Escape - an EU Horizon 2020 project - 824064

An introduction to the CDS services and tools

Mark Allen and the CDS team
Centre de Données astronomiques de Strasbourg, France

last update 01.02.2021

Mark Allen and the CDS team
Initial development of the tutorial

Jenny G. Sorce
updated with a new template for ASTERICS with the new CDS portal, new plots

Katharina A. Lutz
updated to Aladin v10

Ada Nebot
updated for VO school in Nice, November 2019

Katharina A. Lutz
updated to Aladin v11 and ESCAPE template

Contents

1 Introduction

The CDS hosts three major services, accessible through the CDS portal:

- **SIMBAD**: The astronomical database SIMBAD contains more than 12 million objects. For each object it provides basic measurements (type, coordinates, proper motion, radial velocity, spectral type, distance, magnitude), cross-correlations and bibliography.

- **VizieR**: The VizieR catalogue service provides access to more than 20,000 catalogues, being the most complete library of published astronomical data tables available online.

- **Aladin**: The interactive sky atlas Aladin allows to visualize astronomical images and to superimpose entries from different catalogues and databases. It allows to visualize data from SIMBAD, VizieR, and from archives and databases distributed around the world. Furthermore, you can open your own tables or images. There are two versions of Aladin: Aladin desktop and Aladin Lite, which runs in the browser.

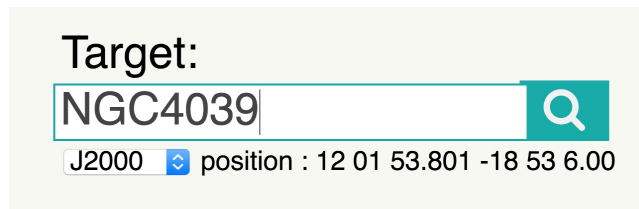In addition, the CDS has been offering a cross-match service (X-match ![XMATCH]) since November 2011.

2 Goal of this tutorial

This tutorial shows how to use the CDS tools to gather information on specific astronomical objects. We will:

- Search for information on NGC4039 in the CDS portal

- Search for data concerning NGC4039 in Aladin

- Compare the coverage of Sky Surveys and select interacting galaxies that have SDSS and GALEX data

3 Search for information on NGC4039 in the CDS Portal

Open the CDS Portal and make a query for 'NGC4039'. The result provides an overview of the information and data available for this object in the 3 CDS services: SIMBAD, Aladin and VizieR:

