



Accord sur la conservation des albatros et des pétrels

Cinquième réunion du Comité consultatif

Mar del Plata, Argentine, 13 – 17 avril 2010

Politique de données de l'ACAP

Auteur : Secrétariat

Le présent document est présenté pour examen par l'ACAP et il est possible qu'il contienne des données, des analyses et/ou des conclusions non publiées et susceptibles d'être modifiées. Les données contenues dans le présent document ne doivent pas être citées ou utilisées à des fins autres que les travaux du Secrétariat de l'ACAP, du Comité consultatif de l'ACAP ou de leurs groupes de travail auxiliaires, sans l'autorisation des propriétaires des données originales.

Politique de données de l'ACAP

Objectif

Le présent document propose une politique de l'ACAP en matière d'utilisation et de réutilisation des données transmises à l'ACAP. Il contient une série de recommandations sur la manière dont le Secrétariat devrait mettre en œuvre cette politique pour s'assurer que les données sont bien gérées et leur paternité correctement attribuée, et que les propriétaires des données stipulent la manière dont leurs données doivent être utilisées.

Contexte

Aux termes des articles III (1) d) et g), V a) et f) et VI (2) e) de l'Accord, les Parties sont tenues d'élaborer des systèmes de collecte et d'analyse de données et d'échange d'information, et d'initier et de soutenir les recherches dans le but d'atteindre et de maintenir un statut de conservation favorable pour les albatros et les pétrels. Pour aider les Parties à cet égard, le Secrétariat propose l'adoption d'une politique d'ensemble en matière de données pour garantir que les données transmises par les Parties à l'ACAP et d'autres propriétaires de données soient correctement gérées et consultées. On trouvera à l'Annexe 1 un complément d'informations générales sur la nécessité d'une politique de données et sa mise en œuvre.

Politique proposée

La politique présentée ci-après est proposée pour adoption.

Afin de s'assurer que les données transmises à l'ACAP soient correctement gérées et consultées conformément aux stipulations des propriétaires de données, le Secrétariat:

1. mettra en place un catalogue de métadonnées pour décrire tous les ensembles de données transmises au Secrétariat.
2. veillera à ce que les limitations d'utilisation de données qui s'appliquent à des données ou des produits existants soient également appliquées à tous les futurs produits utilisant ces données.
3. permettra la consultation et le partage des données, sauf disposition contraire des propriétaires de données. Les propriétaires de données pourront stipuler la manière dont leurs données seront consultées et utilisées dans la section « Restrictions d'utilisation » de l'enregistrement contenant les métadonnées pertinentes.
4. appliquera un modèle de licence Creative Commons pour le partage des données et la réutilisation des données téléchargées.
5. archivera les ensembles de données transmis dans leur format original, qui sera lié à l'enregistrement des métadonnées pertinentes.

Recommandation :

Il est recommandé :

Que le Comité consultatif approuve la politique de données présentée ci-dessus pour adoption par les Parties lors de RdP4 [MoP4], ou par vote intersessions.

Informations générales sur la nécessité d'une politique de données et options pour sa mise en œuvre

En 2008, CC4 a adopté des « Règles pour la consultation et l'utilisation des données sur le statut et les tendances/les sites de reproduction transmises à l'ACAP et maintenues par l'ACAP » (Rapport final CC4, Annexe 13). Toutefois, étant donné le développement du portail Web, et en prévision de nouveaux types de données (par exemple, les informations sur la capture accessoire), il est désormais nécessaire pour l'ACAP d'avoir une politique de données qui englobe toutes les données transmises à l'Accord.

Une politique claire sera particulièrement utile pour les données qui pourraient mettre en jeu des non-Parties à l'ACAP, comme les ORGP [RFMOs], en particulier lorsqu'un Mémoire d'Entente a été signé. Elle permettra également à l'ACAP de partager éventuellement des informations avec d'autres réseaux de données, par exemple, la Global Biodiversity Information Facility (GBIF) [Système mondial d'informations sur la biodiversité], ce qui permettrait d'incorporer les ensembles de données de surveillance pour plusieurs espèces et régions. La GBIF invite l'ACAP à adhérer à la GBIF en juin 2009, mais aucune action n'a été prise en raison de l'absence de conseils d'orientation clairs de la part des Parties à ce sujet.

