Click or tap here to enter text.

INTEGRAL SCIENCE LEGACY ARCHIVE (ISLA) DEFINITIONS AND PRIORITIES

Prepared by	The INTEGRAL Archive Scientist and approved by the INTEGRAL Project Scientist and the INTEGRAL Mission Manager.
Document Type	RS - Requirement Document / Specification (System, Subsystem, Unit, Equipment level)
Reference	ESA-SCI-SO-RS-002
Issue/Revision	1.2 .
Date of Issue	31/01/2024
Status	Issued

Table of Contents

List of Abbreviations

1. INTRODUCTION

The INTEGRAL Legacy Science Archive (ISLA) will be the final and official archive of all data collected during INTEGRAL's science operations. The archive will be hosted at the European Space Astronomy Centre (ESAC), at the ESAC Science Data Centre (ESDC). It will provide public access for the raw data of the main instruments IBIS and SPI, as well as the X-ray and optical monitor JEM-X and OMC. Furthermore, it will provide science-ready products of all instruments as well as data from the radiation monitor (IREM) and anti-coincidence shielding (SPI-ACS). The goal is to have the data presented and formatted in such a way that any interested public user can download, access, and use the data without expert knowledge in INTEGRAL. To that end, ISLA will use "science portals", which will provide tailored science products for specific source types, allowing for maximal scientific return with minimal expertise by the user. To allow users to understand and use the data efficiently, ISLA will provide documentation for all available products.
In addition to serving data, products, and documentation, ISLA will also be linked to an online reprocessing tool, which will allow basic on-the-fly data reduction and extraction. This tool will allow (expert) users to obtain custom made data products like light-curves, spectra, and images.
ISLA will be integrated into the larger science archive environment of ESA within ESDC, and thereby benefit from possible synergies with other (high-energy) archives.

In this document we specify the requirements of a finished ISLA as well as the priorities for each task. The INTEGRAL Science Operations Centre (ISOC) is leading the development with the goal to have a basic functional archive online for public access by the end of the nominal scientific operations phase in December 2024. During the following two years of post-operations, ISLA development will continue, adding further functionality and products to ISLA, within the remaining resource and cost constraints.

2. SCIENTIFIC INSTRUMENTS

2.1. IBIS

The Imager on Board the INTEGRAL Satellite (IBIS) consists of two detector planes, the INTEGRAL Soft Gamma-Ray Imager (ISGRI) and the Pixellated Imaging Caesium Iodide Telescope (PICsIT). IBIS science covers a wide range of astrophysical topics and is the workhorse on-board INTEGRAL for light-curves, soft-gamma ray spectra and wide-field images. Providing science ready basic products from IBIS, in particular ISGRI, is therefore of the highest priority to ISLA.

2.2. SPI

The Spectrometer on INTEGRAL (SPI) provides a unique view into the universe, with its capability of performing high-resolution spectroscopy in the Gamma-ray (keV to MeV) range. SPI's science ready spectral products are of very high interest to the science community at large. Providing these products via ISLA is therefore also of high priority.

2.3. JEM-X

The Joint European X-ray Monitor (JEM-X) consists of two identical coded masks casting shadows on two independent micro-strip gas chambers, co-aligned with IBIS and SPI. JEM-X provides coverage to lower energies than the two main instruments and plays a crucial role in identifying and monitoring X-ray and Gamma-ray sources. JEM-X light-curves, spectra, and wide-filed images will be provided by ISLA with a high priority.

2.4. OMC

The Optical Monitoring Camera (OMC) provides simultaneous Johnson V-band photometry with the X-ray and Gamma-ray instruments. ISLA will make public a OMC-PI provided source catalogue of all variable OMC sources, as well as light-curves for all objects seen by OMC. Providing these data has medium priority.

2.5. IREM

The INTEGRAL Radiation Environment Monitor (IREM) measures the proton and electron fluxes around the spacecraft and is used to inform the instrument about the current radiation environment. The time series data of IREM is processed and stored by the Paul Scherrer Institut (PSI) at http://srem.psi.ch/html/irem_summaryplots.shtml, ESTEC, and ISDC. ISLA will provide a copy of the data for legacy purposes. Providing these data has low priority.

