

# SIA2/SODA upgrades and Extending SIA2 into a generic DataSetSAP

---

F.Bonnarel (CDS)

acknowledgement : DAL working group



# SIAP2 feedback

## Where it can be discussed ?

- SIAP2.0 adopted 2015/12/23 (6.5 years!!!)
- Discussion at interops and on github since 2018, May 2020, May 2021
- SIA on GitHub : <https://github.com/ivoa-std/SIA/>
- Discussion possible on the IVOA wiki too

[https://wiki.ivoa.net/twiki/bin/view/IVOA/SIAP-2\\_0-Next](https://wiki.ivoa.net/twiki/bin/view/IVOA/SIAP-2_0-Next)

- Time to go to a proposal
- Do not merge PR yet !



**IVOA Simple Image Access**  
**Version 2.0**

**IVOA Recommendation 2015-12-23**

Working group  
Data Access Layer Working Group  
This version  
<http://www.ivoa.net/documents/SIA/20151223>  
Latest version  
<http://www.ivoa.net/documents/SIA>  
Previous versions  
Author(s)  
Patrick Dowler, Douglas Tody, François Bonnard  
Editor(s)  
Patrick Dowler, François Bonnard

# SIA github issues

**10 Open** ✓ 0 Closed

Author ▾ Label ▾ Projects ▾ Milestones ▾ Assignee ▾ Sort ▾

**extension of SIA-style protocol usage outside the image/cube "camp"**

#10 opened on 13 May by Bonnarel

**It is not possible to query SIA services by MOC**

#9 opened on 11 May by Bonnarel

**SIA2 cannot discover rebinned data like SIA1 was able to do**

#8 opened on 8 May by Bonnarel

**Possible confusion between FORMAT and RESPONSEFORMAT parameter**

#7 opened on 8 May by Bonnarel

**1 shot discovery (and then access) to cutouts was possible in SIA1 but no more in SIA2**

#6 opened on 8 May by Bonnarel

**POS=RANGE examples inconsistent with spec**

#5 opened on 7 May by pdowler

**input PARAMETERS values are case sensitive**

#4 opened on 7 May by Bonnarel

**NO Wild-carding of the input PARAMETERS values exists**

#3 opened on 7 May by Bonnarel

**No input PARAMETER exists to select the RELEASE DATE**

#2 opened on 7 May by Bonnarel

**input PARAMETERS with limited list of values : better description**

#1 opened on 7 May by Bonnarel

# I ) SIA2 errata

- POS=RANGE examples inconsistent with spec (Pat Dowler)

PR created to fix that

- Possible confusion between FORMAT and RESPONSEFORMAT parameter to be clarified (Alberto Micol)

PR created to clarify this

- Typo in 1.3

«COLLECTION and FACILITY currently provide query parameters provide selection on service defined set of strings. »





# | |)SIAP2 parameters: availability of list of possible values

Several SIAP2.0 parameters have a limited list of possible values

- Some have lists limited by protocol (and obscure)
  - POL (Stokes, LINEAR, etc..)
  - DPTYPE (image, cube, visibility, timeseries ;..)
  - CALIB : levels
  - FORMAT : fits, jpeg , png, etc..
- Some have free string values
  - COLLECTION (HST, WISE, etc...), FACILITY (VLT, Keck, Chandra), INSTRUMENT (ACS, MEGACAM, etc.)
- PARAMETERS less useful if we have no prior idea of their possible values. **This information is often missing in services**

**Proposal to change MAY in SHOULD or MUST in the self description subsection**

**PR created**

**Question : how do we retrieve that ? Query without parameter ? MAXREC=0 ? Other ?**



# Lack of release\_date parameter (PyVO / CADC)

- ObsCore has an optional obs\_release\_date parameter.
- SIAP2 doesn't provide corresponding QUERY PARAMETER
- PR created, RELEASEDATE parameter to check this
- If the service doesn't have the obs\_release\_date parameter : empty response



## II.) Lack of flexibility on parameter value « style » (PyVO /CADC)

STRING QUERY PARAMETER don't allow wild carding or  
incompletion.