Il est proposé qu'une politique de données de l'ACAP comprenne des règles concernant la collecte des métadonnées, le partage des données, la publication, les conditions d'utilisation, et l'archivage, et qu'elle tienne compte, dans la mesure du possible, des règles d'accès aux données convenues lors de CC4.

Métadonnées

Les métadonnées sont, dans leur plus simple expression, « des données au sujet de données ». Bien que l'ACAP définisse les méthodes de collecte des données et les indicateurs de qualité des données, par exemple, en ce qui concerne les données sur les comptages de population, il n'existe pas, à l'heure actuelle, de descriptions des ensembles de données détenus par l'ACAP. Un catalogue de métadonnées fournirait cette fonctionnalité.

Il convient de noter que l'élaboration d'un catalogue de métadonnées propres à l'ACAP comprend le coût initial du développement de l'infrastructure et le coût permanent du chargement et de la gestion du contenu. On estime que deux semaines de travail seraient nécessaires pour faire progresser ces tâches pendant chacune des trois premières années.

Les données ne peuvent être découvertes que si le catalogue de métadonnées est accessible publiquement. Cela ne signifie pas que les données qui y sont décrites doivent être publiques ; le catalogue constitue simplement un mécanisme permettant de découvrir l'existence des données, tant restreintes que publiques. Il existe plusieurs catalogues de métadonnées globales et l'un d'eux peut être utilisé pour concevoir un catalogue adapté aux besoins de l'ACAP. On prévoit qu'à l'avenir, des enregistrements de métadonnées locales pourront être facilement exportés vers d'autres catalogues de métadonnées, en utilisant des normes de données et des protocoles d'échange bien établis.

Étant donné les exigences relativement simples pour un catalogue de métadonnées de l'ACAP, il est recommandé qu'un catalogue local soit élaboré à l'aide d'une structure initiale, l'identificateur des métadonnées, le titre, le propriétaire des données, le gardien ou le créateur des données, leurs coordonnées, les restrictions d'utilisation éventuelles, le résumé et les mots-clés.

Les gardiens devraient pouvoir saisir et mettre à jour les métadonnées. Les gardiens pourront être, soit:

- une personne mentionnée dans un enregistrement de métadonnées pertinentes, responsable des données ; soit
- une personne qui a éventuellement recueilli les données ; soit encore
- une personne qui sera le contact technique ou administratif pour toute personne ayant besoin de plus amples informations sur un ensemble de données.

Enfin, le Secrétariat devrait jouer un rôle de gardien de l'accès afin de garantir la qualité des enregistrements de métadonnées.

Gestion et publication de données

L'ACAP n'est normalement pas le propriétaire des données contenues dans sa base de données. Les enregistrements de métadonnées guideront le Secrétariat en ce qui concerne la communication ou non de données contenues dans la base de données de l'ACAP, conformément aux exigences du propriétaire des données.

L'ACAP produira de temps à autre des produits fondés sur les données existantes. L'ACAP veillera à ce que toutes les restrictions d'utilisation de données spécifiées par le propriétaire des données soient consignées dans le produit final. Par exemple, des méthodes appropriées d'agrégation des données doivent être appliquées pour les données sensibles.

Le catalogue de métadonnées doit fournir une liste des restrictions d'utilisation pour simplifier la réutilisation des données et empêcher la communication de données /d'une publication non conforme(s) aux restrictions appliquées par le propriétaire des données. On trouvera à l'Annexe 2 une liste des restrictions éventuelles ou des règles concernant les données.

