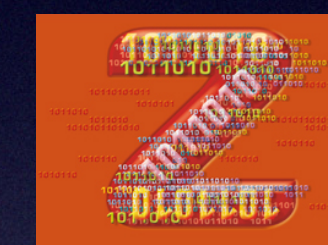




MASse de données En **ASTRO**nomie et **ASTRO**physique (Massive data in Astrophysics)

Coordinators : S. Maabout , C. Surace



DISCUSSION



- National/international context

- PETASKY
- ASOVF

- BigSkyEarth
- ASTROInformatics
- OV

- Qualified data

- EROS2
- EXOPLANETS data
- LSST (test/Use Cases) DR9-SDSS.
- GAIA (Sim)
- VO (accessibilité Vizier)

- Platforms, systems

- Systeme BD classiques (objet, relationnelles, hierarchique ,...- (instancié par QSERV...))
- HADOOP /SPARK- MAP REDUCE
- GALACTICA plateforme de tests et benchmark (origine PETASKY)
- CC-IN2P3 (Tests et benchmarking PanStarr) (tests de passage à l'échelle)

- Visualisation

- GLNEMO2
- TULIP
- ALADIN (HIPS)



DISCUSSION



- Themes :

- Data Storage

- Query processing

- Data search
- Data Acces

- Data processing/mining

- AstroStats

- implication of data mining technics to data search/query



What did you expect ?

- **Goals of MAESTRO :**

Organise and build network with consistent and co-working community of at least two institutes.
Researchers and engineers for the best methods to be used in Astrophysics

- **Proposal :**

- Disposal of astrophysical data to be tested and technical platforms
- Disposal of basic and more complicated queries for astrophysical needs

- Enhance exchanges of experience and methods. Hadoop, Spark, SciDB, Vertica, ...

- Data Mining in Astrophysics (sampling and algorithms for better selection)
- Exploration (partial results) versus data interrogation (complete results).
- Visualisation of big data in Astrophysics.
- ... open to ideas
 - Teaching

Survey

- Expectation ?
 - technical
 - GALACTICA ? How to use
 - astrophysical
 - Data and data access
- Teaching :
 - School Of Statistics
 - AstroInformatics
 - Listing of AstroStat teachers
 - Listing of Astroinformatics teachers
- Astrostat -> Bigdata
 - Astroinformatics
- Site of material disposal / tools
- Financial support ?

Discussion

- Utilisation des outils assez difficile.
 - Formation - Webinar
 - Comment adapter les outils aux utilisateurs. (adapter les outils, définir interfaces)
 - Avoir un point de contact et support Outils-Astro (AstroTulip, Glnemo2Fits)
 - cela prend du temps
- Outils perennnes
- GLNEMO2 : Pb de documentation pour utilisation.
 - FITS : comment visualiser les données.
- TULIP : arbre de données
 - Quoi faire avec un arbre de données.
- Teaching
 - Utiliser TULIP oui !!, mais il faut connaitre les graphes-> qu'apportent les graphes au domaine astrophysique.
 - Astrophysique : Domaine de « problématiques » pour des recherches en informatique.
- Interactions ; Il est nécessaire d'avoir un contact Métier/stat/mining pour reconnaître les choses / outsiders importantes

Discussion

- La technologie avance
 - La technologie d'analyse repose sur les modeles des données (pour faciliter les analyses)
 - Adapter des techniques aux modèles de données
- L'OV a une finalisation de standardisation de mots clés pour décrire les données (plus que FITS)
- c'est une description interoperable des données.
- indexation spatiale des données -> apport par rapport au traitement des données
 - quelle est la bonne indexation pour différentes requetes
 - Cone-Search
 - autre
- **implication of data mining technics to data search/query**

Actions

- Définir un dictionnaire ASTRO - INFORMATIQUE
 - UCD -
 - Data Model -
 - Descriptions -
 - Données -
- Définir les avantages nouvelles technologies ou paradigmes (ie: SPARK) dans le passage à l'échelle.
 - Bonnes pratiques. (plutot applicatif)
 - Classe de problèmes traités par ie : SPARK (plutôt théorique)
- Définir des problématiques astrophysiques
 - machine Learning-> extraction de signaux. (ie)
- Définir une roadmap pour pérenniser les Ecole d'été Astroinformatics/Astrostatistics (?) (1 année sur 2)
 - Algorithmique approximative
 - Echantillonnage

Actions

- Liste des cours ou cours d'Astrophysique pour les informaticiens
- Liste des cours BigSkyEarth
- Liste des possibilités de publications (dans quels série) (A&A T. Forveille - IPAG , Nabila Aghanin)
 - TLDKS (Springer-Kader) numéros thématiques
- Listes de personnes references capable de faire des tutorials
- Reference vers des enregistrements et des videos MOOC.
- Sondage de participation aux groupes de travail.

Collaboration

- Astroinformatics
- CDS / OV
- RDA

mardi 5 juillet 2016

HEURES	ÉVÉNEMENT
14:00 - 14:15	Bienvenue - Direction de l'IRIT
14:15 - 16:00	Astrophysique et big data
14:15 - 14:30	› présentation MAESTRO - <i>Christian Surace, Laboratoire d'Astrophysique de Marseille</i>
14:40 - 15:30	› Big Data in Astronomy: astroinformatics and data driven approach - <i>Giuseppe Longo - Università degli Studi Federico II, Napoli (I), Giuseppe Longo</i>
15:30 - 16:00	› PETASKY - <i>Farouk Toumani, LIMOS</i>
16:00 - 16:30	Pause
16:30 - 17:30	Visualisation (Salle des thèses)
16:30 - 17:00	› TULIP : outils de visualisation - <i>D. Auber LaBRI</i>
17:00 - 17:30	› GLNEMO2 - <i>Jean-Charles Lambert, Laboratoire d'Astrophysique de Marseille</i>

mercredi 6 juillet 2016

HEURES	ÉVÉNEMENT
09:00 - 09:15	Accueil Café - accueil café
09:15 - 10:45	Méthodes / cas d'usage (Salle des thèses)
09:15 - 09:45	› Les données EROS-2, un cas d'étude pour l'archivage et la fouille de données - <i>Jean-Baptiste Marquette, Institut d'Astrophysique de Paris</i>
09:45 - 10:15	› Utilisation d'une plateforme Hadoop/Spark pour des données astrophysiques - <i>Christian Arnault, Laboratoire de l'Accélérateur Linéaire</i>
10:15 - 10:45	› Ecole d'Astrostatistique - <i>D. Fraix-Burnet -Univ Grenoble</i>
10:45 - 11:00	Pause café
11:00 - 12:30	Astrophysique et big data (Salle des thèses)
11:00 - 11:30	› from CDS to IVOA and RDA : Big data in Astrophysics - <i>Francoise Genova, cds</i>
11:30 - 12:30	› Discussion sur les attendus de MAESTRO - <i>Christian Surace, Laboratoire d'Astrophysique de Marseille</i>
12:45 - 13:45	Déjeuner - Déjeuner
14:00 - 15:00	Discussion - échanges (Salle des thèses)
15:00 - 16:00	reunion du comité de pilotage MAESTRO - Réunion du bureau de MAESTRO (reunion restreinte) S. Maabout, C. Surace, E. Ganglier, N. Novelli, P. Poncelet, F. Toumani F. Genova



Pour recevoir les infos
Inscrivez vous

[http://www.madics.fr/intranet/
reinscription/](http://www.madics.fr/intranet/reinscription/)

En précisant l'intérêt pour MAESTRO

Merci à l'IRIT d'accueillir les journées MAESTRO
Merci à l'équipe Pyramide pour l'organisation