Proposal for an evolution of the SIA protocol.

(M.Louys, P.Fernique, T.Boch, F.Bonnarel + …)

The purpose of this note is to propose extensions of the SIAP protocol (http://www.ivoa.net/internal/IVOA/IvoaDAL/ACF8DE.pdf)
in order to allow some kind of structured description of metadata for images or observations provided by an image server or an observation archive. One main requirement for that is to allow VO tools to use this extension protocol as the Data Access Layer for metadata trees, intended for hierarchical display of observation metadata like in the AVO prototype. In the same time we are looking for backward compatibility with SIA.

1. Introduction

The background of this proposal relies on lessons learnt from one and a half year of experience with SIA (or SSA) and the IDHA format within the context of the AVO prototype development. Developed in parallel these two metadata descriptions have common features and advantages as well as very big differences.

Here are the common points:

- They both solve the "availability" problem of images in a given cone search for users and client software developers.
- They aim to provide a standardized description of the images helping the user to "make his/her choice" among available images.
- They both express in VOTable

The differences lay in the way the metadata are structured in the VOTable.

- The main argument for SIA is simplicity: each image is basically described by a tuple in the unique table. Easy to generate it allows data providers and archive managers to develop interfaces based on this protocol quickly.
- The main argument for IDHA Format (Image Data Hierarchical Access, http://www.euro-vo.org/twiki/bin/view/Avo/MetadataTree) is to allow for hierarchical display of the observation metadata as a "metadata tree" in the AVO. It is based on the IDHA datamodel (Images distribuées hétérogènes pour l'Astronomie - http://alinda.u-strasbg.fr/IDHA/lastmodel). The possibility of grouping observations according to some common criteria such as bandpass or filter name, epoch of observation, Observing Program has received satisfactory return from the users, in particular during demonstration and tutorials in front of the AVO science Working team. Extension of this tree structure under the observation level describing various packaging modes of the same observation (datacube slices or preview) has been presented at the January 2004 AVO demonstration and welcomed.

The main drawback of IDHA format may be complexity for both generation and parsing. Another criticism has been made on the lack of flexibility of this tree organisation in the VOTable. That's the reason why we propose a new way -somewhat halfway between IDHA format and SIA- of describing images and Observations. In this description we are actually mapping a datamodel in VOTable. Basic principles we use to do that are exposed in another note (in preparation). Currently we are using IDHA datamodel classes, but we can use the IVOA Observation model in the future (reference). The Observation class is associated to the primary (SIAP like) table. Linked additional datamodel classes are associated to secondary tables. In practice, we propose to use a basic record to describe an observation using the standard SIA structure. A couple of additional fields will also be defined. Some of

these fields can then be used as internal keys towards other tables, where they are repeated as VOTable references. Records sharing a common value for these common fields are then associated.

