

Metadonnées disciplinaires

l'International Virtual Observatory Alliance



F. Genova, Journées Métadonnées 25/11/2004



1

Pourquoi conserver les données

- Objectifs scientifiques majeurs
variabilité/évolutions, statistiques, échelles spatiales/temporelles, couplages,...
- Données multi-échelles, multi-capteurs
- Optimisation du retour scientifique (exigence des agences)

Des données de toutes origines doivent pouvoir être utilisées par toute la communauté ou des communautés différentes qui collaborent

- Informations massives, hétérogènes, distribuées
observations, résultats, modèles, compilations, ...
- Réutilisation \Rightarrow **données \oplus documentation**



F. Genova, Journées Métadonnées 25/11/2004



2

Tirer parti des évolutions techniques

- Capacités de stockage et gestion de l'information
- Irruption fracassante du WWW
 - Permet de rendre les informations accessibles
 - Intégration données/documentation
 - Navigation
- ...vers la Grille, de ressources (calcul, stockage), mais aussi de services

Mais sont indispensables

- Travail de fond sur le contenu et les fonctionnalités (validation, maintenance)
- Définition des métadonnées et standards d'échange



Les acteurs

- Agences
 - Conservation, diffusion des données
- Projets
 - Données disponibles et utilisables
 - sélection, mise en forme, « mémoire »
 - Prise en compte dès le début des projets
- Journaux
 - Résultats des recherches
- Centres de données
 - Bases de données à valeur ajoutée, outils généraux et spécifiques



Données d'observations

Des dizaines de To d'observations

- très hétérogènes (images, spectres, séries temporelles, cubes de données)
- Distribuées dans des archives d'observatoires tout autour du monde

Et aussi des services (cf CDS)



F. Genova, Journées Métadonnées 25/11/2004



5

La mise en réseau des informations



Observations

Observatoires sol et spatiaux
Grands relevés

Résultats

Publications dans les journaux
électroniques



F. Genova, Journées Métadonnées 25/11/2004



6

Les standards en astronomie

- Une petite discipline
- Habitude de définir des standards
 - Au départ un petit groupe expérimenté qui a besoin d'échanger de l'information
 - Effet boule de neige
 - Ex: FITS, bibcode, ReadMe (catalogues, tables publiées, logs d'observatoires, grands relevés)
- Mise en réseau des informations dès le démarrage du WWW



F. Genova, Journées Métadonnées 25/11/2004



7

Retrouver, intégrer l'information

- Description des données (métadonnées)
- Standards d'échange
- Exemple: les données tabulaires
ReadMe
 - les tables publiées sont devenues des données*
- Un standard international: CDS, journaux, autres centres de données



F. Genova, Journées Métadonnées 25/11/2004



8

Le standard de description des données tabulaires

Table = Catalogue, Table publiée, Grand relevé, 'log' d'observatoire

The screenshot shows a web browser window with the URL <http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A&2BAS/134/115>. The page content includes:

J/A&2BAS/134/115 Medicina 6.70GHz methanol maser survey (Slyzbb 1992)

The Medicina survey of methanol masers at 6.7 GHz
 Slyzbb U.I., Val'kov I.E., Ralchenko S.V., Usenkov M.A., Palagi F.,
 Tideni G., Cavazzi M.
 <Astron. Astrophys. Suppl. Ser. L14, L15 (1999)>
 #1999AAS...134...115P

AD Keyword: Maser ; Surveys ; Radial velocities ; Radio sources

Keywords: maser - ICM; molecules ; H II regions - radio lines: ICM

Description:
 File table2 contains the name of the sources and their coordinates (1950) observed, but not detected, with the Medicina Radiotelescope at 6.7GHz in 1995, March. Source are associated with one of nine type of phenomena typical of star forming regions (see note 3 for code description). The number of non detected sources is 387.