Partage de données

La Global Biodiversity Information Facility (GBIF) encourage la publication et le partage de données sur la biodiversité (voir Partenariats GBIF <http://www.gbif.org/governance/partnerships/>). Ces portails permettent de découvrir facilement d'autres ensembles de données qui pourraient être applicables à l'ACAP. Ils améliorent également la visibilité des participants, permettent une plus grande reconnaissance et encouragent des collaborations susceptibles d'ajouter de la valeur aux ensembles de données en incorporant des données provenant de plusieurs sources (Costello, 2009). Le Secrétariat recommande que l'ACAP adhère à la GBIF dans le cadre d'une collaboration croissante de la biodiversité.

Toutefois, il est également reconnu que les données transmises à l'ACAP sont extrêmement précieuses, les possibilités de collecte de données dans des endroits éloignés étant souvent limitées. C'est pourquoi le contrôle de l'accès aux données et de leur utilisation est d'une importance critique pour les propriétaires de données.

On devrait demander aux propriétaires des données d'indiquer de manière explicite les données non publiques et d'insérer des clauses de temporisation [sunset clauses], ce qui permettrait une diffusion rapide des données.

Questions concernant les conditions d'utilisation

La mention de la paternité des données est importante. L'acquisition des données utilisées par l'ACAP a souvent été coûteuse et les organismes de financement exigent généralement les mentions appropriées. L'organisation Creative Commons (<http://creativecommons.org>) a élaboré des conditions de licence relativement simples qui peuvent être appliquées à de nombreux types d'objets, y compris des données et des images. Il convient toutefois de noter que l'application rigoureuse des conditions de licence pour l'utilisation du contenu d'un site Web quelconque est pratiquement impossible à faire respecter.

Les deux catégories de licence applicables à l'ACAP sont les suivantes :

- **Mention de la paternité :** Chaque fois qu'un ouvrage est copié ou redistribué sous une licence Creative Commons, il convient de mentionner le créateur, par exemple, pour les données, les produits de données, les rapports publics de l'ACAP.
- **Mention de la paternité - Fins non commerciales - Pas de produits dérivés :** Permet de copier, distribuer, afficher, et interpréter seulement des copies exactes de votre œuvre, pas d'œuvres dérivées fondées sur elle – mais uniquement à des fins non commerciales, en mentionnant le créateur, par exemple, photos appartenant à des particuliers reproduites sur le site Web de l'ACAP.

Il est proposé que le symbole de Creative Commons pour ces deux catégories de licence soit placé aux points de téléchargement appropriés sur les pages du portail de données.

Il est recommandé, en règle générale, que le Secrétariat sollicite la transmission d'images sans conditions d'utilisation à titre onéreux (généralement commerciale), étant donné que l'ACAP n'a ni la capacité ni le désir de faire respecter l'application rigoureuse de ces conditions.

Archivage

Les données sont précieuses et la perte de données est une préoccupation constante pour toutes les organisations. Une organisation comme l'ACAP nécessite un accès permanent aux données pour leur réutilisation ou l'application de meilleures méthodes d'analyse. Pour assurer la disponibilité de ces données, l'ACAP doit disposer d'un service d'archivage pour les ensembles de données dont le stockage n'est pas garanti dans d'autres dépôts de données pour toute période plus courte que la durée de vie de l'ACAP. Ceci a pour corollaire que les données non détenues par l'ACAP qui doivent être utilisées dans le cadre de la communauté de l'ACAP doivent être facilement accessibles (soit comme données publiques soit par le biais d'accords d'échange de données).

L'ACAP archivera tous les ensembles de données, qu'ils soient destinés à un usage public ou privé, dans leur format natif, avec une fiche de métadonnées appropriées contenant le nom du propriétaire, les méthodes de collecte, et le délai pour la diffusion publique (le cas échéant). Ces données pourront être ajoutées à la base de données de l'ACAP si elles ajoutent de la valeur pour la communauté de l'ACAP. Il est suggéré que le format natif de soit limité à des formats de données couramment utilisés comme texte, MS-Word, MS-Excel, PDF, JPEG.