Let's start with an example: suppose we have a SIAP server providing 2MASS data (eg the CDS one). That's the kind of output we have:

```
   <?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE VOTABLE SYSTEM "http://us-vo.org/xml/VOTable.dtd">
<VOTABLE ID="v1.0">
  <DESCRIPTION> SIAP output for Aladin server </DESCRIPTION>
  <RESOURCE type="results">
    <INFO name="QUERY_STATUS" value="OK"/>
      <TABLE>
        <FIELD ID="Observation_Name" ucd="VOX:Image_Title" datatype="char" arraysize="*" />
        <FIELD ID="CentralPoint_RA" ucd="POS_EQ_RA_MAIN" datatype="double"  />
        <FIELD ID="CentralPoint_DEC" ucd="POS_EQ_DEC_MAIN" datatype="double"  />
        <FIELD ID="Naxes" ucd="VOX:Image_Naxes" datatype="int"  />
        <FIELD ID="Naxis" ucd="VOX:Image_Naxis" datatype="int" arraysize="*" />
        <FIELD ID="AngularPixelSize" ucd="VOX:Image_Scale" datatype="double" arraysize="*" unit="deg" />
        <FIELD ID="OriginalCoding" ucd="VOX:Image_Format" datatype="char" arraysize="*" />
          <FIELD ID="Filter_Name" ucd="VOX:BandPass_ID" datatype="char" arraysize="*"
/>
        <FIELD ID="Effective_wavelength"  ucd="VOX:BandPass_RefValue" datatype="double" unit="um" />
        <FIELD ID="Minimal_wavelength"  ucd="VOX:BandPass_LoLimit" datatype="double" unit="um" />
        <FIELD ID="Maxima_wavelength"  ucd="VOX:BandPass_HiLimit" datatype="double" unit="um" />
        <FIELD ID="Location"  ucd="VOX:Image_AccessReference" datatype="char" arraysize="*" />
        <FIELD ID="PlateNumber"   datatype="char" arraysize="*" />
        <FIELD ID="ObservingProgram"   datatype="char" arraysize="*" />
        <DATA>
          <TABLEDATA>
    <TR>
            <TD>2MASS_K_981122S_KI0760103</TD>
            <TD>53.054080 </TD>
            <TD>-27.707217 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
            <TD>K</TD>
            <TD>2.160000</TD>
            <TD>2.020000</TD>
            <TD>2.300000</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP.cgi?
out=image&position=53.054080+-27.707217&survey=2MASS&color=K&mode=view]]></TD>
            <TD>981122S_KI0760103</TD>
            <TD>2MASS</TD>
           </TR>
           <TR>
            <TD>2MASS_H_981122S_HI0760103</TD>
            <TD>53.054080 </TD>
            <TD>-27.707217 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
            <TD>H</TD>
            <TD>1.660000</TD>
            <TD>1.510000</TD>
            <TD>1.780000</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-TTP.cgi?
out=image&position=53.054080+-27.707217&survey=2MASS&color=H&mode=view]]></TD>
            <TD>981122S_HI0760103</TD>
            <TD>2MASS</TD>
           </TR>
           <TR>
            <TD>2MASS_J_981122S_JI0760103</TD>
            <TD>53.054080 </TD>
            <TD>-27.707217 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
            <TD>J</TD>
            <TD>1.240000</TD>
            <TD>1.110000</TD>
```

```
            <TD>1.360000</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/ nph-HTTP.cgi?
out=image&position=53.054080+-27.707217&survey=2MASS&color=J&mode=view]]></TD>
            <TD>981122S_JI0760103</TD>
            <TD>2MASS</TD>
           </TR>
           <TR>
            <TD>2MASS_K_981122S_KI0770174</TD>
            <TD>53.185894 </TD>
            <TD>-27.762975 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
            <TD>K</TD>
            <TD>2.160000</TD>
            <TD>2.020000</TD>
            <TD>2.300000</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP.cgi?
out=image&position=53.185894+-27.762975&survey=2MASS&color=K&mode=view]]></TD>
            <TD>981122S_KI0770174</TD>
            <TD>2MASS</TD>
           </TR>
           <TR>
            <TD>2MASS_H_981122S_HI0770174</TD>
            <TD>53.185894 </TD>
            <TD>-27.762975 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
            <TD>H</TD>
            <TD>1.660000</TD>
            <TD>1.510000</TD>
            <TD>1.780000</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP.cgi?
out=image&position=53.185894+-27.762975&survey=2MASS&color=H&mode=view]]></TD>
            <TD>981122S_HI0770174</TD>
            <TD>2MASS</TD>
           </TR>
           <TR>
            <TD>2MASS_J_981122S_JI0770174</TD>
            <TD>53.185894 </TD>
            <TD>-27.762975 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
            <TD>J</TD>
            <TD>1.240000</TD>
            <TD>1.110000</TD>
            <TD>1.360000</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP.cgi?
out=image&position=53.185894+-27.762975&survey=2MASS&color=J&mode=view]]></TD>
            <TD>981122S_JI0770174</TD>
            <TD>2MASS</TD>
           </TR>
           </TABLEDATA></DATA>
         </TABLE>
  </RESOURCE>
</VOTABLE>
```

In this format we have no details about the ObservingProgram and some Observation configuration elements: Filter, Telescope, camera, etc …. Some of them (Telescope, camera) are even absent. If we choose to add all these details as new fields in the record, we will reach a high level of redundancy because they can be common to several (in some cases many) observations.

