territoire national ?

M. Guiraud : Le GBIF contribue à la valorisation des collections qui, elles, ont un coût récurrent très élevé.

P. Feldmann : par rapport aux coûts des collections, le coût de la valorisation dans le GBIF est faible et le GBIF est très utile.

En conclusion, une analyse de l'utilisation du portail est à faire. D. Hoffschir suggère de mettre en place un groupe de travail sous l'égide de la FRB et du MNHN.

Autres actions à mener :

- poursuivre le recensement des sources de données (et la saisie des Métadonnées).  
- poursuivre la collaboration avec l'INPN pour identifier les tâches et outils d'intérêt commun, expliciter le flux de données au niveau national et les procédures de validation des données. Il faut enrichir mutuellement les deux dispositifs.

J.-P. Sibley : L'INPN fournit des données de synthèse et transmet les données au GBIF. L'INPN a un référentiel taxonomique qui progresse et qui aide à la validation. Il faut être réactif, mais aussi stable que possible à ce niveau. Travail avec Télabotanica pour le référentiel botanique. Il faut que les données viennent directement à l'INPN pour les observations.

R. Hocdé : Il faut une traçabilité des données et donc que le nom de la structure qui fournit la donnée apparaisse.

J.-F. Silvain : Il faudrait un vademecum, guide de bonne conduite, à destination des chercheurs français, concernant la manière d'alimenter les bases de données (GBIF et INPN), car il persiste une grande méconnaissance des différents systèmes existants. Le GBIF et l'INPN travaillent conjointement sur ces différents points pour produire un document commun.

### **Budget :**

L'ensemble des participants considère que le GBIF (France) doit se poursuivre, mais pour le moment il n'y a pas de visibilité budgétaire après 2012.

D. Hoffschir mentionne la protection du budget 2012 pour la France, mais souligne la difficulté qu'il y aura à financer le GBIF France en 2013. Il revient sur le souhait de créer une structure propre pour le point nodal GBIF (UMS ? voir ci-dessus).

Il faut intégrer le point nodal dans la politique des organismes tout en tenant compte du fait que le GBIF est aujourd'hui intégré dans le contrat d'établissement du MNHN.

M. Guiraud : e-Recolnat (déposé dans l'appel d'offre infrastructures d'excellence) concerne la numérisation des collections naturalistes et présente un lien évident avec le GBIF. Si le projet est financé, il pourrait inclure après coup le GBIF.

### **Questions diverses :**

Si l'IRD soutient le GBIF, il y a des points à corriger (IRD, Régis Hocdé) :

- améliorer la visibilité institutionnelle
- faire attention de bien parler du GBIF comme d'un portail d'accès car il n'y a pas de données « GBIF »
- valoriser l'utilisation des contenus (publications scientifiques)
- faire attention à la multiplication des méta-structures et des portails au détriment de l'acquisition des données

Certains points sont déjà en cours d'amélioration et leur mise en place en cours de réalisation.