

Compte-rendu de la réunion de suivi CDS (12 mars 2003)

Présents: F. Bonnarel, L. Cambrésy, S. Derriere, F. Genova, S. Lesteven, A. Schaaff, M. Wenger

Invité: T. Boch

Excusés: M.G. Allen, P. Dubois, P. Fernique, F. Ochsenbein

Bibliographie

- Retard A&A et ApJ: journaux non reçus à Paris et/ou Strasbourg. Le retard a été rattrapé. ApJ est fait sur version électronique à Paris.
- Vérification des liens sur les objets taggés par les auteurs pour A&A: confié à GT qui entre les données A&
- =g= et =C=: retards à cause du temps passé sur LEDA et le LHS.

==== > revoir les procédures de sélection des catalogues à entrer?

==== > catalogues radio (qui constituent une partie des =C=): post-doc en août

==== > (ajouté au compte-rendu) réexamen des procédures (réunion repoussée, à organiser)

- Acronymes:
 - o création des acronymes: gros retard (démonstration AVO, accumulation des montages de projets, réunion du Conseil, ...)

==== > Résorption débutée

==== > est-il possible de mettre en place un partage des tâches sur la création et/ou de faciliter la tâche (à discuter FG, SL, MW)?

- o Validation du nouveau logiciel de gestion des acronymes (voir ci-dessous): aide demandée à SL
- Travail d'expertise scientifique de LC: à discuter - biblio? catalogues(voir ci-dessous)?
- Formation de Ph. Vonflie à commencer dès que possible (PD)
- Transfert de la machine de travail de MJW (il faut arrêter cdsxb9; trouver une solution pour définition des clés) (action SL)
- NewA est maintenant distribué par Science Direct (OK à Strasbourg, abonnement obtenu pour SL); ils ont promis de rétablir les liens objets/simbad; ne pas y passer trop de temps

Contenu

3 grosses manipulations en cours:

- LEDA: reste 13000 problèmes à traiter. Problème de ressources humaines, à voir avec James? Les mesures ont été rentrées (VR). Attention à l'identificateur 'primaire' qui ne doit pas être systématique LEDA pour les galaxies: corriger la table des identificateurs par type d'objet.
- AAVSO. 5000 étoiles variables. Fichier de correspondance fait. FW a contacté J. Mattei pour les 60 problèmes restants. Acronyme vérifié avec J. Mattei et créé. Lien vers la base AAVSO à définir/construire (MW). Proposer à J. Mattei de mettre des liens de sa base vers SIMBAD (action MW).
- ISO. Repris par ED et FW. Visite ED pour ISO et discussion avec Bernt Vollmer (20-21 mars)
- LHS. CB a terminé le travail. Commentaires à intégrer dans Vizier. Il faut un lien entre SIMBAD et Vizier pour pointer sur ces commentaires.

SIMBAD

- Nouvelle version:
 - o Le stage de S. Jaehn terminé: elle a produit une base SIMBAD complète sans les mesures (deux stratégies possibles pour les mesures: implémentation semblable à l'état actuel, ou réflexion plus approfondie sur les liens avec Vizier). Tout le reste est développé avec une API fonctionnelle et un client Java minimaliste qui permet de faire des requêtes.
 - o Stage 2002 (J.P. Lejal): l'interface de mise à jour est restée en attente. Il reste à décider si elle doit être mise en place rapidement dans le SIMBAD actuel ou si cela attendra la nouvelle version.
 - o Stage d'O. Dellicour (en cours): identificateurs. Noyau du traitement en cours de développement, sur trois types de codage représentant 50% des catalogues. Le nouveau codage est déduit de l'actuel. Problèmes de performance pour le moment. Par ailleurs, les descriptions des identificateurs se retrouvent dans plusieurs fichiers gérés dans des endroits différents. Un objectif à long terme est de regrouper le tout (stratégie à définir). SL souhaite avoir la traduction en expression régulière dans le Dictionnaire. SL fournira sa liste de traductions en expressions régulières, créées pour la recherche semi-automatique de noms d'objets, à MW.
 - o Bibcodes: ils seront stockés comme chaîne de caractères. Refaire la liste des modifications nécessaires (certaines identifiées il y a plusieurs années, par exemple caractères PQR, grands numéros de session des BAAS, ...): SL fournira la liste à MW.
 - o Tous: envoyer la liste de ce que vous souhaitez voir dans la prochaine version de SIMBAD à MW.
- Quelques plantages dans la mise à jour, la plupart résolus, mais il reste le problème des 'fusions à trois' en cours mais non terminé.
- FB a repris les modifications en cours de la mise à jour, en particulier les commentaires objets, des demandes de GJ, d'autres petites modifications. Prévenir rapidement FB des demandes résiduelles (probablement la dernière série de modifications faite par FB avant la nouvelle version). Les commentaires objets doivent être installés dès que possible.