Figure 1: Query for NGC4039 in the CDS Portal

3.1 SIMBAD - Identifiers, Basic Measurements and links to the Bibliography
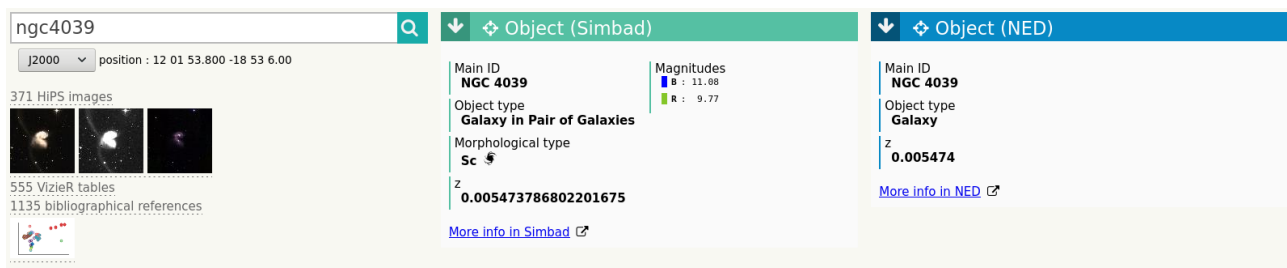

Figure 2: Result of the query for NGC4039 in the CDS Portal

- Click on **More info in Simbad** to see the full SIMBAD information on this object in a new tab.
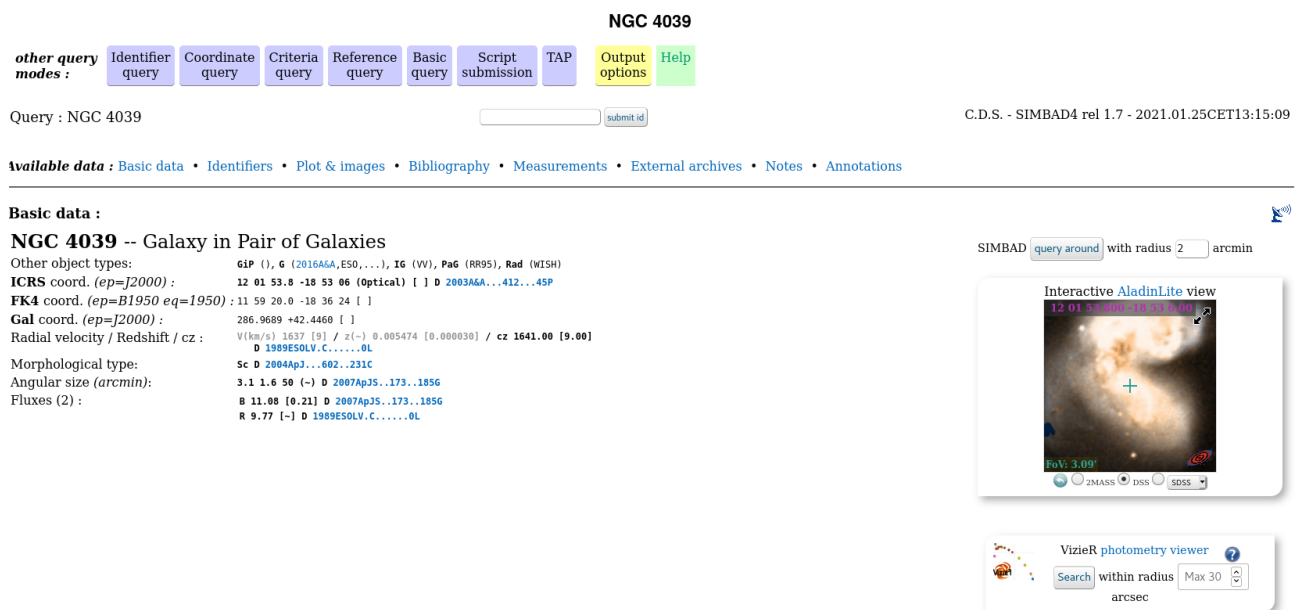

Figure 3: Result of the query for NGC4039 in SIMBAD.

- Note the list of **Other object types**. These types are drawn for the literature and are stored in SIMBAD using a hierarchical classification scheme. The full list of Object Types can be found here: http://simbad.u-strasbg.fr/simbad/sim-display?data=otypes. The main Object Type of NGC4039 is **Galaxy in Pairs of Galaxies (GiP)**.

- Use the **parents** button [parents] to identify the name of the galaxy pair. Sorting by the number of references [#ref 1850 2017] can help bring out the most important ones.

- Follow the link to the SIMBAD entry of the Antennae galaxy pair [NAME Antennae]. There click on the **children** button [children] to identify the name of the two galaxies making up the galaxy pair. Again the number of references might help to find them.

- Visit the SIMBAD entry of the interaction partner of NGC 4039 and use the **References** section to find the earliest listed reference in the literature to this object.

References (1135 between 1850 and 2021) (Total 1135)

Simbad bibliographic survey began in 1850 for stars (at least bright stars) and in 1983 for all other objects (outside the solar system).
Follow new references on this object

Reference summaries :

from: 1850   to: $currentYear

Display or select by : (not exhaustive, explanation here) In table | Title|Abstract|Keyword | Score
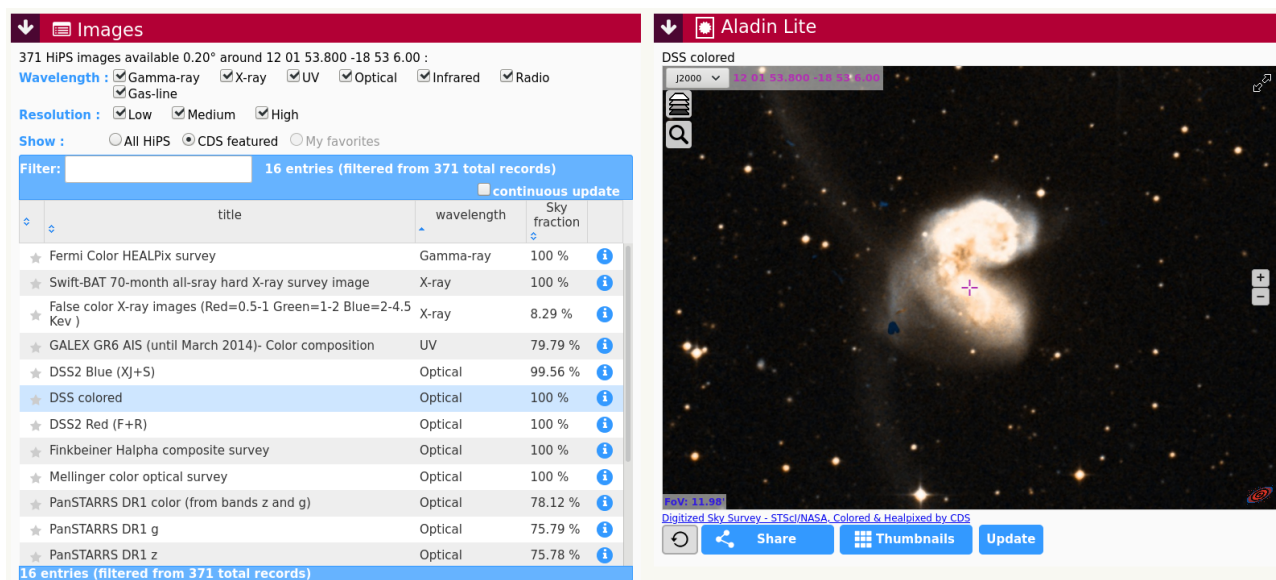
Figure 4: References section

- Return to the CDS portal.