File Summary:

FileName	Line#	Records	Explanations
* ReadMe	60	-	This file
* table1.dat	54	387	List of undetected sources, the typical detection limit (original) level is 0Jy.

Byte-by-byte Description of file: table1.dat

Byte	Format	Unit	Label	Explanation
1- 12	A12	---	Name	Name of the source
14- 15	I2	h	RAh	Right ascension (1950.0)
17- 18	I2	m	RAm	Right ascension (1950.0)
20- 22	F3.1	s	RA ^s	Right ascension (1950.0)
25	A1	---	DE-	Declination sign
26- 27	I2	deg	DEd	Declination (1950.0)
29- 30	I2	deg	DEm	Declination (1950.0)
32- 33	I2	deg	DEs	Declination (1950.0)
35- 39	F5.1	km/s	PV	LSR radial velocity
41- 54	A14	---	Type	Type of associated phenomena (1)

Note (1) the following abbreviations are used:



Une vision commune des données tabulaires publiées dans tous les journaux

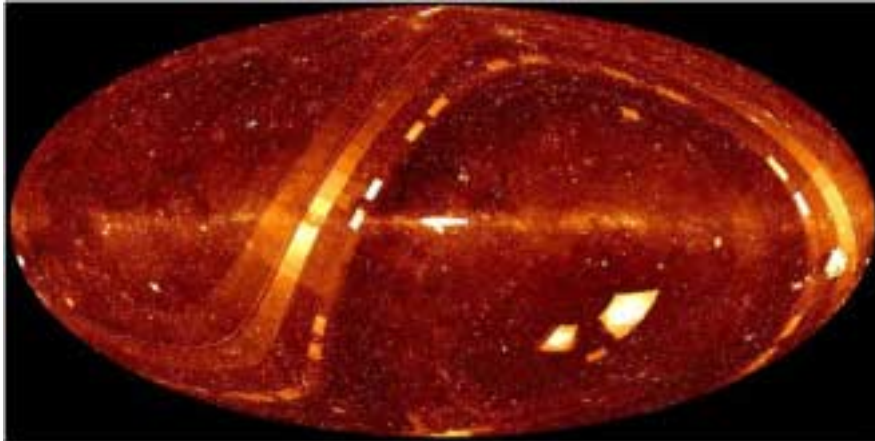
The screenshot shows a website with the title "Catalogues and files available at CDS" and the version "Version of 17-May-2003". The page lists the following categories and counts:

- B. Copies of external databases, regularly updated. (13 catalogues)
- I. Astrometric Data (240 catalogues)
- II. Photometric Data (219 catalogues)
- III. Spectroscopic Data (190 catalogues)
- IV. Cross-Identifications (20 catalogues)
- V. Combined data (97 catalogues)
- VI. Miscellaneous (38 catalogues)
- VII. Non-stellar Objects (201 catalogues)
- VIII. Radio and Far-IR data (69 catalogues)
- IX. High-Energy data (23 catalogues)
- Tables from *Astronomy and Astrophysics* (1039 catalogues)
- Tables from *Astronomy and Astrophysics Supplement Series* (1164 catalogues)
- Tables from *Astronomical Journal* (677 catalogues)
- Tables from *Astronomicheskii Zhurnal (Russian)* (73 catalogues)
- Tables from *Astrophysical Journal* (343 catalogues)
- Tables from *Astrophysical Journal Supplement Series* (436 catalogues)
- Tables from *Monthly Notices of the Royal Astronomical Society* (267 catalogues)
- Tables from *Publications of the Astronomical Society of the Pacific* (104 catalogues)
- Tables from *Pis'ma v Astronomicheskii Zhurnal (Astronomy Letters)* (67 catalogues)
- Tables from *publications from other journals* (142 catalogues)
- Catalogues ordered by their Usual Name (798 catalogues)
- Catalogues with Additional Material





Catalogues, tables publiées, très grands relevés,
liste des observations dans les archives
Liens vers les données



