

# Requirements for an IACHEC Source Database (ISD)

---

## The IACHEC

The IACHEC (*International Consortium for High-Energy Calibration*)<sup>1</sup> is an international consortium, gathering scientists involved in the calibration of the scientific payload of past, operational, and future high-energy astrophysics space missions. It aims to provide standards for high-energy calibration, and supervise cross-calibration among different missions. This goal is reached through Working Groups, where IACHEC scientists cooperate to define calibration standards and procedures. The scope of these Working Groups is primarily a practical one: a set of data and results (eventually published on refereed journals) will be the outcome of a coordinated and standardized analysis of references sources ("high-energy standard candles"). Past, present and future high-energy missions can use these results as a calibration reference.

Key to the IACHEC activities is the definition of the set of standard candles, and the generation and maintenance of a database of high-level scientific products (images, light curves, and/or spectra, as appropriate) used in the calibration analysis. These scientific products shall be available to the IACHEC community, as well as to the community at large, for anybody to be able to reproduce the results published in IACHEC calibration papers, and to verify the corresponding calibration status. Moreover, they shall be easily accessible to the whole community of high-energy calibrators to allow comparison between cross-calibration results obtained by calibration teams of individual missions and the IACHEC Working Groups on various calibration sources/samples.

## Scope of the IACHEC Source Database (ISD)

In order to facilitate the calibration work of IACHEC scientists, the IACHEC "Heritage Working Group"<sup>2</sup> decided in 2015 to create an **IACHEC Source Database (ISD)**. The ISD is defined as the *single repository of high-level scientific data and data analysis procedures used in IACHEC published papers*. The ISD shall be populated by an IACHEC Working Group whenever:

- An IACHEC paper is published on a refereed journal

---

<sup>1</sup> <http://web.mit.edu/iachec/index.html>

<sup>2</sup> <https://wikis.mit.edu/confluence/display/iachec/IACHEC+Heritage+Working+Group>

- Updated calibrations with respect to that used in the aforementioned paper are published, and they have a “significant impact”<sup>3</sup> on the results of the cross-calibration analysis, as verified by the Working Group and documented in a Technical Note, or in a new paper
- Paper and Technical Notes are supposed to be also ingested in the ISD, besides being available from the IACHEC web portal<sup>4</sup>

## Population and maintenance of the ISD

The ISD is designed to be a mere *repository* of high-level scientific products. The IACHEC Working Groups remain solely responsible of its scientific content, i.e. of the production, maintenance, and scientific quality of the data, analysis procedure and documents therein stored. The Working Groups are responsible for timely updating the ISD content as far as the sources they analyze are concerned. Uploading files by individual members of the IACHEC without prior endorsement of an IACHEC Working Group is deprecated.

The ultimate responsible, and point of contact for the ISD content is the IACHEC Chair, who is supported in this task by the IACHEC Working Group Chairs.

## ISD Requirements

### Access requirements (A)

**ISD-A1.** The ISD shall be publicly accessible via Internet

**ISD-A2.** The ISD shall be openly accessible for browsing and file download, without requiring any login credentials (username, password, remote key generators or others)

**ISD-A3.** The ISD shall be accessible for file upload to a selected number of registered users through login credentials (username and password)

**ISD-A4.** The ISD shall be accessible at least 95% of the time under “normal” Internet access conditions

**ISD-A5.** The ISD shall be accessible through the “most common Internet browsers”<sup>5</sup>

**ISD-A6.** The ISD shall be accessible through the “Secure Shell (SSH) Transport Layer Protocol”, allowing access to the ISD database content via batch scripts running `scp`, `wget`, or similar shell commands

### Content requirements (C)

---

<sup>3</sup> In this context, “significant impact” means that the corresponding change in the parameters describing the cross-calibration results/status is comparable or larger than their statistical uncertainties.

<sup>4</sup> <http://web.mit.edu/iachec/papers/index.html>

<sup>5</sup> At the time this document is written, the list of “most common Internet browsers” includes: Apple Safari, Google Chrome, Mozilla Firefox, Opera.

**ISD-C1.** At opening, the ISD shall include the details of the IACHEC Chair as the only user with file upload privileges

**ISD-C2.** Further users can register in the ISD database with file upload privileges by providing the following personal details: name, surname, affiliation, IACHEC Working Group membership, e-mail address, telephone number, username, and password

**ISD-C3.** The IACHEC Chair must approve each new registered user

**ISD-C4.** Users shall be allowed to change and recover their username and password, using the registered e-mail address as credential

**ISD-C5.** The ISD shall be able to store a class of files called "*Source and background scientific products*". This class includes: source+background spectra, background spectra, source+background light curves, background light curves, background-subtracted images<sup>6</sup>