Catalogues

- Très grands catalogues:
 - o installés: GSC2.2 (500 millions objets avec photométrie et positions), USNO-B1.0 (1 milliard objets avec positions, mouvements propres, photométrie)
 - o à venir bientôt:
 - SDSS
 - 2MASS Final (qui devrait être disponible en mars; l'équipe 2MASS a envoyé un échantillon pour commentaires – discussion en cours)
 - DENIS (87% du relevé traité à Paris; programme de filtrage fait par SD; 2200 zones sur 5500 reçues à Strasbourg, mais zone réobservées, strips à problèmes, ...: installation au CDS en attente de la liste de ce qu'il faut mettre en ligne; problème pendant: headers FITS des images, dont l'astrométrie est très pauvre)

- zones blanches détectées par Aladin dans l'USNO-B >> à étudier (pas un pb dû à Aladin ni aux plaques sur lesquelles le test a été fait)
- Catalogues demandés depuis longtemps et installés: NVSS
- Autres catalogues (régime de croisière = 450cat./an)

Interface de Vizier (vérifier que les améliorations sont décrites dans la documentation):

- Visualisation des UCDS
- Recherche à partir d'une carte globale ("Vizier Mine")
- Améliorations de l'interrogation par listes:
 - copie en sortie d'un champ d'entrée (pour pouvoir faire le lien entre la liste d'entrée et ce qui a été retrouvé par Vizier).
 - pouvoir utiliser un résultat de Vizier comme liste d'entrée
 - (par exemple rechercher dans un nouveau catalogue les sources se trouvant autour des objets trouvés dans un résultat précédent)
 - Installation de l'interface de traces VOPlot. À tester.

Les clones de Vizier: 8 actuellement (ADAC Tokyo; IUCAA India; CADC Canada; Cambridge UK; INASAN Moscou; Pekin Chine; UKIRT Hawaii; CFA USA). Ça demande pas mal de travail pour:

- porter les softs, les compiler (c'est pas automatisé, cela pourrait éventuellement se faire si les versions de Solaris et Linux étaient les mêmes sur les différents sites...)
- interagir avec les system managers pour les ajouts de place disque / Sybase management
- communiquer les problèmes
- adapter les softs pour tenir compte des demandes (par ex. sécurité)

Par exemple depuis décembre il y a eu

- hacking au CADC et en Chine ==> beaucoup de messages échangés pour re-ouvrir les ports nécessaires, réinstaller les softs et données
- problèmes de stabilité avec Linux (kjournal qui génère des problèmes)
- réinstallations complètes des données suite à un changement de version de Sybase à Cambridge (UK)

==> La gestion des copies miroirs est problématique parce que nous n'imposons pas de spécifications. À faire évoluer. Demander à installer l'ensemble des clones sur machine standard. Si trop de problèmes, il peut devenir nécessaire de fermer certaines copies.

Utilisation de Vizier: la somme des clones donne une utilisation à peu près équivalente à celle du site principal. L'augmentation du nombre de Kbytes est reliée à l'augmentation de l'interrogation par listes.

Les problèmes actuels et développements en cours:

- entrée des catalogues dans le pipe-line: relecture des ReadMe et insertion dans Vizier demande du temps (entre autres l'assignation des UCDS qui ne peut être totalement automatique)
- en développement:
 - o recherche des catalogues par UCD
 - o sélection sur champs calculés (exemple: coordonnées galactiques)
 - o sujet de stage: recherche multicritère sur les très grands catalogues

