HiPS resource provenance

F.Bonnarel, Pierre Fernique (CDS)

Mireille Louys (Icube, CDS)

What is HiPS ?
(Hierachical Progressive surveys)

- New All sky organisation of pixel data
  - Based on HEALPIX tesselation (projection of Sky on equal area pixels -tiles or diamonds)
  - Adaptative resolution from allsky to full original resolution based on storing of a hierarchy of tiles at different orders
  - Storing and organization based on a tree of including directories for each order

- HiPS is going to become an IVOA standard (Application WG)
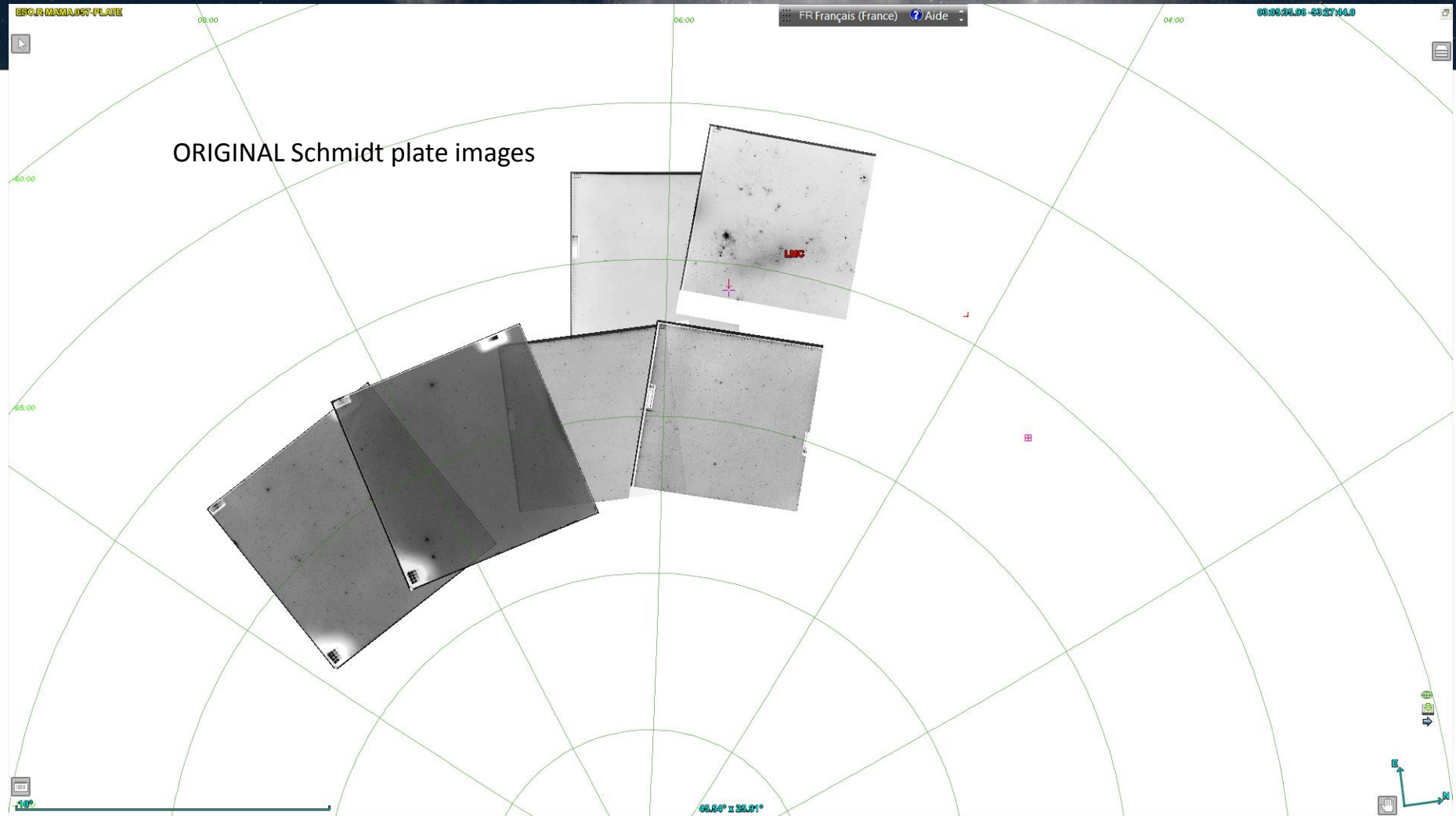- HiPS is a kind of « legacy level » or « very high data organization level » for observational data.