3.2   Aladin - Images

Figure 5: Result of the query for NGC4039 in the CDS Portal

- A DSS (Digitized Sky Survey) image of NGC4039 is shown.

- Note that you can zoom in and out by scrolling your mouse when the mouse pointer is positioned on top of the images.

- Other images can be displayed by selecting them in the left column. Select the PanSTARRS DR1 and 2MASS colour images one after another to see a higher resolution image and an image at infrared wavelength of the galaxy pair.

- Filtering on wavelength and resolution is possible by ticking/unticking the boxes in that same column.

3.3 VIZIER - Catalogues
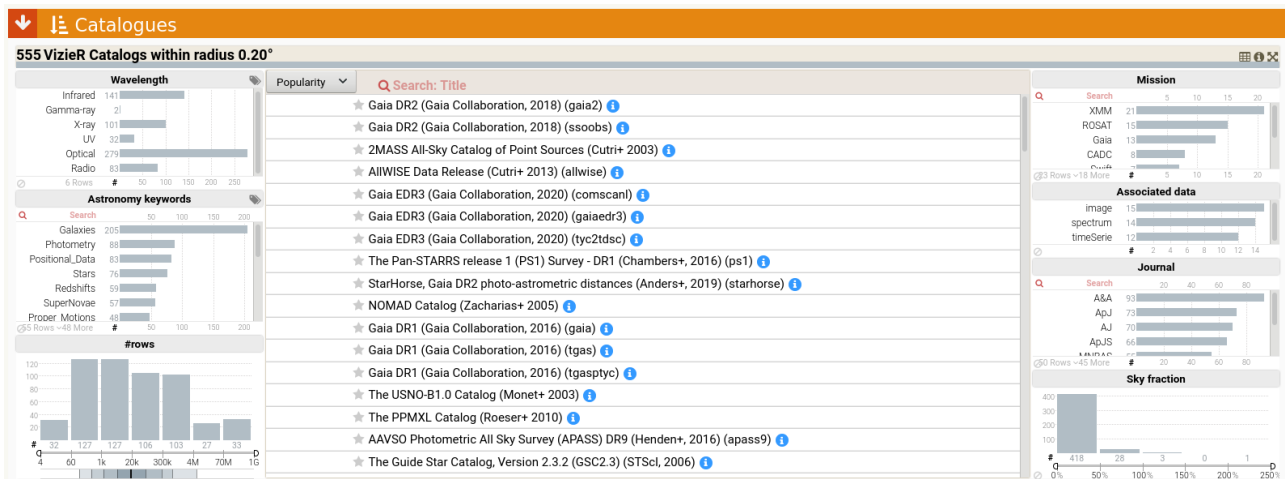
Figure 6: Result of the query for NGC4039 in the CDS Portal

- The list of catalogues is sorted by **Popularity** but can also be sorted by number of rows (**#rows**), **sky fraction** or **year**.

- Restrictions and filtering can be applied by clicking on the left or right columns according to your requirements. How many catalogues provide data in the optical for galaxies and come with images as associated data? Now go back to displaying all catalogues available within 0.2 deg of NGC4039 by clicking the `Show All` button.

- After selecting one catalogue, it is possible to look at it it (either quickly `Quickview` or in VIZIER `Vizir`), to plot column values `Plot` and to send it to Aladin (or other SAMP-aware tools like TOPCAT) `≈`.
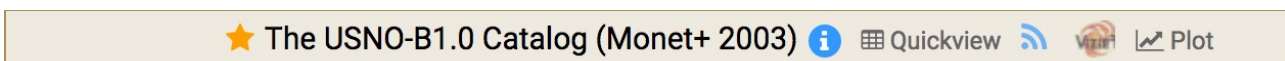
Figure 7: Selection of a VIZIER Catalogue on the CDS Portal.

- Actually the Antennae is listed in the Arp Atlas of Peculiar Galaxies, search for this catalogue by typing 'Arp' into the **Search** box.
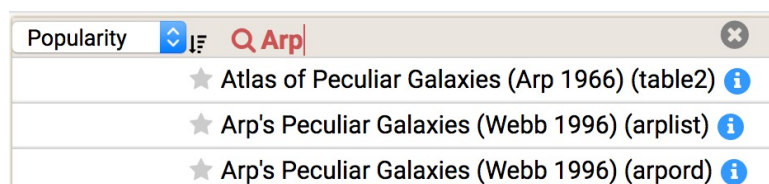
Figure 8: Looking for 'Arp' in the VIZIER catalogues.