F. Genova, Journées Métadonnées 25/11/2004



11

Liens vers
données
complémentaires
(à Strasbourg ou
ailleurs)

CDS Catalogues with additional Material

[CDS](#) - [Simbad](#) - [VizieR](#) - [Aladin](#) - [Catalogues](#) - [Vizier](#) - [Bibli](#) - [StarPages](#) - [AstroWeb](#)

Click on a catalogue name or number to move to the corresponding VizieR page

1 Catalogues with [cube](#)

- [\(VII66\)](#) IRAM observations in pre-star forming regions (Falgarone+ 1998-2001)

2 Catalogues with [image](#)

- [\(VI76\)](#) Simulation Atlas of Tidal Features in Galaxies (Howard+ 1993)
- [\(VII28\)](#) Bell Laboratories HI Survey (Stark+ 1992)
- [\(VII69\)](#) the FIRST Survey, version 1999Jul (White+ 1999)
- [\(VII85\)](#) 1.4GHz NRAO VLA Sky Survey (NVSS) (Condon+ 1998)
- [\(VII71\)](#) The FIRST Survey Catalog, Version 03Apr11 (Becker+ 2003)
- [\(I161\)](#) HST Archived Exposures Catalog (STScI, 2003)
- [\(I162\)](#) ISO Observation Log of validated data (ISO Data Centre, 2001)
- [\(I163\)](#) The Chandra Archive Log (CXO, 2002)
- [\(I164\)](#) XMM-Newton Observation Log (XMM-Newton Science Operation Center, 2002)
- [\(I165\)](#) The MERLIN Archive (MERLIN 1992.)
- [\(I166\)](#) Morphological Types in 10 Distant Rich Clusters (Smail+ 1997)
- [\(I167\)](#) Soft X-ray properties of Seyfert galaxies (Fridman+, 2001)
- [\(I168\)](#) NGC 5548 UBVRi phot. and H α , H β , H γ fluxes (Dietrich+, 2001)
- [\(I169\)](#) ROSAT-BSC galaxy identifications (Zimmerman+, 2001)
- [\(I170\)](#) Stellar composition of CMa R1 (Starvchenko+, 1999)

3 Catalogues with [model](#)

- [\(VII82\)](#) Geneva stellar evolution tracks and isochrones (Lejeune+, 2001)

4 Catalogues with [spectrum](#)

- [\(II47\)](#) Scanzar Abundance in late-type evolved stars (Spinrad+ 1969)
- [\(III84\)](#) SKYLAB S-019 Far-UV Objective-Prism Spectrophotometry (Heizner+ 1979)
- [\(III85\)](#) Spectrophotometry of 1300 stars (Banasik+ 1982)
- [\(III143\)](#) Spectrophotometry of Wolf-Rayet Stars (Tonesi-Dubout+ 1982)



F. Genova, Journées Métadonnées 25/11/2004



12

L'Observatoire Virtuel astronomique

Des données d'observation aux résultats
publiés, en intégrant les très grands relevés

*« an enabling and coordinating entity to foster the
development of tools, protocols, and
collaborations necessary to realize the full
scientific potential of astronomical databases in
the coming decade »*

NVO White Paper, juin 2000



F. Genova, Journées Métadonnées 25/11/2004



13

L'Observatoire Virtuel

- **Piloté par la science** ('science driven')
- Dans une communauté qui a l'habitude de définir des standards d'échanges disciplinaires et du partenariat pour la mise en réseau des informations
- Tire parti des derniers développements techniques
- Peut tirer parti des 'priorités aux STIC'



F. Genova, Journées Métadonnées 25/11/2004



14

Interoperabilité

Accès à une information hétérogène, distribués

Utilisateurs

- Identifier parmi les données disponibles celles qui m'intéressent
- Comment interroger les services dont les données m'intéressent
- Comment utiliser les données que j'ai récupéré

Responsables de services/développeurs

- Comment un autre service peut-il utiliser mes données?
- Comment puis-je utiliser les données d'un autre service?