Generating HiPS from « original » reduced pixel data

- Can be done with Aladin in « hipsgen » mode
- Can be done by other software (have to respect the organization standard)
- Main method differences are due to value estimation in each cell
  - FIRST (NEAREST)
  - MEAN
  - MEDIAN
- Other parameters : in, out, border, fov, skyval, param, etc…. (about 50)
- Parameter file : a config file with a list of param=value statements
- Examples of commands:
  - java -jar Aladin.jar -hipsgen in=/MyImages bitpix=16 pixelCut="-1 100 log
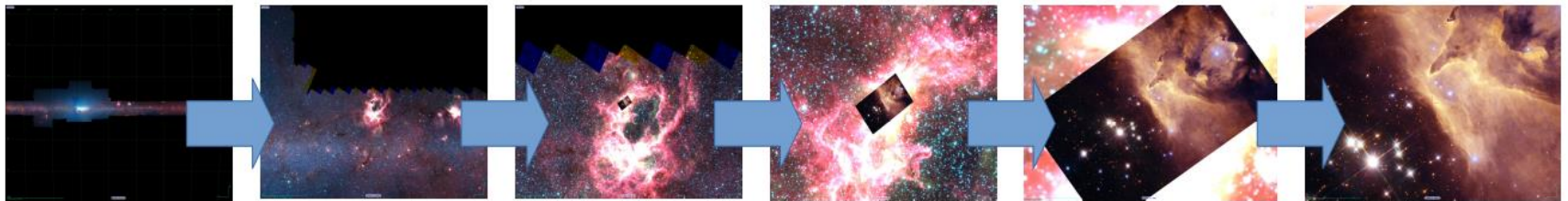 - java -jar Aladin.jar -hipsgen in=/MyImages –param=configFile

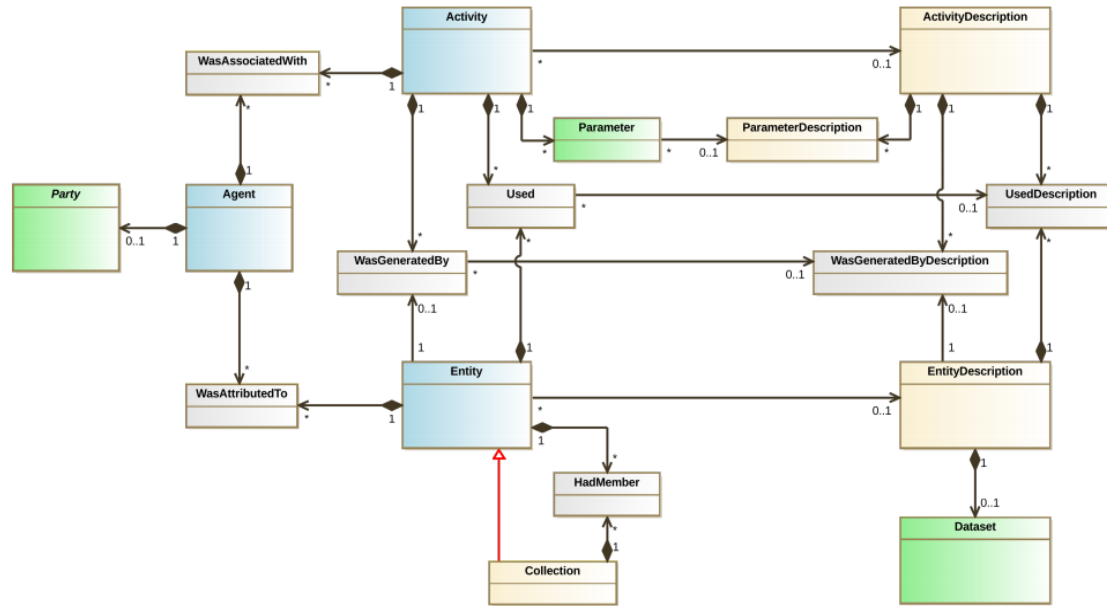# Original data: input HiPS images

ORIGINAL Schmidt plate images

SDSS HiPS allsky view

HiPS Schmidt plate All Sky view

Provenance UML diagram

The provenance of a HiPS dataset. Motivation

- HiPS comes with a « properties » file allowing to describe:
  - the HiPS generation from original dataset
  - some characterisation of the collection (see on slide 5).
- HiPS can be discovered via
  - Registry
  - Aladin
  - HiPSlists at some registered places
- So why do we really need a Provenance ?
    - A good testbed to see if Provenance is able to tackle processes far from the original use cases
    - Include HiPS generation description in more general discovery of datasets by provenance

Describing an HiPS provenance in IVOA/Prov-DM formalism

- Entity (ivo://CDS/P/MAMA/ESO-R,
    [
    prov:Label = « ESO-R MAMA HIPS at CDS »,
    prov:type = « voprov:hips_pixels »,
    prov:description = « This is the HiPS version of ESO Schmidt survey digitized by Mama and processed by CDS »
    prov:description = « ivo://ivoa.net/std/HiPS »,
    hips:HiPS_properties = « http://cds.u-strasbg.fr/hips/p/mama/eso-r/properties.txt »
    ])