2.6. SPI-ACS

The main task of the SPI Anti-Coincidence Shield (ACS) is to reduce the background of the SPI data by filtering out cosmic rays from outside the field-of-view. At the same time, it is one of the most sensitive all-sky Gamma-ray monitors at energies >75keV. Therefore SPI-ACS high-resolution light-curves are made readily available in near-real time at the ISDC, and SPI-ACS has detected thousands of gamma-ray bursts over the course of the INTEGRAL mission (e.g., Minaev & Pozanenko, 2023, MNRAS 525, 2411). Providing the SPI-ACS time-series in ISLA has low priority.

3. REQUIRED DATA PRODUCTS

3.1. Raw data

All raw data (aka CONS data) currently available via the archives at ISDC and HEASARC will be provided in ISLA. ISLA will provide options to download single Science Windows (ScWs) or download large parts of the data in bulk. Providing these data is of very high priority. Raw data download will be integrated with the catalogue and data products described below, in the sense that it will be possible to download all or parts of the data of a specific source. Data selection should be possible interactively, for example via specific search filters (e.g., revolutions), via the light-curve visualization or other visualizations of the data (e.g., mosaics). This functionality is of medium priority.

3.2. Point Sources

3.2.1. Source Input Catalogues

All detected point sources over the mission lifetime need to be listed in a catalogue as a reference for the extraction of data products. The first catalogue to be used in ISLA is the existing INTEGRAL Reference Source Catalogue. This reference catalogue was created before the launch of INTEGRAL, and therefore also contains sources that have not yet been detected by INTEGRAL. It is made and maintained by ISDC. On this catalogue the first available data products will be based.

When new (results) catalogues, created by the respective PI teams, are made available, they will be ingested into ISLA, and the science products will be extracted on any additional sources in these catalogues. ISOC is not taking responsibility for producing these catalogues and instead is relying on input from ISDC or the PI teams for any update. However, ISOC will be able to merge different catalogues from different teams if they follow a common structure (which will be defined and agreed with between ISOC and the respective PI teams in the interface control documents).

3.2.2. IBIS and JEM-X pipeline data products

ISOC is developing a pipeline to quickly process the raw data and create products for ISLA. The pipeline is intended to be based on the MMODA scripts developed by ISDC and will eventually run in the ESA DataLabs environment provided by ESA D/SCI-SAS. The pipeline will ensure data products provenance[1] and quality control, and the DataLabs environment will provide longevity and legacy value and synergies with the scientific data processing of other ESA missions.

All products described in this Section will be displayed in an interactive plot on the ISLA website. They will also be available for download, either via the website or data queries.

[1] The exact requirements on the FITS headers and associated information for each data product will be defined in a separate document.

For each detected catalogue source ISLA will provide ISGRI and JEM-X light-curves with a time resolution as fast as one ScW in at least two energy bands (with exact energy bands defined in the interface control documents). The light-curves will cover the whole mission lifetime. ISLA will also provide light-curves with a revolution-by-revolution binning in the same energy bands. These light-curves will be produced by the ISOC-ISLA pipeline and have high priority.

In addition to light-curves of detected sources, upper limits based on an all-sky grid will be given, averaged over the whole lifetime of the mission (*cf* the current upper limits in the ESA-provided Upper Limit server HiLiGT). Additionally upper-limits on faster time resolution will be created per source and made available, where appropriate, as part of the source light-curve.

Additionally, time-resolved upper-limits on the same all-sky grid will be generated if possible. The upper limits will be calculated by ISOC and have high priority. However, it is estimated that significant development effort is necessary to produce the limits, pushing their publication to a later release of ISLA.

Mosaics and images of each observation will be provided in the same energy bands as the light-curves, based on the same ISOC-ISLA pipeline. The images have medium priority.

Spectra on a revolution-by-revolution basis will be produced with the ISOC-ISLA pipeline for ISGRI, PICsIT, and JEM-X for each detected source. The spectral binning for ISGRI and JEM-X will be defined in the interface control documents with the PI teams. Providing these spectra has high priority.

Products for PICsIT will be provided by the PI teams, the exact format is defined in the interface control document.

In addition to these standard products, products with higher time resolution (sub-ScW for light-curves, and sub-observation for spectra) will be provided for sufficiently bright sources. The good-time intervals (GTIs) for these higher resolution products will be calculated dynamically, based on the signal-to-noise (S/N) ratio in each ScW and the science case of the source type. The products will be shared through corresponding “science portals”, which cater to specific needs for specific science cases. Creating these products will be done by ISOC but requires a major effort and has medium priority.

