

Réunion biblio

20 septembre 2018

Présents : Cécile, Fabienne, Philippe, Patricia, Emmanuelle, Catherine, Tiphaine, Aline, Magali, Evelyne, Mihaela, Anaïs, Gregory, Esther, Bernd, Gilles, Mark et Soizick

Excusés : Marianne, Caroline, Pierre

1. Points d'information

- Conseil du CDS le 16-17 octobre :

Seul un membre du conseil sera absent cette année.

Mark doit finaliser le rapport avant la fin du mois.

En plus des points habituels (suivi et évolution du travail), il y aura un point particulier sur le Plan de Reprise d'Activités (PRA), ce qui avait été demandé l'an dernier.

- Projet européen ESCAPE : European Science Cluster of Astronomy & Particle physics ESFRI

Le CDS participe à différents Workpackages mais principalement au WP 4 qui concerne les données dont Mark est le responsable.

- Les journées FréDocs : 3 au 5 octobre 2018 à Albi avec comme sujet : « Démarches innovantes en IST : expérimenter, proposer, (se) réinventer ». Aline Eisele y participe cette année.

- La journée inter-réseaux organisée par le CNRS avec comme thèmes : "Interopérabilité et pérennisation des données de la recherche : comment FAIR en pratique ?" le 27 novembre 2018.

2 présentations du CDS :

- « Observatoire Virtuel et FAIR, des principes fondamentaux à la pratique » par André Schaaff

- « Les documentalistes et la FAIRisation des données scientifiques : un travail d'équipe inter-métiers » par Soizick Lesteven

- UAI (20 au 31 août 2018) :

Il y avait un stand IVOA auquel le CDS a fortement participé pour montrer ses services.

Discussions avec le publisher Springer qui nous demande un état des lieux afin de mettre en place un contrat.

Marianne, présente sur le stand, a travaillé avec M. Schmitz sur le dictionnaire de nomenclature.

- Journée Doc-Astro à Paris le 29 mars 2018, première rencontre fructueuse avec toutes les bibliothécaires et documentalistes travaillant en astronomie en France : 16 personnes, 10 labos.

La présentation à 5 voix du travail des documentalistes du CDS a été très remarquée. La prochaine rencontre se tiendra à Strasbourg au printemps 2019.

2. Cross-identification de SIMBAD avec Gaïa DR2

Cécile nous fait un compte-rendu de l'opération de cross-identification entre les sources de Simbad et le catalogue Gaïa DR2. Sur les 6 millions de X-id potentielles, 4,5 millions ont été réalisées dans Simbad (objets ayant une bonne astrométrie - qualité A, B ou C - dans Simbad et sans autre voisin à moins de 3").

Toutes ces informations se trouvent sur les news du CDS :

http://cdsweb.u-strasbg.fr/news.php?fn_mode=fullnews&fn_incl=0&fn_id=702

Les retours sont très positifs. Il a été noté que si ce travail s'est globalement bien passé c'est parce qu'il avait été bien préparé par toute l'équipe, notamment grâce à l'ajout systématique de coordonnées plus précises dans Simbad ces dernières années.

3. Point sur les différentes équipes :

- DJIN : Bonne remontée du MNRAS malgré tous les problèmes de réception des données provenant de OUP, le traitement des journaux principaux se maintient dans un délais raisonnable (1 à 3 mois). Le problème du MNRAS semble résolu pour le moment.

- CoSim : Cécile note une forte augmentation des $=g0=$, il y en a actuellement 900 en attente, normalement on en compte 300. Ceci est dû au nombre important de tables provenant de Vizier. Il faudrait de l'aide d'un astronome pour le traitement/tri/évaluation des $=g0=$. Cécile en fait hors réunions $=g=$ pour diminuer le retard.

Emmanuelle propose que l'équipe Vizier mette des statuts «possible egal v » (à définir) sur les tables de Vizier pouvant être entrées dans Simbad. Elle pourrait les traiter directement, sans passer par les réunions et le statut $=g0=$, comme elle le fait déjà sur ses références avec une vérification de Cécile. Ces tables seraient disponibles bien plus rapidement dans Simbad.
(rappel : $=v=$ les objets de ces tables sont déjà dans Simbad).

-Dictionnaire : Bernd rappelle que Marianne et Fabienne sont toutes les 2 co-responsables du dictionnaire, au même niveau, et qu'elles y travaillent toutes les 2 à 100 % (Anaïs a modifié les outils pour leur permettre de travailler en même temps). On constate une diminution du nombre d'acronymes en attente. Il y a toujours du retard mais celui-ci s'absorbe progressivement. Cette procédure convient aux documentalistes. Ils apprécient que Bernd mette des commentaires plus détaillés.
Pour accélérer le traitement des références examinées lors des réunions $=g=$, Bernd propose de mettre directement les statuts et les commentaires lors du traitement des acronymes.