Les ensembles de données, qu'ils soient reçus sous forme de fichier unique ou de fichiers multiples, seront compressés en un « fichier archivé », préservant ainsi les structures de fichier et les types de données. Chaque ensemble de données sera enregistré dans un catalogue de fichiers.

Références bibliographiques

ACAP (2008). *Rapport final CC4*. Annexe 13, ppe 102-103. www.acap.aq
Costello, M. J. (2009). Motivating Online Publication of Data. *BioScience* **59**: 418-427

PROJET DE RÈGLES POUR LES DONNÉES

Les règles ont été divisées en deux séries (accès et de filtrage) et peuvent être combinées de quelque façon que ce soit, mais certaines combinaisons ne seraient ni pratiques ni logiques.

Niveaux d'accès aux données

Les niveaux d'accès n'ont d'incidence ni sur le copyright ni sur la mention de paternité. Ils sont sauvegardés avec les données.

Les règles peuvent être appliquées à un ensemble de données complet ou à un nombre quelconque de lignes de données individuelles dans un ensemble de données. Une seule règle est choisie par donnée ou ensemble de données.

Une clause de temporisation est une date prédéfinie fixée par le gardien au cas où une règle de données peut être modifiée. Généralement utilisée lorsque les données peuvent être utilisées librement après une période d'accès limité, par exemple, lorsque les données/l'analyse ont été publiées.

Règle	Nom de la règle	Objet	Questions de sécurité et commentaires
1	Accès public	Toutes les données sont accessibles à n'importe qui par téléchargement. Tous les accès sont enregistrés, que l'utilisateur ait ouvert une session ou non.	La paternité des données peut être perdue lorsque les données sont réutilisées par des utilisateurs qui n'adhèrent pas aux conditions de téléchargement. Note : Ce niveau est déjà utilisé pour toutes les pages publiques du portail de données de l'ACAP.
2	Accès contrôlé	Toutes les données sont accessibles à n'importe qui par téléchargement mais l'utilisateur doit ouvrir une session au moyen d'un code d'identification valide.	Il est possible d'appliquer une clause de temporisation optionnelle à cette règle, ce qui permettra d'utiliser la règle 1 pour les futurs accès.
3	Accès contrôlé et limité	L'accès aux données est réservé à des utilisateurs connus, soit un gardien de données de l'ACAP et/ou d'autres utilisateurs externes, sélectionnés par le Secrétariat de l'ACAP.	Cette règle peut être utilisée pour la diffusion de données à d'autres organisations. Il est possible d'appliquer une clause de temporisation optionnelle à cette règle, ce qui permettra d'utiliser la règle 2 pour les futurs accès.
4	Gardiens de données seulement	L'accès des données est réservé aux gardiens de données de l'ACAP.	Règle utilisée pour échanger des données entre les groupes de travail. Considérée comme distincte de la règle

			3, cette règle a pour but d'indiquer clairement le niveau d'accès qui a été fixé.
5	Secrétariat de l'ACAP	L'accès des données est strictement réservé au Secrétariat et aux responsables de l'ACAP.	Néant

Agrégation de données et règles de filtrage

Ces règles peuvent être appliquées aux données à n'importe quel niveau d'accès. On peut appliquer n'importe quel nombre de ces règles à un ensemble de données.

Règle	Nom de la règle	Objet	Questions de sécurité et commentaires
1	Pas d'agrégation	Toutes les données sont telles qu'elles ont été communiquées.	Le traitement par défaut de toutes les données communiquées.
2	Agrégation spatiale	Données agrégées en fonction de la sélection de certaines répartitions spatiales [spatial binning] de taille ou de région	Les données peuvent être agrégées pour réduire la capacité d'analyser les données.
3	Résolution spatiale	La précision des positions de données est réduite.	On peut garder le nombre original d'enregistrements.
4	Résolution temporelle	Les dates de collecte de données/d'observation peuvent être agrégées en unités de temps plus grandes.	Les données peuvent être agrégées pour réduire la capacité d'analyser les données.