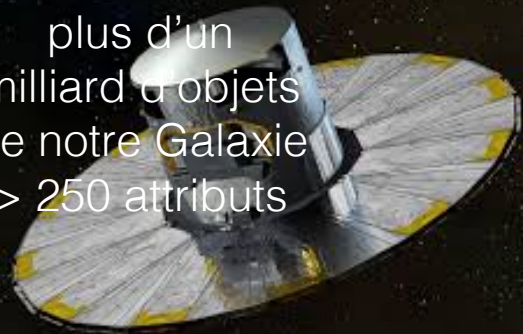




MADICS

Christian Surace  
LAM  
Sofian Maabout  
LaBRI

# Données Astronomiques



GAIA  
plus d'un milliard d'objets de notre Galaxie  
> 250 attributs



Pan-Starrs  
6,000 deg<sup>2</sup> per night.  
several terabytes per night for the full telescope.  
hundreds of Peta bytes.



LSST (2020) : 3 Billions pixels every 17 seconds.  
140 petabytes



SKA TELESCOPE  
**SQUARE KILOMETRE ARRAY**  
Exploring the Universe with the world's largest radio telescope  
Choose your local minisite



generate an exabyte a day of raw data,

Rapprocher les  
partenaires  
astrophysiques et  
informatiques



### Données Astrophysiques

Mise à disposition de données  
astrophysiques

Mise à disposition de requêtes type

Favoriser les échanges d'expérience et de  
methodes. Hadoop, Spark, SciDB, Vertica,

Data Mining en Astrophysique (sampling  
and algorithms for better selection)

Exploration des données, visualisations  
versus interrogation directe)

Programme Vendredi 23 Juin 2017 (25 participants / 13 laboratoires / 3 instituts)

#### Introduction

10h00 : MAESTRO (présentation) Surace./Maabout

10h10 : Plateforme Galactica - Frédéric Gaudet

10h30 : Evolution of Data Management Systems for Big Data Applications - Abdelkader Hameurlain

#### Données Astrophysiques :

11:00 SKA : Chiara Ferrari

11h30 LSST : E. Gangler

12h00 Lunch

#### Machine Learning

13h30 : Exploring the spectroscopic diversity of type Ia supernovae with DRACULA: a machine learning approach - Emille Ishida

13:45 : Classification of reliability for redshift measurements - Sara Jamal

14h00 : Clustering pour détection d'amas ouverts avec Gaia - Mario Morvan

14h15 : Galaxy morphology with CNNs using transfer learning - Alexandre Boucaud.

#### Organisations

14h30 BIGSKYEARTH (présentation) - E. Gangler/ K. Zeitouni

14h50 Cosmostatistics initiative : (<https://asaip.psu.edu/organizations/iaa/iaa-working-group-of-cosmostatistics>) - Emille Ishida

15h10 Wrap up, sondage, conclusions

16h00 fin du meeting



Mise à disposition de données astrophysiques

Mise à disposition de requêtes type

Favoriser les échanges d'expérience set de méthodes. Ex: Hadoop, Spark, SciDB, Vertica, ...

Data Mining en Astrophysique (sampling and algorithms for better selection)

Exploration des données, visualisations versus interrogation directe

COS BigSkyEarth, CosmoStatistics initiative

A travers l'infrastructure Galactica, SKA, LSST

Pas totalem abordé

Présentations clustering/classification mais pas algo approchée

Pas du tout abordé

# Bilan

- Réseautage:
  - ❖ Meilleure connaissance du tissu national/international → Aide au montage de l'école thématique/intervenants
  - ❖ Réponses communes à l'appel PEPS
  - ❖ Sollicitation du CNRS pour une journée d'étude sur actions à mener en astro-informatique
  - ❖ Sollicitation pour faire partie de comités scientifiques
- Manque de formations spécifiques à l'image de la bio-informatique

# Actions envisagées pour 2018

- Journée(s) Maestro avec Focus: technologies/visualisation
- Ecole thématique
- Proposer aux lauréats de l'appel PEPS Astroinformatique de présenter leurs travaux lors d'une (ou de) journée(s) dédiée(s).
- Proposer un atelier en association avec la conférence EGC

# Astro Info

## Ecole thématique 2018

Le traitement de la demande au CNRS est en cours.  
Premiers retours positifs.



# programme

Module	Intervenants (nom et qualité)	Forme (cours, atelier, TD, TP...)
Données Astrophysiques	3 intervenants + 1 étudiant ou post-doc	Cours, TP, utilisation d'outils de visualisation de données, VO
Le Machine learning	3 intervenants + 1 étudiant ou postdoc	Machine learning, Deep Learning
Les plateformes	3 intervenants + 1 étudiant ou postdoc	OV, GALACTICA, LAM
Astro Hackathon	3 projets, 6 intervenants	Projets pour mettre en commun les connaissances

- Les journées « Cours » se dérouleront de la façon suivante : 1 heure d'initiation sur les concepts de base du cours, 1h30 d'interventions, sur le sujet et 3 heures de Travaux Pratiques.