PROCEEDINGS OF SCIENCE

GOLIA: an INTEGRAL Archive @INAF-IASF Milano*

Adamantia Paizis[†‡], **Sandro Mereghetti**
INAF-IASF Milano, Italy
E-mail: ada@iasf-milano.inaf.it
E-mail: sandro@iasf-milano.inaf.it,

Diego Götz
AIM/Service d'Astrophysique - CEA Saclay, France
E-mail: dgotz@cea.fr

Mark Gaber
Former ISDC, Versoix, Switzerland
E-mail: mark.gaber@bluewin.ch

Mauro Fiorini, Roberto Regni Ponzeveroni, Lara Sidoli
INAF-IASF Milano, Italy
E-mail: fiorini@iasf-milano.inaf.it
E-mail: regni@iasf-milano.inaf.it
E-mail: sidoli@iasf-milano.inaf.it

Stefano Vercellone
INAF-IASF Palermo, Italy
E-mail: stefano@ifc.inaf.it

We present the archive of the *INTEGRAL* data developed and maintained at INAF-IASF Milano. The archive comprises the public data currently available (∼8.5 years of data). *INTEGRAL* data are downloaded from the ISDC, Geneva, on a regular basis as they become public and a customized analysis using the OSA 9.0 software package is routinely performed on the IBIS/ISGRI data. The scientific products include individual pointing images and the associated detected source lists in several energy bands, as well as light-curves binned over 100 s in the 17–30 keV band for sources of interest. The whole database (raw data, products, ad-hoc documentation and tools) enables a local interactive and easy access to the hard X-ray long-term behavior of a vast sample of sources. We make the analysis tools used to build such an archive publicly available.

An INTEGRAL view of the high-energy sky (the first 10 years) - 9th INTEGRAL Workshop and celebration of the 10th anniversary of the launch
15-19 October 2012
Bibliotheque Nationale de France, Paris, France

*Based on observations with *INTEGRAL*, an ESA project with instruments and science data centre funded by ESA member states (especially the PI countries: Denmark, France, Germany, Italy, Spain, and Switzerland), Czech Republic and Poland, and with the participation of Russia and the USA.

†Speaker.

‡AP thanks the ISDC staff since most of the know-how needed to set-up such an archive was gained during her years at the ISDC (2000-2005). We acknowledge the Italian Space Agency financial support via contract *INTEGRAL* ASI/INAF I/033/10/0.

PoS(INTEGRAL 2012)151

1. Introduction

The INTErnational Gamma-Ray Astrophysics Laboratory, *INTEGRAL* (Winkler et al. 2003), is a medium sized ESA mission successfully launched on October 17th, 2002. Its payload consists of two main gamma-ray instruments, the spectrometer SPI (Vedrenne et al. 2003) and the imager IBIS (Ubertini et al. 2003), covering the 15 keV – 10 MeV band. IBIS consists of two layers, the lower energy one, IBIS/ISGRI (15 keV – 1 MeV, Lebrun et al. 2003) and the higher energy one, IBIS/PICsIT (0.175–10 MeV, Labanti et al. 2003). Co-aligned to SPI and IBIS are two X–ray monitors JEM–X (4–35 keV, Lund et al. 2003) and an optical monitor OMC (500–600 nm, Mas-Hesse et al. 2003).

In order to increment and ease the exploitation of *INTEGRAL* data at INAF-IASF Milano, we undertook the task of preparing and maintaining an *INTEGRAL* archive, GOLIA (Giant On-Line *INTEGRAL* Archive). GOLIA's aims are twofold: provide a local database of the available public data to facilitate personalized analysis of already downloaded data, and offer easy-to-browse IBIS/ISGRI data products for a quick and efficient view of the hard X–ray sky. All of which is locally and interactively available at INAF-IASF Milano. The scripts we have used to build GOLIA are publicly available. See Paizis et al. (2013) for a detailed description of GOLIA.