- There are two tables for Webb 1996. Using the info button ⓘ, you'll see that the **arplist** table contains a list of all galaxies with some additional information. Select the **arplist** table by clicking on it and then on 🌀. This sends to the VIZIER detailed query page.

- Make a first query on this table by clicking on **submit**. Examine the output as html. Go back to the previous page.

- Modify the query preferences to add extra coordinate columns in J2000 decimal degrees and to obtain the whole catalogue:
  - Remove the restriction on searching only around 'NGC4039' by clicking on the **Clear** button below **Target Name (resolved by Sesame) or Position** at the top of the page.
  - In **Preferences** on the left side of the page, add extra coordinate columns in decimal degrees by ticking **J2000** and **decimal** boxes.
  - Change the maximum (**max**) number of rows to **unlimited** and tick **All columns** to get the whole catalogue.

Figure 9: **Preferences** panel.

  - Submit again.

The following steps are optional for the workflow of this tutorial, but they might come in handy for other problem sets.

- When satisfied, click on **Save in CDSportal**, then click on the **Save** button. The file is now

  saved in your personal user space on the CDS portal.
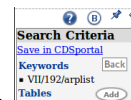
- Click on **Go to MyData** and download a copy in **VOTable** format on your desktop. This file can be used later in the tutorial.

Filename	Nb. rows	Size	Comment	Metadata	Upload date	Download ⬇
vizier_VII_192_arplist_20161220 ✎	592	176.5 KB	Your comment	Set metadata	20/12/2016 at 15:43:44	VOTable csv tsv fits

Figure 10: Personal User Space on the CDS Portal.

3.4   How do the data get into the CDS databases?

As you have seen in the last sections, the CDS databases (SIMBAD, VizieR, the HiPS server for images) host large amounts of a variety of data products. To provide these data to the astronomical community, about 12 data stewards curate the data. This involves collecting data from journal papers,

receiving data directly from large surveys (such as e. g. Gaia), assessing and assuring the data quality and cross-matching sources described by newly published papers/data with sources already available in the databases. You can make this a smooth process by making sure that all sources in your papers appear with correct coordinates and names already known by SIMBAD.

4   Search for data on NGC4039 in Aladin

Open Aladin with at least 1GB memory allocated to the save virtual machine. To do so use the following command line: `java -Xmx1024m -jar Aladin.jar`

Aladin offers two ways to retrieve data: through the DATA TREE on the left hand side of the Aladin window and through the SERVER SELECTOR window, which can be opened via **File → Open server selector...** or with **CTRL + l**.

Data sets in the DATA TREE are colour coded in green or orange depending on whether or not they are available in the region currently visible in the main viewing window.

(a) The DATA TREE: green entries are available for the sky region currently visible in the main viewing window (not shown).

(b) The SERVER SELECTOR window: here SIMBAD has been queried for galaxies within 14 arcmin around NGC 4093.

Figure 11: The two entry points to retrieving data in Aladin: The DATA TREE (left) and the SERVER SELECTOR (right).

- Start by typing "NGC4039" in the **Command** line and press **Enter**. In the main viewing window the coloured DSS image of the Antennae galaxies appears. As for the Aladin Lite window in the CDS portal you can zoom in and out by scrolling your mouse.

- Make a contour map of the image using the **cont** button next to the main viewing window. Increase the number of contours to better represent the image.

- Overlay a SIMBAD plane showing only the galaxies by selecting the SIMBAD tab in the SERVER SELECTOR and choosing **Galaxy** as the **Display filter** (see Figure 11b).

- Change the colour of the SIMBAD plane using the button **Properties**.

9

- Using the **Select** tool ![select icon], select some of the SIMBAD points either by clicking on one point or by holding the left mouse button and drawing a rectangle. The sources within the rectangle are then selected and displayed as a table below the image. This window can be detached with the icon ![detach icon] in the bottom right corner of the table window. Note that the data point belonging to a row in the table blinks in the main viewing window when hoovering over the table row with the mouse.

- In the DATA TREE, go to **Images** → **Infrared** → **2MASS** and load the 2MASS colour image. To do so click on the 2MASS colour image entry and the loading window appears (see Figure 12). In the loading window, check that the box for progressive loading is ticked ![progressive] and load the image by clicking **Load** ![Load].


Figure 12: Loading the 2MASS colour image from the DATA TREE.

- Once the 2MASS image is loaded it appears in the stack, which is located to the right of the main viewing window. The stack informs you of all planes that are currently loaded into Aladin. These layers are shown in the main viewing window as if you were looking from top to bottom through the stack. As it was most recently loaded, the 2MASS plane is now shown as the "background" image in the main viewing window. If you want to go back to the DSS colour image being the "background" image, tick the little grey box next to the DSS image entry in the stack. If you now use the opacity slider ![slider] next to the 2MASS plane in the image stack, you can change the opacity of the 2MASS plane and thus compare it easily to the DSS image. If you wanted to overlay the DSS image on the 2MASS image, you would have to change their order in the stack by drag-and-drop.