Coordination internationale

- International Virtual Observatory Alliance
<http://www.ivoa.net>
15 membres, derniers arrivés: Hongrie, Espagne
- Chaque projet a ses propres objectifs (qui dépendent des financements obtenus ou espérés)
- L'IVOA met en contact les projets (téléconfs régulières), définit une feuille de route (ex: calendrier de démos scientifiques), coordonne les discussions sur les standards
- Intense activité de définition de standards
Réunion interoperabilité, Strasbourg, Octobre 2003
120 personnes! (30 en janvier 2002); deux réunions/an
- Un Groupe de travail UAI pour 'adopter' les standards (cf FITS)





F. Genova, Journées Métadonnées 25/11/2004



17

<http://www.ivoa.net/>



F. Genova, Journées Métadonnées 25/11/2004



18

Définition des standards et des méthodes

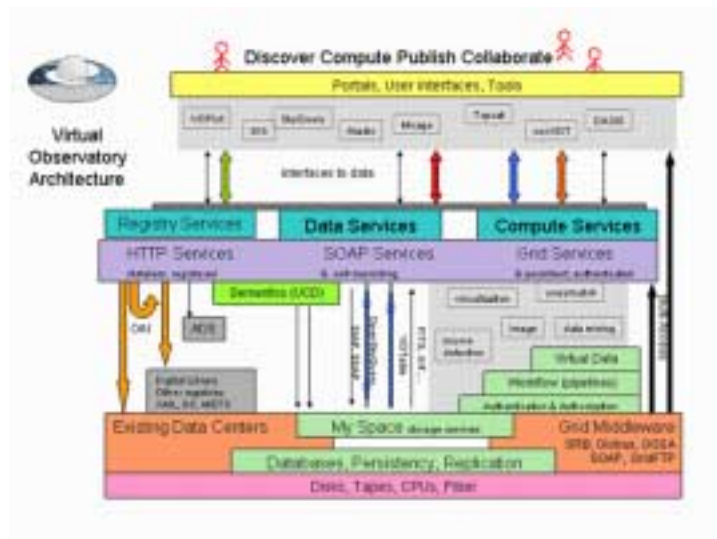
- Groupes de travail de l'IVOA
- Site collaboratif Twiki
- Listes de diffusion – parfois très (trop?) actives
- Réunions deux fois/an
- Une procédure d'adoption des standards dérivée de celle du GGF et simplifiée
 - Standards évolutifs
 - Adoption finale par un groupe de travail de l'Union Astronomique internationale (en cours de constitution)



Lien
'Community'



Lien
'Documents'



IVOA WGs

- DM (Data Models)
- WS (Web Services) & GRID
- Registry
- DAL (Data Access Layer)
- VOQL (Query Language)
- UCD (Unified Content Descriptors)
- VOTable
- Interest Group Theory
- Interest Group Applications



Démonstrations: Utilisation des standards

- Les projets de l'IVOA organisent des démonstrations annuelles, phasées
- Démonstrations de prototypes qui permettent de valider les standards et de proposer des évolutions
- Par exemple: données complexes, hétérogènes, distribuées >> le modèle de données et la couche d'accès aux données doivent permettre de hiérarchiser les données





'First Science' Demo

- Prototype VO tools for science
 - Led to Discovery of Type 2 QSOs – A&A 2004
- Enabled by real gains in standards for:
 - Data access
 - Manipulating image and catalogue data
- Employed *many VO tools* to access and use VO compatible data from many sources



F. Genova, Journées Métadonnées 25/11/2004



25

2MASS
ESO-WFI
Chandra
VLT-ISAAC
HST-ACS
DSS
My Data

Data available at selected point are highlighted in tree

Field of view outlines are plotted automatically

Image metadata