2. A walk through GOLIA

2.1 Building the archive (archive *Owner*)

All the scripts described in this section are downloadable from our web page[1].
Our IBIS/ISGRI processing and post-processing steps, performed by the archive *Owner*, are described below.
1. Download: *INTEGRAL* public data are downloaded via the ISDC Data Centre for Astrophysics (Courvoisier et al. 2003)[2].
2. Process: the standard OSA 9 software package is run on the public data of the IBIS/ISGRI instrument, looping over the Science Windows (hereafter ScWs) via the Shell scripts `launch_IMA.sh` and `launch_LCR.sh`. The following steps are performed:
2-1 Image reconstruction (IMA): for each pointing a standard analysis is performed in four energy bands: 17–30, 30–50, 17–50 and 50-100 keV. The results of this analysis are the starting point for pointing-based, $\sim$2 ks, light-curves and maps/mosaics for a deeper analysis.
2-2 Catalogue creation: for each ScW, a catalogue is created merging the bright sources discovered from the previous IMA step (with detection significance higher than 3) and sources of interest, *regardless* their detection level. This allows to detect short bursts from interesting sources that would be missed in a single pointing ($\sim$2ks). Successful tests have been made on known GRBs.
2-3 Light-curve extraction (LCR): 17–30 keV standard 100 s bin light-curve extraction is performed for all the sources of each ScW catalogue, including sources of interest not necessarily detected in the $\sim$2 ks image.
2-4 Clean and close down: relevant files are zipped and intermediate level products canceled.

[1] http://www.iasf-milano.inaf.it/$\sim$ada/GOLIA.html
[2] http://www.isdc.unige.ch/integral/archive#DataRelease