→ Nothing done yet

STRING QUERY PARAMETER are case sensitive only

→Nothing done yet



## II ) Adding a MOC parameter

- Select data set in SMOC, TMOC or STMOC areas
- PR proposes MOC = « ascii MOC syntax »





## II ) SIAP 2 Discovery and access :

SIAP1 versus SIAP2 : virtual data versus axis completion

- SIAP1 had « cutout » and « mosaic » modes beside « archive » mode
  - 1 shot before access but only spatial
- We now have :
  - SIAP2.0 or ObsTAP
  - + SODA : for cutouts only (all axes)
  - +DataLink (Service descriptor and/or {links} table)
  - → 2 shots before access (instead of 1)



## II )SIAP 2 Discovery and access :

SIAP1 versus SIAP2 : virtual data versus axis completion

- perfectly possible to provide functionality by replacing the full retrieval or datalink url in « access\_url » by a SODA url.
- SODA URL parameters are similar to SIA ones.
  - When SIA Parameters values constrain the discovery, SODA parameters force the cutout dimensions.
  - See Hips2FITS SIA2 service in Aladin Desktop
- PR proposes a RETRIEVEMODE parameter (FULL/CUTOUT)
- Default is FULL or DataLink



# Other features planned in the 2.0 spec introduction

- Full metadata (cube-DM-oriented ?) endpoint
  - depends on cube-DM achievement ?
  - only provide FITS-headers ? Provenance ?
  - proposal/prototype by CADC : integrated in SODA (see November interop presentation)



# III ) Extending allowed dataproduct\_type

## (1) relaxing

- Currently limited to « image » and « cube »
- Why not relaxing this ? (could be useful for generic queries
  - No real issue for timeseries, spectra
  - If we include spectra it could serve as SSA2 (ObsCore compatible)
  - Visibilities and event lists ?
  - Measurement/catalog: It's for discovering catalogs as a whole (see ESO ? , ASKAP ? CADC?)



## III )Extending allowed dataproduct\_type

- It's a different protocol
- Propose to call it « DatasetSAP » (DsSAP)
- PR created





# III ) Extending allowed dataproduct\_type

@@ -472,7 +453,9 @@ \subsubsection{INSTRUMENT}

472 The INSTRUMENT parameter is a string-valued parameter that specifies the name  
of the instrument with which the data was acquired. The value is compared with  
the instrument\\_name from the ObsCore data model.

473

474 \subsubsection{DPTYPE}

475 - The DPTYPE parameter is a string-valued parameter that specifies the type of  
data. The value is compared with the dataproduct\\_type from the ObsCore data  
model. For the SIA \{query\} resource, the only values that should be returned  
for dataproduct\\_type are \textit{image} and \textit{cube}, so this parameter  
can be only really be used to select one of these.

476

---

453 The INSTRUMENT parameter is a string-valued parameter that specifies the name  
of the instrument with which the data was acquired. The value is compared with  
the instrument\\_name from the ObsCore data model.

454

455 \subsubsection{DPTYPE}

456 + The DPTYPE parameter is a string-valued parameter that specifies the type of  
data. The value is compared with the dataproduct\\_type from the ObsCore data  
model. In contrast to SIA2.0 which allows only dataproduct\\_types  
\textit{image} and \textit{cube}, all the other dataproduct\\_types are  
allowed, in such a way that we can constrain the service to retrieve  
visibility data, timeseries and event lists. Using DPTYPE with the "spectrum"  
value can be considered as an upgrade of the SSA protocol.

457 +

458 + ObsCore extension have been defined for visibility data and TimeSeries. It is  
possible to query services using these extensions by optional query parameters  
defined in Appendix B.