2. Factorization of common features

What we do first is to add a Telescope and Camera field (and remove wavelength details because they will be factorized) beside ObservingProgram and VOX:BandpassID (which is close to Filter). At this point we do not add details, just Names.

```
<!DOCTYPE VOTABLE SYSTEM "http://us-vo.org/xml/VOTable.dtd">
<VOTABLE ID="v1.0">
  <DESCRIPTION> SIAP output for Aladin server </DESCRIPTION>
  <RESOURCE type="results">
    <INFO name="QUERY_STATUS" value="OK"/>
      <TABLE>
        <FIELD ID="Observation_Name" ucd="VOX:Image_Title" datatype="char" arraysize="*" />
        <FIELD ID="CentralPoint_RA" ucd="POS_EQ_RA_MAIN" datatype="double"  />
        <FIELD ID="CentralPoint_DEC" ucd="POS_EQ_DEC_MAIN" datatype="double"  />
        <FIELD ID="Naxes" ucd="VOX:Image_Naxes" datatype="int"  />
        <FIELD ID="Naxis" ucd="VOX:Image_Naxis" datatype="int" arraysize="*" />
        <FIELD ID="AngularPixelSize" ucd="VOX:Image_Scale" datatype="double" arraysize="*" unit="deg" />
        <FIELD ID="OriginalCoding" ucd="VOX:Image_Format" datatype="char" arraysize="*" />
            <FIELD ID="FilterName" ucd="VOX:BandPass_ID" datatype="char" arraysize="*"
/>
        <FIELD ID="Location"  ucd="VOX:Image_AccessReference" datatype="char" arraysize="*" />
        <FIELD ID="PlateNumber"   datatype="char" arraysize="*" />
            <FIELD ID="ObservingProgramName"   datatype="char" arraysize="*" />
            <FIELD ID=”TelescopeName” datatype=”char” arraysize=”*” />
            <FIELD ID=”CameraName” datatype=”char” arraysize=”*” />
     <DATA><TABLEDATA>
    <TR>
            <TD>2MASS_K_981122S_KI0760103</TD>
            <TD>53.054080 </TD>
            <TD>-27.707217 </TD>
            <TD>2</TD>
            <TD></TD>
            <TD>20.000000 20.000000</TD>
            <TD>image/fits</TD>
                <TD>IR K</TD>
            <TD><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP.cgi?
out=image&position=53.054080+-27.707217&survey=2MASS&color=K&mode=view]]></TD>
            <TD>981122S_KI0760103</TD>
            <TD>2MASS</TD>
                <TD>CTIO (Chile)</TD>
                <TD>NICMOS3(HgCdTe)_J</TD>
          </TR>
              <TR>…….
      </TABLEDATA></DATA>
      </TABLE>
</RESOURCE>
```