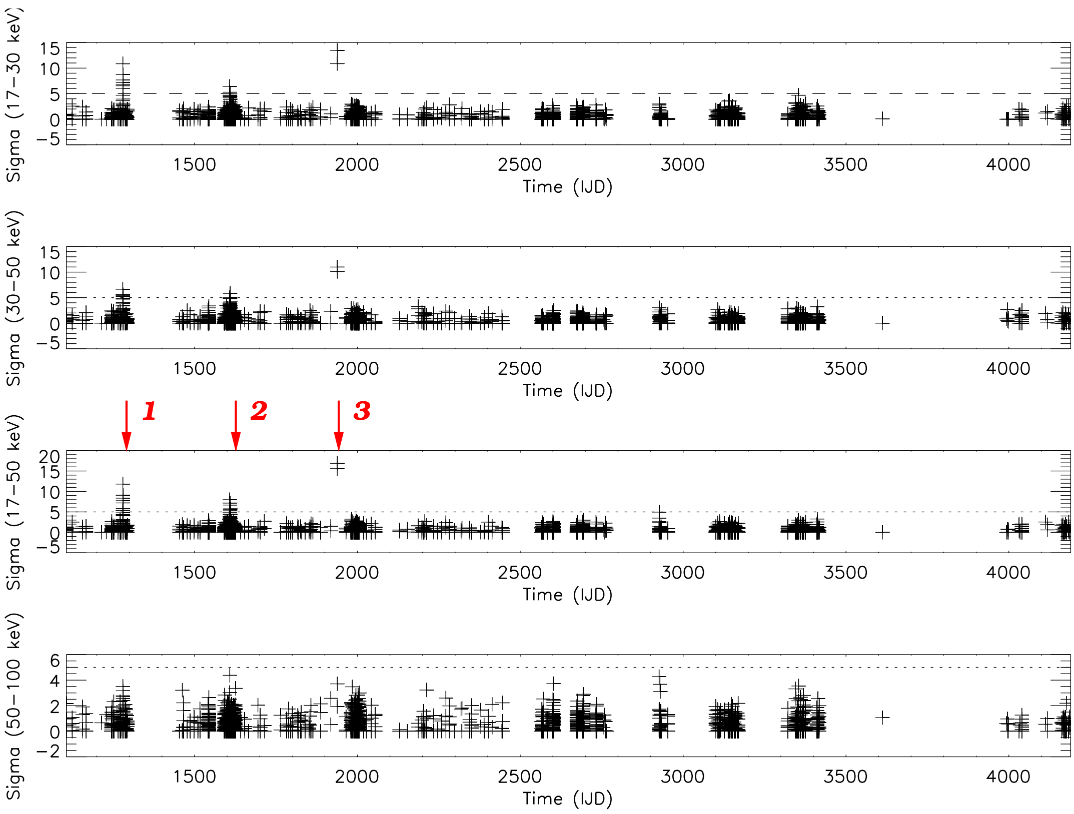

Figure 1: Whole-archive ∼2 ks light-curve obtained for IGR J11215−5952. The plot shows all the *INTEGRAL*/IBIS public available data: 2442 pointings, 4.8 Ms, from revolution 0030 (Jan 2003, IJD∼1107) to 1060 (Jun 2011, IJD∼4190). Sigma ≧5 (horizontal line) is a detection (from Paizis et al. 2013).

3. Post-process: for each pointing, the IMA analysis step produces images in several bands and lists (FITS files) with position, flux and significance of the sources. Since a map file and a source list file are derived for each ScW, all the information about a given source is spread over many ScWs. The OSA 9 tool src_collect allows to collect IMA source results into a single table. However it enables to retrieve results for one source at a time and the results are saved in a FITS table. We have developed a PERL script (ima_pick.pl, also available on our web page) that extracts in one go single ASCII files for *all* the sources present in the archive, one file per source. Each ASCII file contains the results of the imaging step. The tool can be run in a cumulative way (the ASCII files can be updated with the new incoming results) and is the starting point for quick visualization by the *Users*.

2.2 Using the archive (archive *User*)

Each *User* can access GOLIA in many ways. The mere **Download** provides *INTEGRAL* public data for a personalized analysis, avoiding multiple time and space consuming downloads. Once the **Process** step is done by the *Owner*, the *User* can access analyzed ScW images to build customized mosaics and source search for a deeper analysis. For bright sources (or sources of interest) overall light-curves binned at 100 s can be collected with the ISDC tool lc_pick (see Fig. 2 left, and

Section 3). The stacking of the products of the "Process" step (mosaics and 100 s bin lightcurve collection) can be time consuming if a very large number of ScWs is involved, as in standard OSA usage. The ASCII files produced via the Post-process step are easily readable / plottable by several software tools. We note that the introduction of the **Post-process** step, via the ima_pick tool (not part of the OSA package), leads to a huge simplification in the visualization of the image products. The produced ASCII files are a very easy-to-handle starting point and indeed whole-archive (8.5 year) light-curves can be extracted in matter of seconds for whichever source originally in the catalogue. We have developed two interactive IDL tools (readarcres.pro, readnew.pro) to read and plot such files. The former tool allows us to handle / plot single source ASCII files, providing long-term light-curves in terms of count-rate or detection significance (as in Fig. 1 for the detection significance case). The latter, readnew, allows us to investigate the NEW sources selecting them by coordinates. Indeed, ima_pick collects the results based on the source name, creating a SourceName.dat file. Every time the analysis software detects a new source, the name *NEW_1* is given (with incremental number for more new sources), hence in the whole catalogue we obtain several *NEW_1*, *NEW_2*, etc., sources that are *not* the same, but have different positions. All these sources will be collected in the same ASCII file by ima_pick and the readnew script enables the search of NEW sources by coordinates, grasping the occurrences of each NEW source in the whole archive (see Fig. 2 right, and Section 3).

3. GOLIA: a starting point for new investigations

In this section we give examples of published and work-in-progress scientific results based on GOLIA.

Supergiant Fast X-ray Transients (SFXTs) Fig. 1 is an example of a whole-archive light-curve obtained with our interactive IDL script. The plot shows all the IBIS/ISGRI publicly available data of the transient IGR J11215−5952. With such a long-term view, outbursts, trends, periodicities can be easily spotted. Indeed our archive allowed us to discover that the Supergiant Fast X-ray Transient IGR J11215−5952 displays periodic outbursts (Sidoli et al. 2006). The source was discovered in 2005 (arrow 3 in Fig. 1, Lubiński et al. 2005), but thanks to our systematic re-analysis of *INTEGRAL* archival observations, we could discover two previously unnoticed outbursts (1 and 2 in Fig. 1). This discovery triggered by our archive has led to further investigations, such as the study of the source orbital period (Sidoli et al. 2007), the development of a stellar wind model for OB supergiants to explain the transient and variable nature of SFXTs (Ducci et al. 2009), and its application to public *INTEGRAL* observations on 14 SFXTs (Ducci et al. 2010).