459

---



# III ) Extending allowed dataproduct\_type

```
42 \section{Introduction}
```

```
43 - The Simple Image Access (SIA) protocol defines several capabilities to support  
discovery and access to astronomical image datasets of any dimension. Typical  
image datasets include 2-D spatial images, spectral data cubes, and cube and  
hypercube data of higher dimensions as well as derived image data products.  
The underlying ObsCore data model is a simplified view on the typical image  
datasets derived from observational data, which have some combination of  
spatial, spectral (including velocity and redshift), time, and polarization  
axes.
```

```
44 - For complete access to datacubes, the SIA-2.0 specification makes use of  
features defined in DataLink \citep{std:DataLink}. It also makes use of  
AccessData services, as well as custom data services.
```

```
45  
46  
47
```

```
@@ -52,45 +45,26 @@ \section{Introduction}
```

```
52 \label{fig:architecture}
```

```
53 \end{figure}
```

```
54
```

```
55 - SIA defines data discovery and metadata capabilities that work with other DAL  
services to enable image and data cube access. The basic interface for the  
capabilities defined in this specification are described in DALI  
\citep{std:DALI}. DataLink can be used with SIA for finding access URL(s) for  
files, related resources, and data services such as AccessData (in  
development). SIA services also support VOSI-availability and VOSI-  
capabilities \citep{std:VOSI} resources.
```

```
56
```

```
35 \section{Introduction}
```

```
36 + The Dataset Simple Access (DsSAP) protocol defines several capabilities to  
support discovery and access to astronomical datasets of any type and  
dimension. Typical datasets include spectra, timeseries, 2-D spatial images,  
spectral data cubes, and cube and hypercube data of higher dimensions as well  
as derived image data products an event list or visibility data. The  
underlying ObsCore data model is a simplified view on the typical image  
datasets derived from observational data, which have some combination of  
spatial, spectral (including velocity and redshift), time, and polarization  
axes.
```

```
37 + For complete access to datacubes, the SIA-2.0 specification makes use of  
features defined in DataLink \citep{std:DataLink}. It also makes use of  
AccessData services such as SODA \citep{2017ivoa.spec.0517B}, as well as  
custom data services.
```

```
38  
39  
40
```

```
45 \label{fig:architecture}
```

```
46 \end{figure}
```

```
47
```

```
48 + DsSAP defines data discovery and metadata capabilities that work with other  
DAL services to enable image, data cube and other types of dataset access. The  
basic interface for the capabilities defined in this specification are  
described in DALI \citep{std:DALI}. DataLink can be used with DsSAP for  
finding access URL(s) for files, related resources, and data services such as  
SODA. DsSAP services also support VOSI-availability and VOSI-capabilities  
\citep{std:VOSI} resources.
```

```
49
```



# III ) Extending allowed dataproduct\_type

56

57 + Virtual Observatory access to astronomical images has been available via the  
SIA-1.0 protocol for over a decade.

58 + Many such services have been implemented since 2002, and SIA-1.0  
`\citep{std:SIAP}` was formally standardized as an IVOA Recommendation in 2009.

59 + SSA `\citep{std:SSAP}` played a similar role for spectra and also used specific  
metadata in the query response.

60 + SIA-2.0 `\citep{std:SIAv2}` was multi-dimensional and fully integrated with the  
modern VO architecture and related

61 + standards, but restricted to images and data cubes.

62 + DsSAP is an extension of SIA2 to other dataproduct types. It can be seen as a  
server side parameter based proxy to an ObsTAP service.

63 `\subsection{Motivating Use` `\subsection{Changes from SIA-1.0 to SIA-2.0}`

Virtual Observatory access to astronomical images has been available via the SIA-1.0 protocol for over a decade. Many such services have been implemented since 2002, and SIA-1.0 `\citep{std:SIAP}` was formally standardized as an IVOA Recommendation in 2009. The legacy SIA standard however pre-dates much of the VO technology developed since 2002, and is limited to two-dimensional images. SIA-2.0 is multi-dimensional and fully integrated with the modern VO architecture and related standards.

SIA-2.0 differs from legacy SIA-1.0 in the following aspects:

`\begin{itemize}`

`\item` The capabilities for dynamic access to image datasets are expanded in scope, but are separated from data discovery and download of whole image datasets. A separate **"AccessData"** specification currently under development will define the more advanced dynamic data access functionality. Automated virtual data generation and discovery (as in SIA-1.0) is not currently



# SODA

- Errata
- MOC parameter
- Rebin/reproject
- Choose `dataproduct_type` of retrieved data
- Metadata extraction (fits header, cube ,  
provenance?)