A summary of these requirements is given in Table 1.

Table 1: Summary of priorities and effort for point source products for IBIS and JEM-X

Product	**Priority**	**Effort**	**Comment**
ScW resolved light-curve	High	Low	
Observation resolved light-curve	High	Low	
Higher time resolution light-curves	Medium	High	Exact resolution depending on S/N and science case
Observation resolved spectra	High	Low	
Time resolved spectra	Medium	High	Exact resolution depending on S/N and science case
Revolution resolved mosaics	High	Low	
Mission-average upper limits (all-sky)	High	Medium	Update to existing upper limits in HiLIGT
Faster time-resolution upper limits per point source	Medium	High	Exact resolution depending on science case
Time-resolved empty sky bins upper limit	Low	High	Exact resolution TBD (possibly yearly)

3.2.3. SPI data products

Given the funding situation and lack of necessary expert knowledge at ESAC, no detailed SPI data products for point sources are envisioned at this point.

3.2.4. OMC data products

Clean light-curves for all OMC detected sources and identified counter-parts of INTEGRAL detected high-energy sources will be provided. These data will be produced by the PI team at the Centro for Astrobiologia (CAB).
OMC images with 11x11 pixels resolution for each target for which a light curve is provided will also be provided by the PI team. Ideally, these images can be linked to a webservice providing higher resolution optical images, like Aladin or ESASky.

3.2.5. Output Results Catalogues

A result source catalogue or catalogues of detected INTEGRAL sources will be provide for all science instruments. These catalogues should provide information about detections of all INTEGRAL's X-ray and Gamma-ray instruments, marking which instrument detected which source accordingly. This catalogue will be ingested into ISLA and needs to provide information on coordinates, source classification (including an unclassified category) and, ideally, average source flux in at least two energy band per instrument. It should agree with the catalogues used for data product extraction in ISLA. Furthermore, each catalogue source should have a mission-average spectrum associated. Additionally, a similar catalogue for the OMC will be provided.
For JEM-X the DTU team is actively working on a catalogue, which will contain all significantly detected sources in one ScW of JEM-X, together with their coordinates, uncertainties on the coordinates, corresponding source in the reference source catalogue and fluxes at least two energy bands (exact bands and units are described in the interface control document). Furthermore, the catalogue will list all weak sources detected in 6-month integrated mosaics.
For the IBIS catalogue, a first version will be based on the catalogue by Krivonos et al. (2022, MNRAS 510, 4796) or by Bird et al. (2016). The IBIS team is working on new catalogues for ISGRI and PICsIT, which will become available in the future. Details about the contents of the catalogue will be defined in the interface control document.
The OMC team is working on the final catalogue of optically variable sources observed by OMC, based on the first catalogue by Alfonso-Garzón et al. (2012, A&A 548, 79A).

3.3. All Sky Maps

All maps in ISLA will be displayed interactively and will allow to query coordinates for their values. For IBIS and JEM-X they will be produced through the ISOC-ISLA pipelines, while they will be contributed by the PI teams for SPI.

3.3.1. IBIS

ISLA will contain an up-to-date mission-average exposure map of ISGRI and PICsIT as well as an all-sky mosaic in three energy bands (20-35 keV, 35-60 keV, 60-200 keV, as for the individual source light-curves). ISLA will also provide an all-sky sensitivity map, which forms the basis for the upper limits. Producing these maps has high priority.

3.3.2. JEM-X

Like IBIS, ISLA will contain an up-to-date mission-average exposure map, as well as an all-sky mosaic and sensitivity map for the JEM-X monitors. The mosaic will be created in the same energy bands as the light-curves and has high priority.

3.3.3. SPI

ISLA provide all-sky maps of various energies centred around important Gamma-ray lines, like 511 keV, Al-26, Fe-60, Na-22 and others (exact bands will be definded in the interface control document). These maps are likely provided as community contributed products.
ISLA will also contain an up-to-date exposure map of SPI.