Beside these new fields which contain only character string values (= “names”) we will now define new tables for each of these, in order to give details. Let’s add a new “ModelExpansion” resource after the results one, containing as many new tables than we need classes, in the following way:

```
<RESOURCE type="ModelExpansion">
       <TABLE ID=”Filter”>
        <DESCRIPTION>
          Description of the filter characteristics
        </DESCRIPTION>
          <FIELD ref="FilterName"/>
        <FIELD ID="Minimal_wavelength" name="Minimal_wavelength" datatype="float" precision="F7" width="11"
ucd="SPECT_WAVELENGTH" unit = "um">
            <DESCRIPTION Minimal wavelength of the filter bandpath />
         </FIELD>
        <FIELD ID="Maximal_wavelength" name="Maximal_wavelength" datatype="float" precision="F7" width="11"
ucd="SPECT_WAVELENGTH" unit = "um">
            <DESCRIPTION Maximal wavelength of the filter bandpath />
        </FIELD>
        <FIELD ID="Effective_wavelength" name="Effective_wavelength" datatype="float" precision="F7" width="11"
ucd="SPECT_WAVELENGTH " unit = "um">
            <DESCRIPTION Effective wavelength of the filter bandpath />
        </FIELD>
        <FIELD ID="Identifier" name="Identifier" datatype="char" ucd="ID_FILTER">
            <DESCRIPTION Reference ID of the filter, according to the Instruments settings. />
```

```
        </FIELD>
        <DATA><TABLEDATA>
            <TR>><TD>IR
K</TD><TD>2.020000</TD><TD>2.300000</TD><TD>2.160000</TD><TD>K</TD></T
R>
          <TR>><TD>IR J</TD
<TD>1.110000</TD><TD>1.360000</TD><TD>1.240000</TD><TD>J</TD ></TR>
          <TR>><TD>IR H</TD
<TD>1.510000</TD><TD>1.780000</TD><TD>1.660000</TD><TD>H</TD ></TR>
        </TABLEDATA></DATA>
      </TABLE>
       <TABLE ID="ObservingProgram">
   <DESCRIPTION>
     This is a few resource information for  the    ObservingProgram
   </DESCRIPTION>
    <FIELD ref="ObservingProgramName"/>
   <FIELD ID="Organisation" name="Organisation" datatype="char" ucd="CURATOR">
     <DESCRIPTION Name of Organisation(s)  performing Observing Program />
   </FIELD>
   <FIELD ID="Begin" name="beginning_date" datatype="char">
     <DESCRIPTION Begin date of the survey />
   </FIELD>
   <FIELD ID="End" name="end date" datatype="char">
     <DESCRIPTION End date of Observing program />
   </FIELD>
   <FIELD ID="SpectralDomain" name="SpectralCoverageName" datatype="char" ucd="SPECT_RANGE">
     <DESCRIPTION General spectral domain (Optical X-ray ...) />
   </FIELD>
     <DATA><TABLEDATA>
            <TR><TD>2MASS</TD><TD>UNIVERSITY OF MASSACHUSETS / IPAC
(CALTECH /
JPL)</TD><TD>07/06/1997</TD><TD>15/02/2001</TD><TD>Optical</TD></TR>
     </TABLEDATA></DATA>
    <TABLE ID=”Telescope”>
         <FIELD ref=” TelescopeName” />
        <FIELD ID=”Longitude”  datatype=”char” ucd=”POS_EARTH_LON “ arraysize=”*”>
                    <DESCRIPTION    Terrestrial longitude />
        </FIELD>
       <FIELD ID=”Latitude” datatype=”char” ucd=”POS_EARTH_LAT” arraysize=”*”>
                    <DESCRIPTION   Terrestrial latitude />
       </FIELD>
       <FIELD ID=”Height” datatype=”float” ucd=”OBSTY_ALTITUDE” arraysize=”1” unit=”m”>
                   <DESCRIPTION Altitude />
        </FIELD>
        <FIELD ID=”Description” datatype=”char” ucd=”DATA_LINK” arraysize=”*”>
                    <DESCRIPTION  Bibliographic information about the telescope />
         </FIELD>
         <FIELD ID=”Organisation” datatype=”char” ucd=”CURATOR” arraysize=”*”>
                    <DESCRIPTION  Institute responsible for this telescope. />
            </FIELD>
           <DATA><TABLEDATA>
          <TR><TD> CTIO (Chile)</TD> +70.815</TD><TD> -30.165</TD><TD></TD><TD></TD><TD></TD><TD></TR>
          </TABLEDATA></DATA>
       </TABLE>
</RESOURCE>
```

The advantage of such a structure is to deliver metadata in a prospective manner and to be backward compatible. SIA Vers 1 services will interpret only the first table of metadata. SIAP evolution services will interpret all the tables with factorized information and reference pointer access in the XML document. They will be able to generate metadata trees or other structured descriptions in a flexible way.

3. Extended details on a single observation

There are other piece of metadata which can be given outside the main “result” table , but which are directly linked to a single observation. We propose to group the additional tables used to describe these metadata in a new “ObservationDetails” RESOURCE. A typical example of such additional table will be Astrometry or WCS. We can choose to have this information in the main “results” table or to shift that to a dedicated table further in the document. This is useful to directly identify WCS field in one single structure. This can also allow the management of several versions of WCS for the same Observation. This concept can be extended to other kind of calibrations (e.g. photometry).

Another critical field is the acref one: actually the case where one single observation can be retrieved by one single fixed URL is not general enough. (see examples in an informal note that one of us (FB) posted on the DAL and DM mailing lists -http: //www.ivoa.net/forum/dal/0075.htm). If a server produces several different kind of products for a single observation: preview, cutout, access to a full retrieval page, we propose to describe all possible retrieval facilities in as many Packaging records as necessary. Each of them may (or may not) require a specific set of parameters in order to generate URL templates. A possible solution for URL templating has been proposed in FB informal note and is used in the following example. In this example the single URL directly given in the acref field is just a default URL. (Note: Packaging is here encompassing the info previously described in a Table/Class misnamed StorageMapping in the IDHA model and IDHA format).