**ISD-C6.** The ISD shall be able to store a class of files called "*Instrumental responses*". This class includes: spectral redistribution matrices, effective area vectors, exposure maps

**ISD-C7.** The ISD shall be able to store a class of files called "*Data analysis procedure*". This class includes any data analysis procedure applicable to the stored scientific products as in ISD-C5 and ISD-C6

**ISD-C8.** The ISD shall be able to store a class of files called "*Graphical products*". This class includes all graphical products uploaded together with the data and analysis procedures, even if the procedures do not generate them directly

**ISD-C9.** The ISD shall be able to store a class of files called "*Documents*". This class includes: a) the PDF file of the IACHEC paper published using the uploaded data; b) any Technical Notes describing updates of the published results following updated calibrations; c) a README file (in ASCII format) with the list of the uploaded files, as well as a one-line description of the nature of each file

**ISD-C10.** Users with file upload privileges shall be allowed to upload the files of all classes via a single tarfile. The tarfile shall conform to a standard directory structure specified in the document described in ISD-M2. The tarfile shall include an ASCII file with the metadata describing the uploaded tarfile: source name, source type, instruments, IACHEC Working Group, and surname of the first author of the corresponding IACHEC paper. For each upload, a new entry in the database shall be created – otherwise stated, all uploads corresponding to the same set of metadata shall be kept in the database

### *Interface requirements (I)*

**ISD-B1.** The ISD interface language shall be English

### **Querying**

**ISD-B2.** The ISD interface shall allow querying the content of the ISD by *source name*

**ISD-B3.** The ISD interface shall allow querying the content of the ISD by *source type*

---

<sup>6</sup> All high-level scientific products described in this document are assumed to be compliant with standard OGIP formats, unless specified otherwise.

**ISD-B4.** The ISD interface shall allow querying the content of the ISD by *instrument*

**ISD-B5.** The ISD interface shall allow querying the content of the ISD by *IACHEC Working Group*

**ISD-B6.** The ISD interface shall allow querying the content of the ISD by *surname of the first author of an IACHEC paper*

### **Displaying**

**ISD-B7.** The results of a query by *source name* shall be a directory structure where each directory corresponds to an observation on that source in the ISD. Each directory shall include: source+background and background scientific products, instrumental responses, data analysis procedures, graphical products, and documentation

**ISD-B8.** The results of a query by *source type* shall be a directory structure where each directory has the structure described in ISD-B7 for all the sources belonging to that source type

**ISD-B9.** The results of a query by *instrument* shall be a directory structure where each directory corresponds to an observation performed with that instrument in the ISD. Each directory shall include: source and background scientific products, instrumental responses, data analysis procedures, graphical products, and documentation

**ISD-B10.** The results of a query by *Working Group* shall be a directory structure where each directory is the structure described in ISD-B7 for each the sources published by that Working Group

**ISD-B11.** The results of a query by *surname of the first author of an IACHEC paper* shall be a list of the documents uploaded together with the paper(s) led by the IACHEC scientist carrying that surname

**ISD-B12.** If multiple file sets corresponding to the same set of metadata exist in the ISD, the user shall be offered the option of selecting which one s/he would like to display (default: the latest one).

A scheme of a possible implementation of a typical query/display sequence is shown in Fig.1

