Highlights on Spanish Astrophysics X, Proceedings of the XIII Scientific Meeting of the Spanish Astronomical Society held on July 16 – 20, 2018, in Salamanca, Spain. B. Montesinos, A. Asensio Ramos, F. Buitrago, R. Schödel, E. Villaver, S. Pérez-Hoyos, I. Ordóñez-Etxeberria (eds.), 2019

The Gran Telescopio Canarias Archive.

Enrique Solano[1], **José Manuel Alacid**[2]

[1] Departmento de Astrofísica, Centro de Astrobiología (INTA-CSIC), ESAC Campus, Camino Bajo del Castillo s/n, E-28692 Villanueva de la Cañada, Spain.
[2] Spanish Virtual Observatory.

Abstract

The Gran Telescopio Canarias (GTC) archive is operational since November 2011. The archive, developed and maintained at Centro de Astrobiología (INTA-CSIC) in the framework of the Spanish Virtual Observatory project, provides access to raw, calibration and science ready data and has been designed in compliance with the standards defined by the International Virtual Observatory Alliance (IVOA) to guarantee a high level of data accessibility and handling.

In this presentation I will describe the main capabilities the GTC archive offers to the community, in terms of functionalities and data collections, to carry out an efficient scientific exploitation of GTC data.

1 Introduction

The GTC archive is not something new but there is a long history behind it. In fact, the first GTC archive meeting took place in 2003, fifteen years ago. Already at that time, the need of a Virtual Observatory (VO) compliant archive to ensure the long term preservation as well as the optimum scientific exploitation of GTC data was clearly identified.

Eight years later, in November 2011, the first operational version of the archive was released[1]. The GTC archive has always been managed in the framework of the Spanish Virtual Observatory[2] and is part of the CAB Scientific Data Centre. In this context it is important to remark that the GTC archive is the only entry point for the national and international astronomical community to access GTC data.

At the time of writing the GTC archive contains more than 160 000 science files of five instruments (OSIRIS, CanariCam, CIRCE, EMIR and HIPERCAM) and is intensively used

[1] http://agencias.abc.es/agencias/noticia.asp?noticia=1002094
[2] http://svo.cab.inta-csic.es

by the community as demonstrated by the fact that users from the five continents have made hundreds of queries to download thousands of files.

2 The GTC archive. Characteristics and functionalities

2.1 Data transfer and ingestion

GTC data are periodically transferred through a secure connection from the telescope in La Palma to the archive in Madrid. At their arrival, a number of quality control tests are carried out to ensure data integrity and metadata coherence. Problematic files are inspected and remedial actions are taken in agreement with GTC staff. After passing the tests, metadata information is extracted from the FITS header and stored in a database while FITS files are moved into the data storage system. Once data and metadata have been successfully ingested, they are automatically available to the general public through the web a VO interfaces.

In order to guarantee a safe long-term preservation of GTC data, a well-defined backup and data safekeeping policy have been established.

As of January 2018 the GTC archive also provides access to private data, identified according to their observing date. Principal Investigators are provided with a login/password to access their private data. They are also emailed whenever new data of their programmes arrive. At present (October 2018) the archive contains 93 private programmes with more than 9 000 science files.

2.2 Web interface. Data query

The GTC archive is friendly enough to be potentially used by a wide variety of users. From the home page, the user can get information on the archive through three different channels: Reading the system overview section, checking the FAQ system or directly asking GTC archive staff using the available HelpDesk.

The input query form is very simple and allows typical queries by list of objects names and/or coordinates, observing date, type of instrumentation, and programme. Special programmes like the ESO-GTC or the GTC Large Programmes have their specific entry access.

In addition to basic metadata, the table of results obtained after a query also includes links to the paper where the observations have been published as well as to NED and the CDS portal. These are powerful functionalities to easily get additional information on a GTC object. The output fields can be ordered by date, instrument, program or product identifier and the number of results shown per page can also be customized.

As said in the abstract, one of the main requirements of the GTC archive is its compliance with the Virtual Observatory, with the objecting of taking advantage of the VO standards and protocols. The Multi-Object Coverage (MOC[3]) is one of these standards. It helps users in the identification of the sky regions covered by GTC observations. MOC not only facilitates the visualisation of the footprint of a given survey but also, making use of a

[3]http://www.ivoa.net/documents/MOC/20140602/REC-MOC-1.0-20140602.pdf

VO tool like Aladin, permits operations to know, for instance, the region in common between two surveys or the part of the sky that contains observations of only one of the surveys.

A remarkable strength of the VO protocols is their ability to discover information using VO tools. By providing seamless access to thousands of archives and services, VO-compliance facilitates research projects that, otherwise, would be extremely difficult to perform outside the Virtual Observatory. In this sense, SIAP (Simple Image Access Protocol) has been implemented to access GTC images from VO tools.

Figure 1: GTC archive query form

2.3 Reduced data

Reduced data are of fundamental importance for archives as they enhance their use by the community and provide a higher visibility of the project results. The bad news is that there is no official reduction pipeline for the GTC instruments. To overcome this problem three channels have been defined to feed the GTC archive with science-ready data products.

- Reduced data provided by the community: On a monthly basis, GTC archive staff query ADS looking for papers containing GTC observations. Once identified, we contact the first author inviting him to send us their reduced data using a system similar to ESO-Phase 3 but much simpler. In short, the users have to submit a single FITS file or a

zip file with their data in FITS format in the case of multiple products. As soon as reduced data are ingested in the archive, they are linked to the associated raw products and are available for downloading.

At the time of writing (October 2018), we have identified 355 publications containing GTC data and have ingested reduced data for 103 ($\sim$ 30%).

- Use of existing pipelines: In collaboration with GTC staff, we are working in the use of existing pipelines to reduce in a systematic and homogeneous way data from an observing mode or instrument. This is, for instance, the case of CanariCam and EMIR.
- Provision of high level data products: The goal of this line of work is to generate high level products (astrometrically and photometrically corrected images and the associated source catalogues) of the broad-band images obtained with the instrument OSIRIS. We have used 2MASS-PSC as the reference catalogue for astrometry while Pan-STARRS (PS1) was used for the photometric calibration. Corrected images and the associated catalogues will be publicly available soon.

3 Conclusions

The main ideas described in this paper are briefly summarized here:

- The GTC archive is a research resource in operation since 2011. A lot of science can already be done with the current archive contents and functonalities.
- It contains raw, calibration as well as reduced data provided by the community.
- New high level data products (i.e. source catalogues, images processes in a standard and automated way) will be available soon.