Aladin

- Réunion bilan/perspectives le 31 janvier. Un compte-rendu sera diffusé.
- Mise en service officielle de la nouvelle version d'Aladin 1.4 début février, intégrant : les filtres, la recalibration manuelle, la compatibilité VOTable, le stack illimité et foldérisé. Les clones (sites VizieR) ont été mis à jour quelques jours plus tard. Aucun problème à noter.
- Poursuite des développements Aladin pour finir d'intégrer les idées et le code issus du proto AVO. Principalement il s'agit de la récupération du concept de l'arbre GOOD pour n'importe quelle région du ciel (développement TB) et visite à Pune pour l'intégration de VOPlot dans Aladin a priori sur le même principe que les tracés SED dans le prototype AVO (S. Kale + PF). En cours...
- Amélioration de l'interface pour les filtres. Mini-réunion Aladin fin février dont sont sorties quelques idées pour simplifier l'écriture des filtres. Développement en cours (TB).
- Mise à jour de la doc. Pages HTML, FAQ, Copyright ont été mis à jour. Le manuel de l'utilisateur est en cours (DE).
- Ergonomie de l'interface et tests de configuration: discussion avec APM. Il serait souhaitable de passer une série de tests de fiabilité sur les configurations JAVA 1.3 et 1.4. Les essais d'APM sur Linux/KDE/JVM.1.3 sont inquiétants. Tests à faire sur MacIntosh. Tests: PF, autres?
- GOODS et interopérabilité:
 - o nouvelle version de la sortie XML qui corrige quelques bugs et corrige des incohérences par rapport au modèle IDHA. Prise en compte d'un localisateur de fichier dans le document. Publié dans la liste DataModel. La documentation va être écrite en collaboration avec MGA d'ici fin mars.
 - o Contact avec D. Tody pour intégration SIA. Les données CDS seront présentées et utilisées dans les démos NVO de Sydney. Réunion Data Access Layer prévue, non confirmée, recontacter D. Tody.
- Données
 - o avec JM: DSS2 et MAMA – tests complets pour le moment pour une plaque sur 4. Reprendre un test systématique des plaques entrées au début (mise en place de la procédure).
 - o reprise des sauvegardes sur bandes (nouvelles sauvegardes, reprise de sauvegardes sur bandes 'agées')
 - o données 2MASS: courant avril (sera fait par FB dans le nouveau serveur)
 - o MSX (pour AstroVirtel): données LMC intégrées. En attente d'un serveur ftp à IPAC pour le reste.
 - o SDSS (Early Data Release): étude d'installation dans la base test; problème de temps de transfert à étudier. Volume total 2.4 TB. Il faudra étudier et acheter un nouveau système RAID.
 - o CFHTLS: il va falloir faire une demande au Steering Group qui inclurait plusieurs options pour l'accès aux données, accès à distance ou copie partielle des données (de type sortie Terapix) Strasbourg.
 - o Intégrer les époques 3 et 4 des données ACS dans le serveur de démo (demande MGA).

AVO/projets

- IDHA: réunions régulières sur les aspects MARSIAA. La modification du modèle de données suite à la réunion de janvier à l'ESO reste à faire, programmée pour les jours qui viennent. AO a repris une nouvelle version du programme. Les tests continuent avec de nouveaux jeux de données avec C. Bot. Plusieurs stages démarrent (techniques markoviennes floues, ACI), thèse en cours, thèse demandée en BDI, reprise de l'analyse des données TERAPIX. Réunion ACI GRID Nice le 22/4 (préparation le 8).
- UCD: document de bilan, peu de commentaires, contact avec C. Davenhall qui promet ses commentaires pour la semaine prochaine. Reprise systématique de l'arbre par APM, avec vérification de toutes les associations. Adaptation des outils. SD étudie la mise en oeuvre des UCD 'atomiques' et des suffixes (mutuellement exclusifs / orthogonaux). A faire: organisation de la réunion de Tiger Team de Mai. UCD/DataModel: images – repartir des mots-clé FITS (sans référence a priori au data model) (à proposer à APM?); en parallèle terminer l'action de corrélation du modèle avec les UCD existant. Présentation au Workshop Electronic Service and Semantic Web (Budapest) en mai.
- Web Service: mis en ligne: Sesame, GLU (resolver de tags pour renvoyer des URL); en cours: GLU (accès au Dictionnaire pour récupérer les informations). Présentation 'GLU dans Web Service' proposée à la conférence de Budapest, le papier sera envoyé à la liste Registry.
- Registry (T. Linde, Cambridge, 19-20 mars) (SD, AS, MGA). Point stratégique à faire. Cerner les spécificités du GLU en ce qui concerne ses fonctionnalités registre. PF a rédigé un texte de bilan sur le GLU. Modifications en cours (PF) dans Aladin: passage en Java de l'interrogation des sites distants dans Aladin. Il est important d'être réactif sur les mailings lists (un groupe de travail restreint va être constitué, AS proposé)
- FP6 Europe: EuroVO (15 avril, ESO, ESA, CDS, AstroGrid), VONet (6 mai), VOTech (automne)
- ACI MD: participants potentiels IRIT, LSIIT, LORIA, IRIT, OCA Bijaoui, OP (M.L. Dubernet). Définition du contenu en cours.

Colloques et information utilisateurs

- ADASS: à Strasbourg 13-15 octobre. Repris par MW. Contacts repris avec le Palais des Congrès, avec Congrès Louis Pasteur (inscriptions, réservations hôtel, visites, ...) et avec les personnes qui gèrent le logiciel des pages Web (base Sybase).
- Tutorial CDS à remettre à jour

Visiteurs

D. Vandromme (directeur de RENATER)	12 mars
A. Holl	17-30 mars
E. Davoust, B. Vollmer	20-21 mars
G. Eichhorn	3-4 avril
B. Mann	mai
M. et E. Kontizas	juin
W. Osborn, J. MacConnell	juin