

CR de la réunion biblio du 10 octobre 2023 - 14h grande coupole

Présents : Caroline Bot, Marianne Brouty, Catherine Brunet, Mihaela Buga, Esther Collas, Aline Eisele, Pierre Fernique, Ana Fiallos, Coralie Fix, Anaïs Gonneau, Soizick Lesteven, Cécile Loup, Grégory Mantelet, Magali Neuville, Anaïs Oberto, Pierre Ocvirk, Emmanuelle Perret, Arnaud Siebert, Katia Van den Woerd, Patricia Vannier, Bernd Vollmer

Excusés : Laurent Cambrésy, Evelyne Son, Mark Allen

Ordre du jour :

Révision du workflow des tables : attribution et suivi des =e= / =s=

Point ajouté à l'OJ : =q=Acro

Rappel :

- Statistiques 2022 (MNRAS & ApJ pas encore traités pour cette année de sélection)

- =e= Tables qui peuvent intéresser VizieR et dont les données peuvent intéresser SIMBAD. Le statut est donné par l'équipe DJIN. Tables de plus de ~100 lignes (difficilement traitables par DJIN, MRTs, data behind figures, catalogues décrits, données en ligne (http, fits, pdf, csv, tar.gz...) non accessibles...). En cas de doute, il faut mettre un =e= et l'équipe VizieR décide. =e= est modifié en =f= par l'équipe VizieR lorsque les données sont récupérées et que le catalogue est en cours de traitement, puis en =g0= lorsque le catalogue est dans VizieR.

- =s= tables <~100 objets astronomiques, non intégrées dans VizieR, dont les données peuvent intéresser SIMBAD, mais qui ne peuvent pas être traitées par DJIN (xids complexes et/ou données à entrer en mesures – voir [doc Twiki](#) pour les données intéressantes pour SIMBAD). Le statut est donnée par l'équipe DJIN. Le statut est modifié suite à la réunion =g= en =s1=, =s2= ou =L=.

- =q=Acro. Signale le besoin éventuel d'un nouvel acronyme, ou une demande de renseignement à propos d'un nouvel acronyme. Le statut est donnée par l'équipe DJIN (parfois VizieR ou COSIM) Il est à destination de l'équipe Dictionnaire. Ce statut est à utiliser par défaut en cas de doute, pour une demande de renseignement sur un acronyme, nouveau ou pas.

1. Pratiques à ce jour (oct. 2023) pour l'équipe DJIN

1.1 =e=

Cas différents selon les journaux :

- **A&A**
 - Très peu de =e= – car la plupart des tables sont envoyées directement à VizieR (pour les tables uniquement disponibles au CDS = sélection des éditeurs) – mais il y en a quelques uns.
- **AAS (ApJ+ApJS+AJ) et MNRAS**
 - Si MRT (machine-readable table) ou fichier .zip/.tar.gz dont le contenu n'est pas visible : peu importe la taille de la table, le statut =e= est mis
 - Si table avec >~100 objets difficilement identifiables et des données intéressantes pour SIMBAD
→ =e=

- Si table $>\sim 100$ objets mais faciles à traiter dans SIMBAD, on traite la table mais on signale tout de même la table pour qu'elle soit examinée par les équipes suivantes concernées.

A noter : les tables en MRT ne sont pas complètes dans les articles (l'article contient un extrait et la table complète est dans un fichier à part), on ne connaît pas le nombre d'objets d'une table MRT à moins d'aller l'ouvrir et elles ne sont pas analysées via DJIN. De même pour les tables .zip

1.2 =s=

- les « petites » tables ($<\sim 100$ lignes) contenant trop d'objets à créer et/ou de données complexes pour être faites via DJIN.
- Données intéressantes pour SIMBAD (voir le rappel) → =s=

1.3 =q=Acro pour les tables

Si les objets astronomiques d'une table en =e= ou =s= nécessitent un acronyme (nouveaux objets sans ID connus) le statut =q=Acro ou =q=AAcro (acronyme déjà en attente sur une autre référence) est mis au moment de la sélection de la table la plupart du temps.

L'équipe DJIN ou du Dictionnaire peut également ajouter un statut =a0= : besoin d'acronyme repéré mais dont la création est reportée.

