

(in)actions autour de la base de données

Coordination :

Laurence Fleury, Laurence Eymard

ESPRI, IPSL, Palaiseau : gestion des produits satellite, modèle ...

Sophie Cloché, Karim Ramage

SEDOO, OMP, Toulouse : observations in situ, interface utilisateur

Laurence Mastrorillo, Guillaume Brissebrat, Jean-Luc Boichard

Centre Régional AGRHYMET, Niamey : toutes les données, services identiques

Oumarou Moulaye, Hadiza Aboubakar, Brahima Koné

Rappel réunion de lancement de mars 2011

Déjà à la disposition des participants à ESCAPE :

Le système d'information AMMA

- Un système assurant la gestion / distribution de données de différents types (satellite, modèle, in situ géophysique, questionnaire...) <http://amma-international.org/database>
- Une base de ressources bibliographiques

A élaborer avec les WP ESCAPE :

- Un moyen d'échanger les données au sein du projet : aspects « data policy » (D6.1 – M6)
- Echanger oui bien sûr ... mais quelles données ? Avec qui ?
 - définition : M6
 - alimentation : M0 à 48
- Des besoins spécifiques ?
 - demandes : le plus tôt possible
 - évaluation + réalisation selon RH (6 mois CDD)

Règles d'échange de données

Toujours à finaliser car 2 options possibles :

Option AMMA-like

- Proposer une charte de données très proche de celle de AMMA
- Période de réserve des données de 2 ans pour ESCAPE

→ OK pour données géophysiques, agronomiques, etc

→ Spécificités des données SHS, personnelles, etc. non traitées.

Seules mentions :

- Respect lois et règles éthiques de collecte de données + anonymisation des données à la charge du PIs
- Si problème à ces égards relevé par un utilisateur, obligation de le signaler au PI + au contact base de données

Règles d'échange de données

Option AMMA + SHS

→ Intégrer les spécificités des données SHS, personnelles, etc.

→ Problème : sujet non mûr en France mais ... lancement de n+1 groupes de réflexion...

Pistes :

- TGIR Corpus

- groupe de travail droit / MSH (images, données personnelles) (Aix – Dijon – Nanterre – Tours ...)

- Liens avec labos droit // bases de données (ERCIM, UMR 5815), Université de droit Toulouse.

- Guide des bonnes pratiques pour les questionnaires TGE Adonis prochainement (slt technique ?)

- **HELP needed !!** Idées, pistes (+ pertinentes ?) venant des participants à ESCAPE ?

Règles d'échange de données

Proposition (M6 déjà dépassé...)

- Data policy AMMA-like + mention addenda pour certains type de données / WP ? Lesquelles ? Lesquels ?
- Création d'un **groupe de travail base de données pour les SHS** (par mail, téléconf.) dans ESCAPE ??

Axes de réflexion :

- Quelles données SHS dans la base de données ESCAPE ?
- Cercle de diffusion ? (interne ou diffusion après 2 ans)
- Conditions de diffusion ?
- Outils à mettre en place ? Formats ? ...

Traitement différent selon

- données préexistantes (quel propriétaire ?) ou données générées dans le cadre de ESCAPE
- données personnelles ou intégrées

Collection / Corpus de données SHS - ESCAPE

- Rassemblement de données pour les participants au projet ESCAPE (point de contact : Charles ?)
- Possibilité de mettre en place de l'espace, des facilités d'accès (http, ftp...) au SEDOO (+ duplication au CRA ?) pour le WP2.
- **Cas test : résultats de ECRIS ?**
Partage outil : contribution au rapprochement disciplinaire ?
- Eventuellement : outils associés (fouille de données, etc.)
cf. 6 mois de CDD ESCAPE.

Message

SEDOO + IPSL = un **service informatique**
dédié données pour ESCAPE
(au moins 5 informaticien/ne/s et 1 traductrice)

+ CRA, Niamey (2 informaticiens)

Utilisez le !! *

*** Ce message s'adresse aux WP 1, 2, 3 et 4 !**

Infos nouveaux jeux de données

- Données SYNOP 1980 – 1995 insérées dans la BD.
Stations météorologiques synoptiques.
Zone : 50°E – 50°W / 20°S – 40°N
- ERA INTERIM : bientôt accessible via la BD AMMA.
période : 1989 – présent
0,75°x0,75° ou 1,125° x 1,125°

ANNEXES

Règles d'échange de données

AMMA data and publication policy 2005-2009.

La nouvelle en cours d'adoption (prochain IGB ?).

Même philosophie : favoriser les collaborations.

Devoirs des fournisseurs de données :

- Meilleure version possible des données.
- Respect lois et règles éthiques de collecte de données + anonymisation des données.

Droits et devoirs des utilisateurs :

- accessibilité, gratuité des données
- pas de redistribution
- pas d'applications commerciales
- contacts entre fournisseur/utilisateur de données :
 - collaboration (mails automatiques)
 - copublication à proposer / accord minimal : remerciement.
- remerciements aux institutions finançant le programme
- produits issus de jeux de la base de données à fournir à celle-ci.
- ...

Règles d'échange de données

“ This document covers mainly the data collected during the AMMA programme from 2002. Other projects will use and feed the AMMA database. Individual scientists, principal investigator teams and projects will be permitted a reasonable period (with a maximum of 2 years) of exclusive access to datasets which they have collected, allowing them to work on them and produce publications. At the end of this period, the data will be governed by the present policy. Any individual scientist, team or project is required to propose an acknowledgment sentence which will be used in publications using the supplied data. ”

? Pour **ESCAPE** : période d'accès réservée ou non ?

Si oui, pendant cette période : a priori règles AMMA devraient être OK, mais distribution dans le projet seulement → à vérifier.

Tâches dans le cadre de ESCAPE

- Données à échanger

- définition : M6

- alimentation : M0 (www, données préexistantes...) à 48

(n.b. : attention quand contrat achat type : mettre tous participants dans le projet. Ex : images SPOT)

- Des besoins spécifiques ?

- Besoin d'un thesaurus pour données ISEC (actuellement GCMD pas complet pour ces disciplines)

- Chaque questionnaire : base de données relationnelle dédiée sous MySQL (interface de saisie / de consultation / de requêtes...)

- Entretiens directifs, semi-directifs... : archive texte (+ audio), base de données relationnelle des métadonnées, fouille de données...

- Travail sur anonymisation à la charge du PI mais conseil + éventuellement développements possibles (cryptage, hachage...)

- Formats de fichiers : besoin d'extraction format SIG ?