- Now add more colour images form the DATA TREE, for example the allWISE (infrared) and the XMM Newton (X-rays) colour images. Compare the images in a number of different ways:
  - Multiview: **View** → **Create one view per image**, or via the **multiview** icon at the bottom left of the Aladin image window ![multiview]. After selecting e. g. the four by four grid drag and drop the images into the panels, where you want to see them.
  - Align and scale all images by using the **match** icon below the image window ![match].
  - Transparency overlays: return to single view mode. Change the transparency of planes in the stack with the opacity sliders ![slider] as before. Remember that you can move the location of the planes in the stack and thus change the order in which the image are shown.

- All the colour images we have loaded so far are pre-computed. Aladin also allows you to compute your own RGB colour images:
  - For this task, we need fits images rather than the HiPS tiles we have worked with so far. Those fits images may come from your own data or from the Virtual Observatory. We will now search and load images at the location of the Antennae galaxies from the VO using the simple image access (SIA) protocol.

– First, we look for a far-ultraviolet image from *GALEX*: Enter "GALEX Atlas of nearby galaxies" in the **Select** line below the DATA TREE ![select GALEX Atlas of nearby galaxies]. In the DATA TREE you will find the entry "GALEX Atlas of nearby galaxies" under **Others** → **SIA (image)** → **mast.stsci**. Load this entry **in view** ![in view]. A table appears in the stack.

– The newly loaded table contains a row for every image available from this image service in the vicinity of NGC 4039. Find the entry for the file "h_ngc4038-fd-int.fits.gz" and click on the URL in the first column. This will load the corresponding fits image into your stack.

– Next, we are after an infrared image: search for "2MASS large galaxy atlas" in the **Select** line, load the table from **Others** → **SIA (image)** → **irsa.ipac** → **2MASS Large Galaxy Atlas** and select any one of the three entries in the table to load into the stack.

– Finally let's add an X-ray image to the collection: search for "Chandra X-ray Observatory" in the **Select** line. There will be two services available, we will want to use the SIA service. Load the table "in view" from **Others** → **SIA (image)** → **cxc.harvard.edu** → **Chandra X-ray Observatory Data Archive** and select one of the observations with the longest "exptime".

– To create the RGB image, open the **RGB image generator** with ![rgb icon] (see Figure 13a). Enter the images you want to use for each colour component and click **CREATE**. If you want to adjust flux scale and transfer function of the individual images, use the **pixel** ![pixel icon] menu.
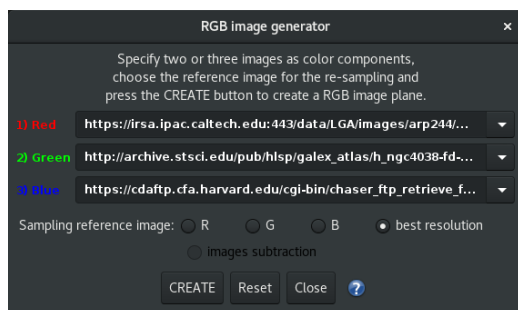

(a) The RGB image generator window.
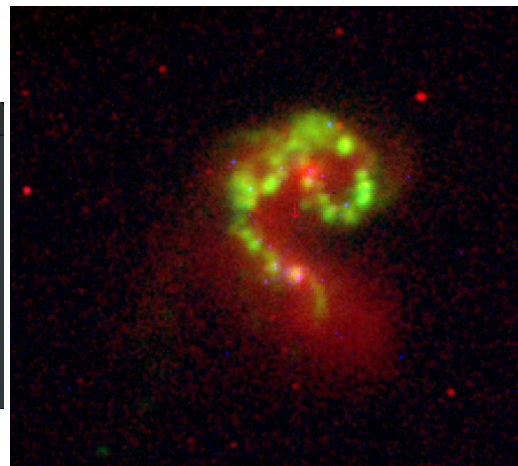
(b) The resulting RGB image for the Antennae galaxies.

Figure 13: Creating RGB images in Aladin.

- Search for more data from the VO on NGC4039:

  – To restrict the collections that the DATA TREE is showing, use the **Select** line below the DATA TREE ![select transient]. For example when searching for transients within the Antennae galaxies you might restrict the search to "transient" and find that the Palomar Transient Factory photometric catalogue has data for transient events in the area. Load the catalogue and find the data points in the main viewing window.

  – More elaborate filters can be created using the **Data Discovery Tree Filter** window, accessed through ![filter icon] next to the **Select** line.

– As before, entries in the DATA TREE, which are coloured in green, have data available in the region currently visible in the main viewing window.