Low Mass X-ray Binaries (LMXB) Visual inspection of light-curves such as Fig. 1 for persistently bright neutron star (NS) LMXBs has been the starting point for the systematic study of the evolution of the transient hard-tails in these sources (e.g. Paizis et al. 2006a). The origin of such tails dominating the spectra above ~30 keV is still debated. Paizis et al. (2006a) proposed for the first time a qualitative unified physical scenario to explain the spectral evolution of NS LMXBs, including the peculiar transient hard tail. In order to study in a quantitative way the evolution of the parameters describing the innermost physical conditions of NS LMXBs, a new Comptonization model was developed (compTB[3], Farinelli et al. 2008), and successfully applied to a sample of

[3]http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/models/comptb.html

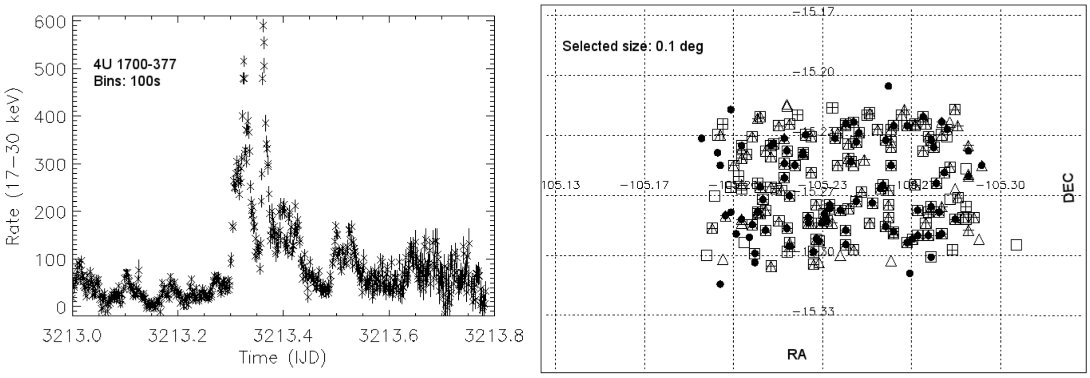

Figure 2: *Left*: 0.8 day IBIS/ISGRI light-curve of 4U 1700–377 (100 s bins). 1 Crab corresponds to a rate of about 173 counts/s in the same band. *Right:* multiple IBIS/ISGRI detections of the recently discovered transient MAXI J1659−152. From Paizis et al. (2013).

bright NS LMXBs studied with *INTEGRAL* (Mainardi et al. 2010).

Rapid (∼100 s) variability In the archive, 100 s bin light-curves in the 17–30 keV band are available for all the sources detected in the IMA part as well for sources of interest. Among others, we have extracted a whole-archive light-curve for the high mass X-ray binary 4U 1700–377, for a total of 18 Ms, 100 s bins. A 0.8 day zoom of this light-curve is shown in Fig. 2, left. The complete light-curve is the starting point for a detailed study of the fast variability over 8.5 years of *INTEGRAL* data in this peculiar source (Sidoli & Paizis, in preparation).

New Sources As discussed in Section 2.2, we have developed an interactive tool that reads the NEW source ASCII files created by ima_pick, grouping them by coordinates (readnew). Fig. 2, right panel, shows multiple detections (from different bands and pointings) of a new source, the recently discovered transient MAXI J1659−152 (Negoro et al. 2010), not yet included in the *INTEGRAL* General Reference Catalogue, hence labeled as NEW by the software. We can trace back in the *INTEGRAL*/IBIS archive sources that were discovered after the creation of the *INTEGRAL* General Reference Catalogue used, dated 2010 April 8th, as well as discover previously unnoticed new *INTEGRAL* sources, such as IGR J15283–4443, discovered by our systematic IBIS/ISGRI analysis (Paizis et al. 2006b) and later on confirmed by *Swift* follow-up observations (Rodriguez et al. 2010).