Example (GOODS WFI survey: Preview Packaging and Cutout Packaging – NB: In the Template indexing case, we have used the convention on the position parameter described in Bonnarel informal note -http://www.ivoa.net/forum/dal/0075.htm -)

```
<!DOCTYPE VOTABLE SYSTEM "http://us-vo.org/xml/VOTable.dtd">
<VOTABLE ID="v1.0">
  <DESCRIPTION> SIAP output for Aladin server </DESCRIPTION>
  <RESOURCE type="results">
    <INFO name="QUERY_STATUS" value="OK"/>
      <TABLE>
        <FIELD ID="Observation_Name" ucd="VOX:Image_Title" datatype="char" arraysize="*" />
        <FIELD ID="CentralPoint_RA" ucd="POS_EQ_RA_MAIN" datatype="double" />
        <FIELD ID="CentralPoint_DEC" ucd="POS_EQ_DEC_MAIN" datatype="double" />
        <FIELD ID="Naxes" ucd="VOX:Image_Naxes" datatype="int" />
        <FIELD ID="Naxis" ucd="VOX:Image_Naxis" datatype="int" arraysize="*" />
        <FIELD ID="AngularPixelSize" ucd="VOX:Image_Scale" datatype="double" arraysize="*" unit="deg" />
        <FIELD ID="OriginalCoding" ucd="VOX:Image_Format" datatype="char" arraysize="*" />
          <FIELD ID="FilterName" ucd="VOX:BandPass_ID" datatype="char"
arraysize="*" />
          <FIELD ID="Location" ucd="VOX:Image_AccessReference" datatype="char"
arraysize="*" ref=”Packaging” />
        <FIELD ID="PlateNumber" datatype="char" arraysize="*" />
          <FIELD ID="ObservingProgramName" datatype="char" arraysize="*" />
        <DATA>
         <TABLEDATA>
           <TR>
             <TD>GOODS-WFI_ICLWP_DEEP2C-FI</TD>
             <TD>53.119485 </TD>
             <TD>-27.803630 </TD>
             <TD>2</TD>
             <TD></TD>
             <TD>0.000066 0.000066</TD>
             <TD>image/fits</TD>
             <TD>ICLWP</TD>
                <TD>><![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP.cgi?
```

```
    out=image&position=53.054080+-27.707217&survey= GOODS-WFI &color=ICLWP
    field=DEEP2C-FI-PREVIEW&mode=view]]></TD>
                <TD>DEEP2C-FI</TD>
                <TD>GOODS-WFI</TD>
               </TR>
          </TABLEDATA></DATA>
    </TABLE>
</RESOURCE>
<RESOURCE type="ModelExpansion">
      <TABLE ID=”Filter”>
            <FIELD ID=….
    ……..
    ……..
</RESOURCE>
<RESOURCE type="ObservationDetails”>
    <TABLE ID=”Packaging”>
          <FIELD ref="ObservationName" >
         </FIELD>
          <FIELD ID="Organisation"  datatype="char" >
         <DESCRIPTION>
             Status of cutout availability
          </DESCRIPTION>
         </FIELD>
         <FIELD ID="desc" name="OrganisationDescription" datatype="char">
            <DESCRIPTION>
             Description of the organisation
          </DESCRIPTION>
          <FIELD ID=" Maximum size"  datatype="int"  >
          <DESCRIPTION>
                Maximum size of a cutout / total size of  a “standalone” image
          </DESCRIPTION>
            <FIELD ID="Indexing" name="Indexing"  datatype="char">
              <DESCRIPTION>
                   Image or Subimage Indexation mode
          </DESCRIPTION>
             </FIELD>
             <FIELD ID=”LinktoPixels" datatype="char"  ucd="DATA_LINK">
        </FIELD>
       <DATA>
        <TABLEDATA>
         <TR><TD> GOODS-WFI_ICLWP_DEEP2C-FI </TD><TD>PREVIEW</TD><TD>Preview of the
    Observation</TD><TD>621</TD><TD>URL</TD><TD> <![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP3.cgi?
    out=image&position=$POS&survey=GOODS-WFI&color=ICLWP&field=DEEP2C-FI-PREVIEW&mode=view]]></TD></TR>
    <TR><TD> GOODS-WFI_ICLWP_DEEP2C-FI </TD><TD>CUTOUT </TD><TD>Cutouts in  the
    Observation</TD><TD>20481</TD><TD>TEMPLATE</TD><TD> <![CDATA[http://aladin.u-strasbg.fr/cgi-bin/nph-HTTP3.cgi?
    out=image&position=$POS&survey=GOODS-WFI&color=ICLWP&field=DEEP2C-FI-PREVIEW&mode=view]]></TD></TR>
    ……..
```

As a matter of conclusion, lets’say that we have assessed that implementation of this proposal is possible for the existing resources available in the Aladin image server and the complex resources used in the AVO Science demos. It would allow the possible display of hierarchical data trees organised following different parameters at user request.