• Now delete all the planes in your stack before continuing. This is not mandatory but it will free some useful memory space and allow you to proceed more easily with the next steps.
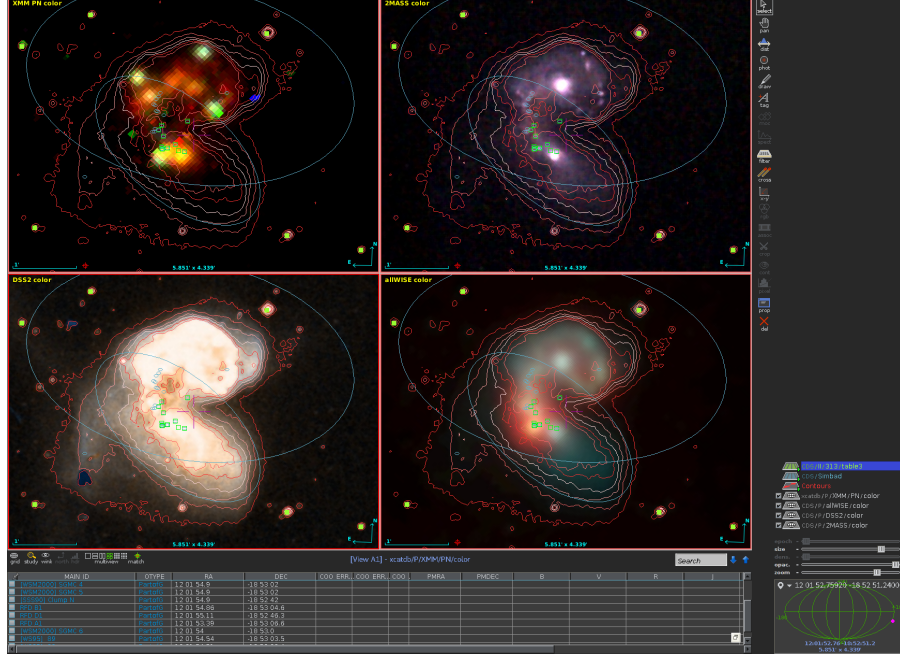

Figure 14: Results of all the above mentioned steps.

5 Compare the coverage of Sky Surveys and select interacting galaxies that have SDSS and GALEX data

Many large sky surveys in Aladin are stored in the **HiPS** format (Hierarchical Progressive Survey), which allows for easy access, browsing and visualisation of image and catalogue data. To describe (non-trivially shaped) regions on the sky **MOC** (Multi-Order Coverage) maps are used. In the following, we will make use of the advantages of these two data structures to easily asses which galaxies in the Arp catalogue of peculiar galaxies have been observed by SDSS and GALEX. If you would like to have an overview of currently available HiPS, have a look at this list. You can load any of them in Aladin by entering their base URL in the **Command** line.

• Select the **SDSS 9 colored** (**Image** → **Optical** → **SDSS**) and **GALEX All Sky Imaging Survey colored** (**Image** → **UV** → **GALEX**) surveys in the DATA TREE and load both the imaging data (tick progressive ✔ progressive) and the MOC of the survey (tick coverage ✔ space cov.). For the moment make the two MOC planes invisible by clicking on their opacity slider.
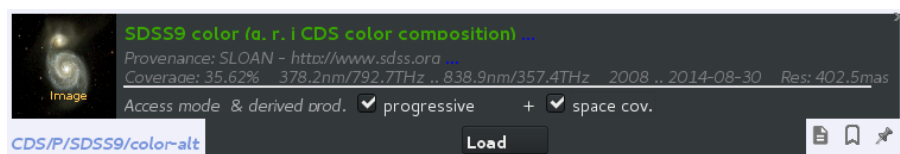

Figure 15: Selecting the SDSS 9 coloured survey and MOC

- Turn on the coordinate **grid** , zoom out and use the **pan** tool to explore the whole sky.

- Now turn the MOC planes back on (click again on the opacity slider). Zoom onto the edges of the surveys and note the way the MOC represents the coverage of the surveys.

- Calculate the intersection of the coverage maps of the SDSS and GALEX surveys using in the menu the item **Coverage** → **Logical operations** or the MOC button on the right of the main viewing window.
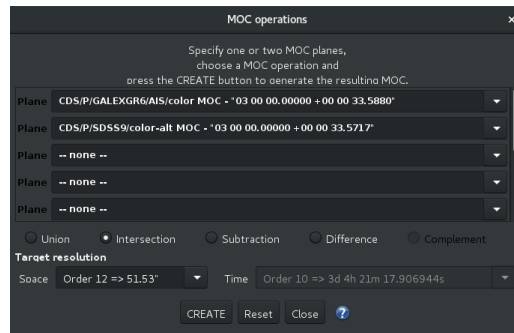

Figure 16: Building the intersection of the coverage maps

- Load the full Webb 1996 Arp catalogue that you saved earlier. You can do this in one of the following three ways:

  - **File** → **Load local file...** .
  - Drag and drop on the stack.
  - Broadcast the catalogue from the VizieR page that shows the HTML version of the table or **MyData** area of the CDS Portal. To do so click on or respectively and allow the browser to connect to the SAMP hub. For VizieR you additionally have to click on the newly appeared grey **Broadcast** button and in your **MyData** area of the CDS Portal you might have to click the same broadcast button again after granting your browser permission to connect to the SAMP hub. Now the data appears in Aladin.