- EntityDescription (« ivo://ivoa.net/std/HiPS »,
    [
    prov:Label = « HiPS generic description »,
    prov:Description = « This is hook for the description of te HiPS data format and organization »,
    voprov:URL = « http://cds.u-strasbg.fr/hips/documentation.html#structure »,
    hips:Data_productype:= « voprov:hips_directory »,
    hips:Level : 3
    ])
    // Relationship
- WasAttributedTo(« ivo://CDS/P/MAMA/ESO-R », « ivo:cds », prov:role= « voprov:creator »)
- Agent(« ivo::/cds » ,[voprov:Name= « CDS »,voprov:contact = « question@astro.unistra.fr »])
- WasAttributedTo((« ivo://CDS/P/MAMA/ESO-R », « ivo:cds », prov:role= « voprov:creator »))

Describing an HiPS provenance in IVOA/Prov-DM formalism

- WasGeneratedBy («ivo://CDS/P/MAMA/ESO-R », « EHG1 », prov:role = «hips:HiPSgeneration »)
- -> Activity(« EHG1 »,
    [
    Prov:Label = « ESO HiPS generation 1 »,
    Prov:startTime = « 2016-07-18 »,
    Prov:endTime = « 2016-07-20 »,
    Voprov:Description = « this activity is final generation of HiPS for ESO Mama survey »,
    Voprov:Description= « HipsgenM »
    ])

- ActivityDescription (« HipsgenM »,
-         [
            Prov:Label = « HiPS Generation MEAN »,
            Prov:Type = « HiPSgen »,
            voprov:Subtype= « HiPS_MEAN »
            voprov:docuLink = « http://cds.u-strasbg.fr/HiPSGEN-Documentation »
- WasAttributedTo(« EHG1 », « Buga », voprov:role=voprov:operator)
- Agent (« buga », [name = « Mihaela Buga », contact= « buga@astro.unistra.fr »])
- WasAttributedTo(« EHG1 », « ivo://CDS », prov:role=« voprov:creator »)
-

Describing an HiPS provenance in IVOA/Prov-DM formalism

- Used (« EHG1 », « ivo://obspm/gepi/ESO-MAMA », Prov:Role = « hips:OriginalData »)
- -> Entity(« ivo://obspm/gepi/ESO-MAMA »,
-          [
  Prov:Label:=« ESO-R Survey Mama GEPI »,
  Prov:Type = collection,
  Prov:Description = « This is the ESO Schmidt survey collection digitized by MAMA at CAI curated byGEPI »,
  Prov:Description= « http://ivoa.net/semantics/datacollection/images »,
  Prov:url = « http//obspm.fr/gepi/ESOR-MAMA »
      ])
- EntityDescription (« ivo://ivoa.net/std/prov/datacollection/images »,
-       [
   Prov:Label= « Image collection »
   hips:Dataproduct_type = image
   Hips:Level= 2
   ])
- WasAttributedTo(« ivo://obspm/gepi/ESO-MAMA »,   « Guibert »,prov:Role =« author »)
- Agent(« Guibert », voprov:name = « Jean Guibert », voprov:contact= « guibert@obspm.fr »)
- WasAttributedTo((« ivo://obspm/gepi/ESO-MAMA », « ivoident://obspm/gepî », prov:Role=« curator »)
- Agent (« ivo://obspm/gepi, [Voprov:name=« GEPI », contact: « lesidaner@obspm.fr »])
  WasAttributedTo:(« ivo://obspm/gepi/ESO-MAMA »,   « ivo://obspm/cai »,prov:Role= « Creator »)

Agent:(« ivoident: ivo://obspm/cai »,[voprov:name= « centre analyse des images », voprov:contact=« lesidaner@obspm.fr »)

☐ Questions/ Work to do

- A lot of questions:
  - Do we have all what we need: types ? Subtypes ? Roles ? Additional attributes ? Software Versions ?
  - Check validity/interest with HiPS developpers
    - They Want to know if they can pick up some prov concepts for HiPS properties (see next slides)
  - Write other serialization (PROV-JSON, PROV-XML or PROV-VOTABLE) and Diagram
  - Use case Integrated In the Working draft ?

New proposed hips provenance properties

- Hips_prov_creator :
- Hips_prov_creator_contact
- Hips_prov_operator :
- Hips_prov_activitysubtype
- Hips_prov_activitystop
- Hips_prov_activitystart = hips_creation_date
- Hips_prov_level
- Hips_prov_parameterset
- Hips_prov_mask

Describing an HiPS provenance in IVOA/Prov-DM formalism

- Used
  - Prov:Role : parameterSet
- -> Entity
  - Prov:ID = ivo://CDS/P/MAMA/ESO#configFile
  - Prov:Label : « CDS ESO HIPS generation config file »
  - Prov: Description : « This is the configuration file of the ID ivo://cds/P/mama/ESO.. Generation »
  - Prov:Description : ID of entity description for parameter files ?
- Used
  - Prov:Role : ancillary_data
- -> Entity
  - Prov:ID ivo://cds/P/mama/ESO#eso.fov
  - Prov:Label :   »ESO plate FOV »
  - Prov:Description : « file containing the exact contours of the FOV of ESO Plates for HiPS generation»
  - Prov:Description : ID of entity description for fov files.