4. Les commandes gsc4sim et find

La commande gsc4sim qui est utilisée par toute l'équipe quotidiennement nécessite d'être étendue à d'autres catalogues. Ce programme (écrit par François pour le gsc uniquement au départ) est difficilement maintenable par Gilles et Anaïs. Il faut le réécrire entièrement.

Il s'agit d'un besoin urgent.

De même pour la commande find dont la sortie ne correspond pas aux besoins des documentalistes.

→ Nécessité de reprendre ces commandes. Il faut organiser une réunion où chacun donne ses besoins pour construire un cahier des charges prenant en compte les attentes de chaque groupe d'utilisateur. Un groupe de travail gsc4sim existe déjà, il faut le réactiver.

5. Nouvelles demandes pour Simbad

- Historique des modifications des mesures dans les DF : ces données sont importantes pour la traçabilité des données et la compréhension de l'historique des objets, notamment dans les cas où des corrections sont nécessaires dans Simbad.

Anaïs propose de faire un équivalent du oldsim avec une sauvegarde 1 fois/mois et une sortie plus lisible et exploitable (à définir avec elle).

Pour le moment on peut retrouver l'info dans les logs.

- Ouvrir le champs « référence d'origine » à autre chose que le bibcode : cela est utile dans les cas où on rentre des données fondamentales provenant d'autres sources que les articles (ex: les coordonnées des GRB venant du Swift-XRT, les données de LEDA, etc...). La proposition retenue est de pouvoir mettre une URL dans ce champs, pointant vers la page d'accueil du site d'où vient l'information. La possibilité de mettre un « faux bibcode » n'a pas été retenue.

- Rawids visibles en interne dans les pages WEB de recherche par référence : cela facilite l'identification des objets cités dans des articles. Ok, Anaïs s'en occupe.

- Identificateurs cachés visibles en interne uniquement : permet d'identifier des objets cités dans la littérature sans multiplier les NAMEs ni propager les identifications incorrectes. Ok, Anaïs s'en occupe.
-> Il reste à déterminer quel ID on utilise pour définir les IDs cachés (VIRTUAL...?).

6. Les noms d'objets dans les KWs, titres, abstracts :

Depuis le début de cette année (2018), toute l'information concernant les références bibliographiques est dans Simbad.

Les liens sur les noms d'objets des titres et mots clés se font grâce aux rawids associés via DJIN, sauf lorsqu'il y a un tag sur l'objet. En cas de lien non actif, il est possible de l'activer en ajoutant un tag. En aucun cas il n'est possible de modifier les données provenant des éditeurs, seul l'ajout des tags est autorisé. Ces tags peuvent être utilisés dans les titre, abstracts et KWs.

Vu les statistiques données par Anaïs, ces liens sont peu utilisés (118 clics en septembre), il n'est donc pas nécessaire d'y passer beaucoup de temps. Mais si vous voulez le faire, voici un rappel des commandes disponibles pour tous en maj :

- `\object{id}` génère un lien vers un objet de Simbad, l'objet est reconnu par Simbad..

Exemple : `%K quasars: individual : \object{PG 1329+412}`

- `\objS{idS}{idL}` génère un lien sur un objet Simbad tout en gardant visuellement l'identificateur venant de la littérature (non reconnu par Simbad) :

Exemple : `%K quasars: individual : \objS{QSO 0450-132}{0450-132}`

- `\objnS{id}` évite de créer un lien sur un objet qui n'existe pas dans Simbad.

Exemple : `%K stars: individual: \objnS{Sun}`

7. Formations

Anaïs propose de (re)faire des formations vi et sql pour les documentalistes qui le souhaitent.

Fabienne propose de faire des réunions pour discuter et (re)définir ce que l'on met dans les données fondamentales et dans les mesures de Simbad. Ceci devrait permettre d'homogénéiser les pratiques entre documentalistes et faire le point sur les pratiques actuelles (les qualités, ce qu'on entre ou ne rentre pas/plus dans Simbad...).

Conclusions :

Il ressort de cette réunion qu'une forte charge de travail pèse sur tous, documentalistes, informaticiens et chercheurs. Tout le monde a une grande volonté de bien faire son travail mais cela devient difficile dans les conditions actuelles de surcharge de travail. Il est vraiment important d'améliorer la communication entre tous, de faire des réunions biblio et/ou de travail plus fréquemment et de redéfinir les priorités. Il est urgent de prendre en charge les besoins en programmes informatiques d'aide à l'ingestion des données spécifiques aux documentalistes.