  Alternatively you can download the catalogue from VizieR through the DATA TREE, enter "Arp Webb" in the **Select** line below the DATA TREE, select the "arplist" table and **Load** the **Whole data**.

- Filter the catalogue to select only the sources that fall within the SDSS+GALEX MOC: **Coverage** → **Filter a table by MOC...**: a new catalogue plane appears and the filtered catalogue contains 425 sources. You can find the number for example in the little information text that appears on the top of the stack when hovering the mouse over an entry in the stack. Note that you can also filter a table by MOC when loading the table initially (instead of loading the entire table).
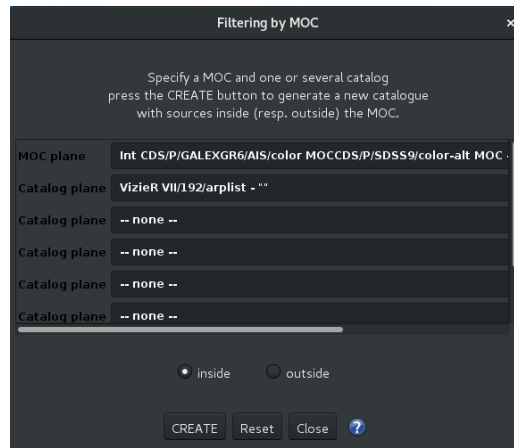
Figure 17: Filtering a catalogue by a MOC

- Visualize the brightest (<9 mag) galaxies of the selected sources by extracting small images from the SDSS survey:

  – Select the brightest galaxies using the **filter** tool ▨: select the **Show brightest stars** predefined filter and edit it with the **Advanced mode** to select object with magnitude below 9. Note that the column is automatically identified with the Unified Content Descriptor ' phot.mag* '.

  – Make sure that only the MOC filtered catalogue is active in the stack and visible in the main viewing window (or other sources may also be filtered). An easy way to ensure that is to delete every catalogue not useful any more.

  – Click on **Apply** and then **Export** to create a new plane consisting only of sources selected by the filter. There are 7 sources with magnitudes below 9.
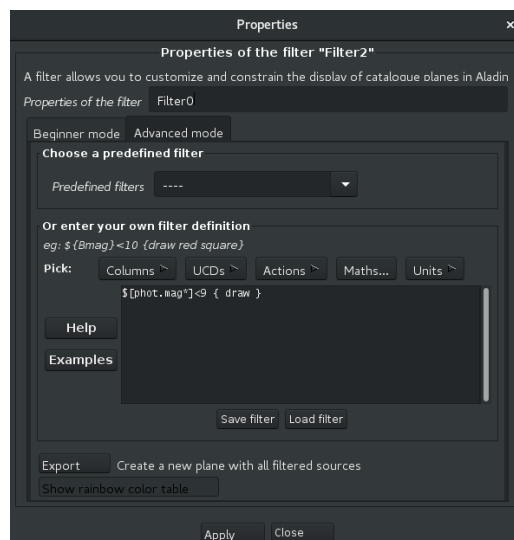

Figure 18: Filtering a catalogue by magnitude

- Make thumbnails of the selected brightest sources: **Tool** → **Thumbnail view generator...**, set the thumbnail size to 14 arcmin.
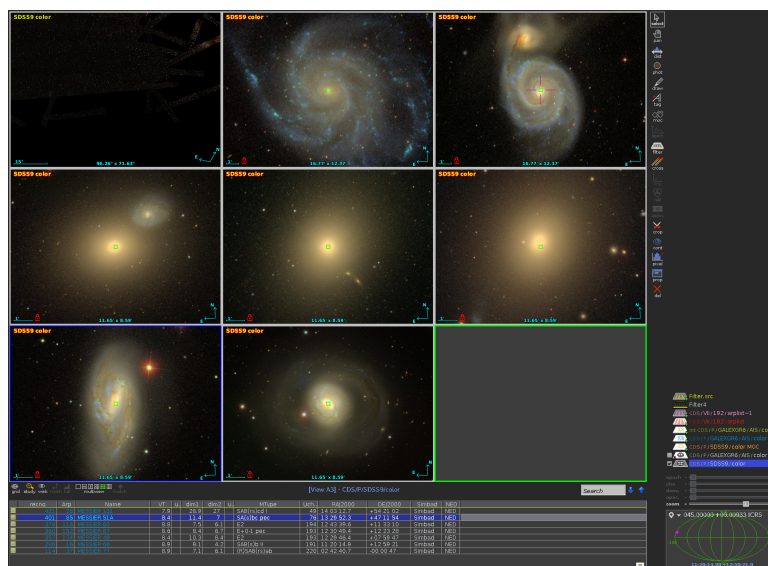
Figure 19: Thumbnails

6 Automation [Optional]

When samples become larger, it is useful to automate some tasks. In this optional section we will explore how to execute some tasks with scripts.