Table 2 Summary of priorities and effort for all-sky images

	Priority	Effort	Comment
All-sky images of			
IBIS exposure map	High	Low	
ISGRI mosaic in 3 bands	High	Low	
ISGRI sensitivity map	Medium	Medium	
JEM-X exposure map	High	Low	
JEM-X mosaic in three bands	High	Low	aka significance maps
JEM-X sensitivity map	Medium	High	
SPI Exposure map	High	Medium	

3.4. Gamma-Ray Bursts

Gamma-ray bursts (GRBs) form part of the core science driver of INTEGRAL and will be represented in ISLA through their own "science portal". The GRB data will be based on the existing INTEGRAL Burst Alert System (IBAS) catalogue (see http://ibas.iasf-milano.inaf.it/), which provides name, position and uncertainty, flux, star time, T90 and some basic products like light-curves. The IBAS catalogues will be ingested into ISLA and displayed interactively, but ISOC will not create their own products for GRBs. Any information not currently in IBAS (like the fluence or redshift) would need to be provided by external partners.

In addition, ISLA, will host a copy of the SPI GRB catalogue by Biltzinger et al. (2023, A&A 675, A175) together with any products that the team can provide in agreement with ISOC and ESDC.

3.5. Contributed products

Any products beyond the capabilities of the ISOC-ISLA pipelines and described in the previous section will be supplied by the PI team or the community. ISLA will have the capabilities to include those products and make them available to the public, mainly via the "science portals".

Together with the products, the documentation and ideally extraction software/scripts should be provided. Examples for such data could be products for fast pulsars or a catalogue of X-ray bursts. These products should also include polarization products (taken via the “Compton mode” of IBIS) for the published sources and preferably the software necessary to create these products.
ISOC takes no responsibility for the data quality or delivery time of these enhanced products. Nonetheless, the products will need to be delivered at milestones to be agreed with ISOC such that proper implementation and documentation into ISLA can still be co-supervised by ISOC during post-operations at the latest.

3.6. Non-science instrument products

3.6.1. IREM

ISLA will provide a copy of the radiation monitor data available at PSI, with the same high time-resolution over the whole mission. Like light-curves from the X-ray and Gamma-ray instruments, the radiation monitor data will be displayed via an interactive plot and be available for download.

3.6.2. SPI-ACS

The total ACS light-curve of the mission life-time will be made available via ISLA, on the intrinsic time resolution of 50 ms, as well as already rebinned data with 2 s and 10 s resolution. The data will be displayed in an interactive plot and available for download in parts or in total. The ACS data will be provided by ISDC and have a low priority for ISLA.

3.6.3. IBIS/Veto

Due to the low effective area and complicated data extraction and calibration, it is currently not planned to provide detailed IBIS/Veto data in ISLA beyond what is available at ISDC.

4. OTHER PRODUCTS AND INTEGRATION

4.1. Publications

ISOC is already curating the INTEGRAL Literature Survey (ILS), which contains a list of published papers which use INTEGRAL. For each paper the used ObsIDs are provided. As a low priority goal, ISLA should provide a link between the ObsIDs and the papers and, ideally, the source products.

4.2. Integration with ESASky

ESASky (https://www.cosmos.esa.int/web/esdc/esasky-how-to) provides easy access to a wide range of data and observations from ESA and non-ESA missions. INTEGRAL is already integrated into ESASky via ISGRI RGB images. As part of ISLA, these images will be updated. INTEGRAL raw data are available for download within ESAsky, but only on a ScW basis. The goal is to have an integration between ISLA high-level products and ESAsky, so that it becomes easily possible to access the ISLA data products via ESASky.

4.3. Digital Object Identifiers

As part of the effort of creating ISLA, all INTEGRAL data will be assigned Digital Object Identifiers (DOIs), which will be linked on the ESA/ESDC websites[2]. Additionally, all high-level products will be assigned DOIs as well.

5. TIMELINE & MILESTONES

5.1. Timeline

The timeline for implementing these requirements is currently under discussion. Important fix points are the end of operations by the end of 2024 and the end of post-operations phase by the end of 2026. It is planned to have a working archive with the basis (high priority) data products available before the end of 2024.

[2] https://www.cosmos.esa.int/web/esdc/doi/

5.2. Milestones

Milestones will be defined in concordance with the stakeholders, contractors, and external teams after effort for each deliverable is estimated and agreed upon. Milestones are expected to be defined in Q1 2024.