4. Build your own GOLIA

In this section we report the main characteristics of GOLIA, in order to give a quantitative idea of the effort involved in such a work. The whole analysis has been made on a single server HP ML330G6, openSUSE 11.4, two processors Xenon Quad-Core E5506, 2.13GHz and RAM 8 GB (June 2011 cost ∼2300 Euro). A total of 12 external disks of 1 Tb each have been purchased to store the data and the results (213 Euro each Tb). Currently a total of 83733 pointings have been analyzed, i.e. starting revolution 0026 (December 2002) to 1079 (August 2011), for a total of about 8.5 years of data. Most disk space is needed for the raw data, the results themselves

PoS(INTEGRAL 2012)151

occupy a smaller fraction (~1.3 Tb for 8.5 years of results versus about 4.8 Tb of 8.5 years of data - excluding SPI raw data). The integrated IBIS/ISGRI analysis time of the 8.5 years of data has been about 8 months with one server (roughly half of which for the IMA part and half for the LCR one). More servers (that can be used in parallel with GRID, the analysis script modular structure allows it) will result in a much lower processing time.

5. Conclusions

Once the setup is up and running, the maintenance of the archive is not an issue and neither is its regular update. Maps, light-curves, raw data, fits and ASCII files can all be accessed by the local *users*, enabling a personalized usage of the archive and triggering further detailed investigations such as the ones previously described. Long term light-curves can be easily extracted also by non-*INTEGRAL* experts (matter of a few seconds for the 2 ks binning) and a systematic view of different classes of sources can be performed, enabling interesting discoveries and trends, as well as providing a complementary view to other on-going missions.The scripts we have used to build GOLIA are available at our web page (URL given in footnote 1). See Paizis et al. (2013) for a detailed description of GOLIA.

References

[1] Courvoisier, T.J.L., Walter, R., Beckmann, V., et al. 2003, A&A, 411, L53

[2] Ducci, L., Sidoli, L., Mereghetti, S., et al. 2009, MNRAS, 398, 2152

[3] Ducci, L., Sidoli, L., & Paizis, A. 2010, MNRAS, 408, 1540

[4] Farinelli, R., Titarchuk, L., Paizis, A., et al. 2008, ApJ, 680, 602

[5] Labanti, C., Di Cocco, G., Ferro, G., et al. 2003, A&A, 411, L149

[6] Lebrun, F., Leray, J. P., Lavocat, P., et al. 2003, A&A, 411, L141

[7] Lubiński, P., Cadolle Bel, M., von Kienlin, A., et al. 2005, The Astronomer's Telegram, 469

[8] Lund, N., Budtz-Jørgensen, C., Westergaard, N. J., et al. 2003, A&A, 411, L231

[9] Mainardi, L. I., Paizis, A., Farinelli, R., et al. 2010, A&A, 512, A57

[10] Mas-Hesse, J. M., Giménez, A., Culhane, J. L., et al. 2003, A&A, 411, L261

[11] Negoro, H., Yamaoka, K., Nakahira, S., et al. 2010, The Astronomer's Telegram, 2873

[12] Paizis, A., Mereghetti, S., Götz, D., et al. 2013, Accepted in Astronomy and Computing

[13] Paizis, A., Farinelli, R., Titarchuk, L., et al. 2006a, A&A, 459, 187

[14] Paizis, A., Gotz, D., Sidoli, et al. 2006b, The Astronomer's Telegram, 865

[15] Rodriguez, J., Tomsick, J. A., & Bodaghee, A. 2010, A&A, 517, A14

[16] Sidoli, L., Paizis, A., & Mereghetti, S. 2006, A&A, 450, L9

[17] Sidoli, L., Romano, P., Mereghetti, S., et al. 2007, A&A, 476, 1307

[18] Ubertini, P., Lebrun, F., Di Cocco, G., et al. 2003, A&A, 411, L131

[19] Vedrenne, G., Roques, J.-P., Schönfelder, V., et al. 2003, A&A, 411, L63

[20] Winkler, C., Courvoisier, T. J.-L., Di Cocco, G., et al. 2003, A&A, 411, L1

PoS(INTEGRAL 2012)151