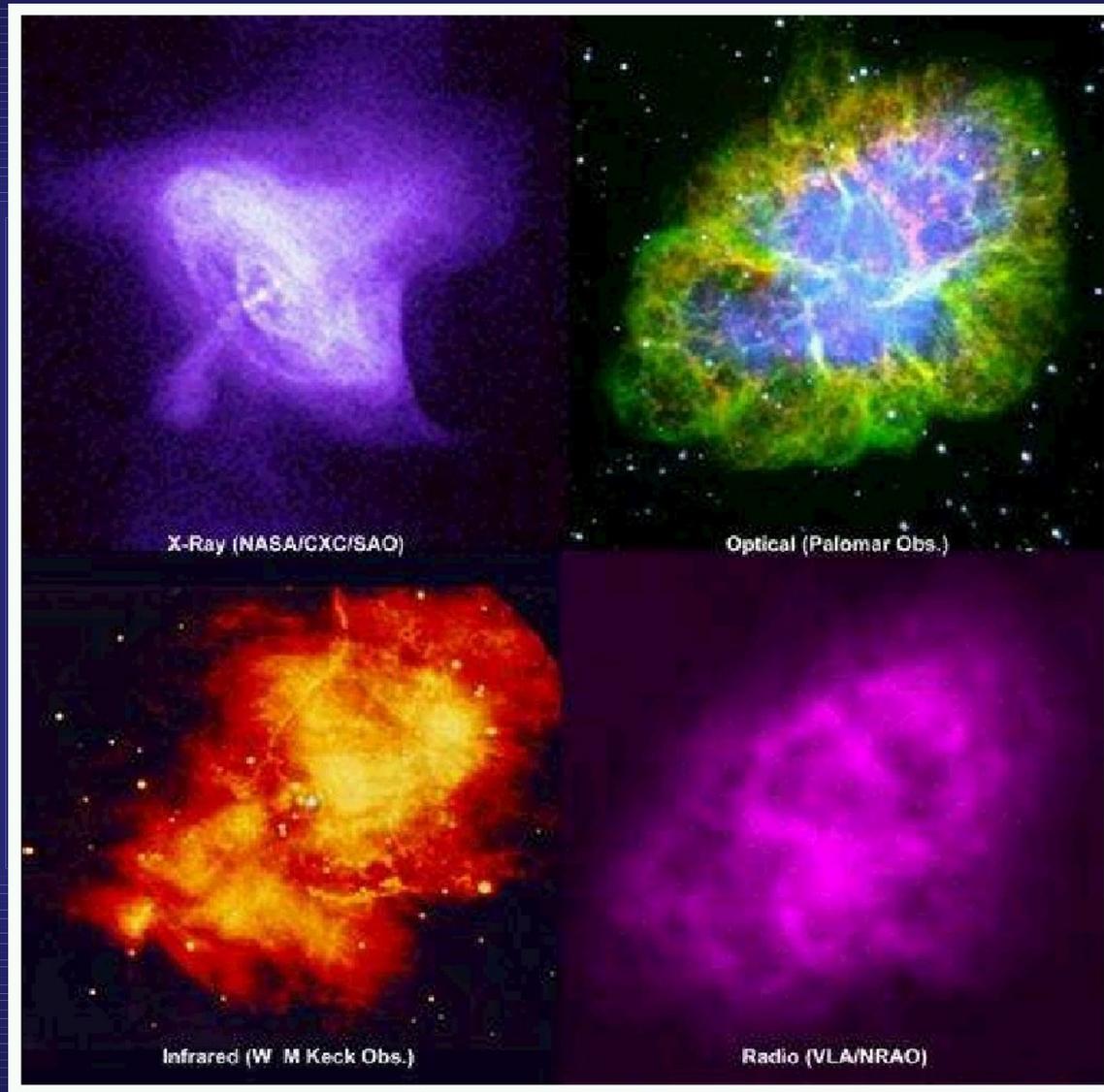


# Utilisation d'un *Workflow* en Imagerie

collaboration AS OV : CDS / ESSI / OCA

# Plus-Value d'une Fédération de l'Information



# *VO Science Cases*

(from Astro Grid Top Ten)

- Processus physiques au sein de la couronne solaire ;
- Intercomparaisons entre les *flares* solaires et stellaires ;
- Détection de naines brunes ;
- Environnement des supernovae ;
- Recherche de galaxies à faible brillance de surface ;
- Relevés extragalactiques profonds ;
- Identification des amas de galaxies ;
- Découverte de quasars à très haut décalage spectral.

# Distribution à grande échelle des galaxies

## • Effets d'environnement •

- **identification** des amas de galaxies à partir de relevés comprenant plus de couleurs (optique + IR) → précision accrue ;
- **croisement** avec les données spectroscopiques, l'émission X et radio ;
- **diagnostiques** sur l'état dynamique (structures) :
  - densité projetée et dispersion de vitesses de l'échantillon de galaxies ;
  - émissivité, température, entropie et métallicité du plasma intra-amas.
- > processus non gravitationnels : ondes de chocs, turbulence, MHD ... ;
- > couplages et rétroactions sur la formation stellaire (cf. morphologie) ;
- > comparaisons aux simulations : corrélations spatiales,  $N_M(z)$ , ...