6.1 Collect information on a sample of galaxies using Aladin

Use an Aladin script to obtain DSS and SDSS image with HST, Chandra, ESO observation log overlays for each of the selected bright galaxies.

- Copy the script Arp_script.ajs from the url
  http://cds.unistra.fr/tutorials/CDS-tutorial/Arp_script.ajs
  to your computer or create a new file with the script shown at the end of this section.

- Create a folder called 'Arp' and edit the script Arp_script.ajs to insert a path in order to save the output files, e.g. ~/Desktop/Arp.

- Open the Aladin macro controller and load the script:
  - **Tools → Macro Controller** then **File → Load script**
  - or cut and paste the script into the top panel of the MACROS window.

- Select all the sources in the bright galaxy catalog:
  - Right click on the plane and **Select all objects in the selected planes**
  - In the MACROS window: **File → Use selected plane sources as params**.
  - Note how the catalogue columns are shown as parameters which can be referred to as $1, $2, etc within the script.

- Click on the first row of the parameters table and execute the script for this row: **Exec current params**.

- Optional: add an SDSS image: remove the '#' to enable download of a SDSS g-band image for each source. Note that this results in an **Could not find any data corresponding to your request** message for objects not covered by SDSS.

- Inspect the output in the Aladin window and also the files written in the Arp folder.

15

- Execute the script for all sources: **Exec all from current**.

- Note that the saved stack files (∗ stack.aj) can simply be dragged and dropped into Aladin for inspection.

This is the script:

```
#AJS
#
reset
grid on

# DSS image:
"ARP-$2_DSS" = get DSS.STScI(POSS2UKSTU_Red,15,15) $3

# SDSS image:
# "ARP-$2_SDSS" = get SDSS(keyword=Filter g) $3

# SIMBAD plane
# "ARP-$2_Simbad" = get Simbad $3 5'

# Observation Logs
viz_logHST=get vizier(logHST) $3
viz_logESO=get vizier(logESO) $3
viz_logChandra=get vizier(logChandra) $3

sync
pause 1

# Write results to files
# export B/hst/hstlog /Path/To/Datafolder/Arp/Arp-$2_HST.xml
save /Path/To/Datafolder/Arp/Arp-$2_chart.png
backup /Path/To/Datafolder/Arp/Arp-$2_stack.aj
#adapt ^^^^^^^^^^^^^^^^^^^^ to your own path
```
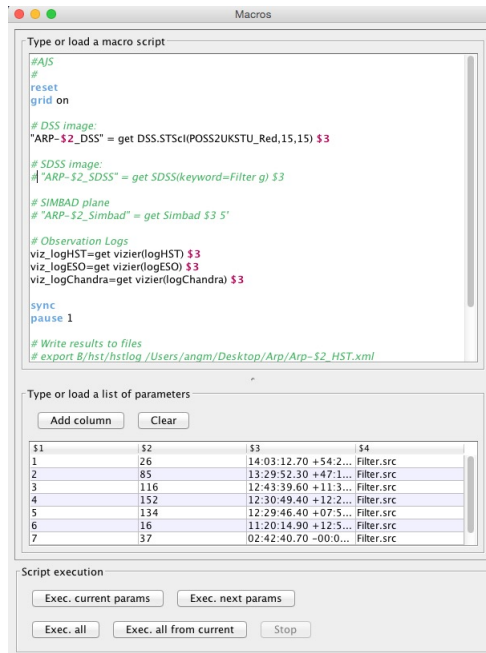
Figure 20: MACROS window & Arp_script.ajs

6.2 Visualise the distribution of Arp's galaxies on the Sky

The scripting language Python is becoming more and more widely used by Astronomers. Consequently, more and more Virtual Observatory services are available through Python. To explore some of the features that Python can provide today, a Jupyter notebook was prepared. These notebooks allow interactive execution of Python code. Go now to

https://mybinder.org/v2/gh/cds-astro/tutorials/master?filepath=Notebooks

and wait a moment for everything to load. Click on the link to

The_CDS_tutorial.ipynb

and follow the instructions there. You can also download the Jupyter notebook from github repository to your machine. It was created to run with the packages as listed in the requirements.txt file. Using pip[1] you can install all these packages with the command `pip install [name of the package]`.

---

[1]https://pip.pypa.io/en/stable/