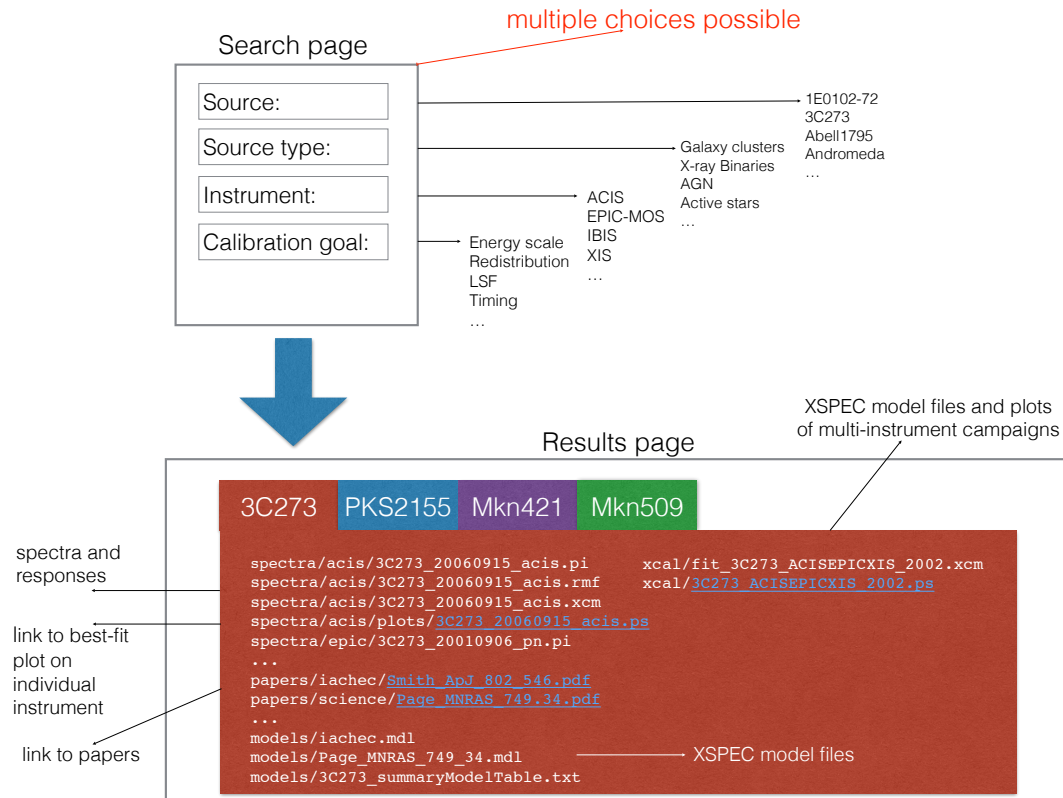


Fig.1 – Scheme of the ISD query/display sequence

## Downloading

**ISD-B12.** It shall be possible to download all files displayed in a list produced by a query as a single compressed tarfile by selecting a check-box in the ISD user interface

**ISD-B13.** It shall be possible to download all files of the same class displayed in a list produced by a query as a single compressed tarfile by selecting a check-box in the ISD user interface

**ISD-B14.** It shall be possible to select for download each individual file in a list produced by a query

## Development and maintenance

**ISD-M1.** The ISD shall be developed using solely open-source, free-of-charge software

**ISD-M2.** The ISD shall be delivered together with a technical document describing the database structure, content, and access, as well as the procedure to upload files to the ISD according to the required directory structure

## Appendix: Standard format of data package for upload

Data stored in the ISD shall be uploaded as a single tar file per source, including all the files relevant to the analysis of that source. The tar file shall be named: `iachec_isd_{$source}_{$working_group}_YYYYMMDD.tar`, where:

- `{$source}` is the name of the source
- `YYYYMMDD` is the date of the upload (in format year/month/day)
- `{$wg}` is a two-letter code indicating the IACHEC Working Group responsible for the upload, according to the following legend:

IACHEC Working Group name	Code
Calibration Uncertainties	cu
Clusters of Galaxies	cg
Contamination	cn
Coordinated Observations	co
Detector and Background	db
Heritage	he
High-resolution	hr
Non-Thermal SNR	nt
Thermal-SNR	ts
White Dwarfs and Isolated Neutron Stars	wd

In the following table the files to be included in the tarfile are listed, together with the recommended naming convention.

Product	Mandatory	Naming convention
Source+background spectrum	Y	<code>{\$source}_{\$instrument}_{\$wg}_src.pi</code>
Background spectrum	Y	<code>{\$source}_{\$instrument}_{\$wg}_bkg.pi</code>
Source+background light curve	N	<code>{\$source}_{\$instrument}_{\$wg}_src.lcu</code>
Background light curve	N	<code>{\$source}_{\$instrument}_{\$wg}_bkg.lcu</code>
Background-subtracted, exposure corrected image	N	<code>{\$source}_{\$instrument}_{\$wg}_sbs.img</code>
Redistribution matrix <sup>7</sup>	Y	<code>{\$source}_{\$instrument}_{\$wg}_src.rmf</code>
Effective area	Y	<code>{\$source}_{\$instrument}_{\$wg}_src.arf</code>
Exposure map	N	<code>{\$source}_{\$instrument}.exm</code>
Spectral analysis script per instrument	Y	<code>{\$source}_{\$instrument}_{\$wg}_src.cmd</code>
Spectral analysis script per source	N	<code>{\$source}_{\$wg}_src.cmd</code>
IACHEC paper <sup>8</sup>	Y	<code>{\$source}_{\$wg}_iachecPaper.pdf</code>

<sup>7</sup> This can be replaced by a response file (RSP) given by the combination of the redistribution matrix and the effective area. The file naming convention in this case is: `{$source}_{$instrument}_{$wg}_src.rsp`

Technical Notes	N	<code>\${source}_\${wg}_tni.pdf<sup>9</sup></code>
README (list of the full content of the tar file)	Y	<code>\${source}_\${wg}_readme.txt</code>

All strings in the file naming convention shall be *lower case*. A free naming convention may be used for graphical products, provided that their name is specified in the README. However, it is warmly encouraged to follow a self-explaining naming convention, as close as possible to the general rules described above.

---

<sup>8</sup> The paper can include also data of other sources, it goes without saying.

<sup>9</sup> With  $i=1, 2, \dots$