# Quelques Requêtes Pertinentes

## • Analyse de catalogues

- croisements entre catalogues (contreparties absentes ou multiples) ;
  - distribution spectrale d'énergie, *redshifts* photométriques
- écarts à l'homogénéité, relations de hiérarchie entre les structures ;
- filtrages adaptés liés à la régularité constatée des propriétés ;
- comparaisons aux modèles et aux bases de données (images, spectres).

## • Analyse d'images multibandes multicapteurs

- détection optimale d'objets de forme plus ou moins connue et stable ;
- fond variable, superposition d'objets, emboîtement de composantes ;
- caractérisation du bruit, échantillonnage et résolution, magnitude limite, cohérence des mesures entre bandes, mosaïques.

> **Requêtes couplées** < e.g. *QSO host galaxies & BH masses*

# Une Réalité Encore Prégnante

- des archives très hétérogènes vis-à-vis de la description des données, de leur qualité et de leur calibration ;
- la fusion de telles données distribuées est dès lors pénible ;
- accroissement exponentiel du nombre de pixels ;
- complexité intrinsèque d'une information multicapteur multi- $\lambda$ ;
  - > besoin d'algorithmes et de méthodes de traitement adaptés <
  - > accès à des chaînes d'analyse éprouvées <
  - > performances en traitement massif <

# Analyse d'Images Multivariées

- complexité des données multidimensionnelles •  
(non-gaussianité, dilution, corrélations)

## 1) réduction de la dimensionnalité

- choix d'un espace de représentation *discriminant*
- adoption d'un modèle de bruit

– approches globales linéaires ou non vs. recherche des sous espaces d'intérêt –

## 2) segmentation des données

- sélection de modèles (cf. représentation compacte)
- choix d'une mesure de distance ; estimation de PDFs

– approches supervisées ou non, hiérarchiques ou pas, avec ou sans frontières floues –

## 3) extraction de primitives et classement

# Fusion de l'Information

## Objectifs

- **combiner** de façon *optimale* toutes les informations dans un modèle unique
- **préserver** la *totalité* de l'information contenue dans le jeu de données
- **améliorer** la qualité du produit final  
(débruitage, super-résolution spatiale et/ou spectrale)

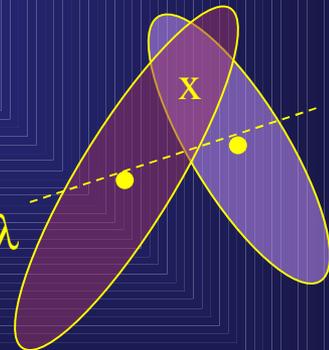
## Stratégie

- une série d'images d'un même champ ;
- une collection de spectres ;
- un cube tridimensionnel.

## Tactique

ex. co-addition : image Khi2 vs. approche probabiliste

marginale *a posteriori* :  $P(X | Y) \propto \int P(X, Y, \omega, \lambda) d\omega d\lambda$



# Systemes d'Analyse de Données

- > pérenniser le savoir-faire et gagner en efficacité <
- > rendre l'expertise accessible au plus grand nombre <

requête (but) sur les données → résultat (objectif)

## 1) Utilisation d'une *bibliothèque* de programmes

mise à disposition uniformisée des outils  
construction graphique d'un enchaînement de tâches

Application AIDA

Techniques de grille : GRID5000, EGEE

## 2) Pilotage supervisé de programmes (ex. moteur PEGASE /INRIA)

un plan d'action à la volée à l'aide d'opérateurs et de règles

# Une Requête Simple ...

- **opérations :** détection et évaluation d'objets connexes dans une image mono-bande (SExtractor).
- **modèle sous-jacent :** taches diffuses disjointes en émission sur un fond lentement variable sans défauts.
- **méthode adoptée :**
  - cartographie du fond ;
  - seuillage par segmentation ;
  - ajustement d'une ellipse de forme ;
  - évaluation du profil azimuthal de brillance ;
  - calcul de mesures de forme et de flux.



# Quelques Perspectives

- Fusion d'images multibandes, multicateurs, multimodales
  - ANR *Space Fusion* / LSIT Strasbourg.
- Extraction et mesures de formes floues emboîtées
  - un premier *webservice* en cours de développement à l'OCA.
- Classification et fouille sur des données massives multidimensionnelles
- Gestion, croisement et exploitation de (très) grands catalogues
  - > OV : migration des données vs. migration des exécutables
  - > Grille : état et suivi de la charge, barrières de synchronisation
  - > AIDA : ergonomie, suivi et journal d'exécution, scénario partiel