Ce statut doit normalement être suivi d'un commentaire explicatif.

A ce jour (oct. 2023), l'équipe DJIN ne l'utilise que pour les objets qui ne peuvent pas être rentrés dans SIMBAD, typiquement pas de coordonnées ou en attente de recitation.

2. Pratiques antérieures et à ce jour (oct. 2023) pour l'équipe Vizier

Est-ce qu'il vous arrive d'enlever des =e= parce que la table ne sera pas mise dans Vizier ? Si oui sur quels critères ?

Rappel/précisions : différentes pratiques selon les époques de traitement et les journaux : Pendant longtemps, toutes les tables de $>\sim 100$ lignes étaient traitées. Ce n'est plus le cas depuis au moins 5-6 ans.

Le =e= est un statut « en attente ». Il signifie que les tables ont été vues (a priori et si pas de backlog, cf. MNRAS et ApJ) mais qu'elles n'ont pas été jugées prioritaires. Dans ce cas là, le =e= n'est pas supprimé et on espère revenir dessus (un peu comme pour les g2)...

Si on décide de rejeter le =e=, le statut devient (e) et il y a généralement une explication pour le rejet.

Lorsque les tables ne sont pas accessibles (tables complètes introuvables, besoin de mot de passe, sites inexistant, ...), on met un =e0= avec un commentaire.

- **A&A**
Très peu de =e=... Traités pour la plupart sauf si les données ne sont pas récupérables.
- **MNRAS** – Backlog : traitement actuel des journaux publiés début 2021.
On fait ce qu'on peut ? - peu de rejet a priori.
- **AAS (ApJS+AJ)**
Journaux à jour – « pas trop » de =e= (même si le temps de traitement pour l'ApJS est beaucoup plus long que pour les autres journaux). Très peu de rejet ; traités pour la plupart sauf si les données ne sont pas récupérables ou données atomiques/simulations (à partir de 2019).

- **ApJ** – Backlog : traitement actuel des journaux publiés début 2022 (mais le backlog a pu être moins important avant 2019-2020...)

En ~2018, traitement prioritaire pour les MRTs qui sont préparées par les auteurs+éditeurs AAS (pression des éditeurs pour qu'on entre plus de MRTs). Cette année-là, nous avons noté l'augmentation de la charge de travail et modifié nos critères de sélection en conséquence – voir [CR de réunion du 7 mai 2019](#).

Depuis cette date, les critères de sélection ont évolué : les données atomiques, les modèles/simulations etc. sont passées en priorité 2 ; y-compris pour les MRTs. La priorité est mise pour les données d'observation et/ou les tables contenant des données intéressantes pour SIMBAD (peu importe le format de la table d'origine).

Depuis 2019, l'équipe Vizier dispose également du statut =e2= (priorité 2) mais qui ne concerne que les tables contenant des objets astronomiques qui sont déjà référencés pour SIMBAD via DJIN et qui ne contiennent pas de nouvelles mesures provenant des auteurs.

Problème : les tables laissées en =e=, =e2= ou (e) ne remontent pas vers l'équipe COSIM, ce qui peut provoquer des problèmes pour SIMBAD.

3. **Décision :**

On abolit la règle **MRT** ⇒ =e= et on précise la règle :

tables de moins de <~100 objets (+- ~10) → =s=

Pour cela, il faut dénombrer facilement le nombre d'objets des tables à récupérer...

=> Lorsque les tables sont dans des fichiers illisibles, le =e= reste appliqué.

=> Anaïs propose d'afficher automatiquement le nombre d'objets des tables MRTs. Il faut un développement au niveau de DJIN.

4. =q=Acro

La question « comment réduire le nombre de =q=Acro » ? a été soulevée.

« Pas de solution miracle » au vu des quantités traitées.

En plus de déjà distinguer les =q=Acro des =q=AAcro, les djinistes sont invités à mettre également plus de =a0=.

Ils devront concerner les acronymes moins prioritaires.

Qu'est-ce qu'un « acronyme moins prioritaire » ?

=> Cela nécessite l'expertise de l'équipe Dictionnaire ? A discuter en réunion =g= lorsque les cas se présentent ?