The Brazilian Science Data Center

An experience in the context of 'Open Universe'

Ulisses Barres de Almeida (CBPF)
on behalf of the BSDC Team*

What is the BSDC?

- A **novel data center initiative**, developed under the ICRANet-Brazil framework, with support of CBPF and the AEB.

What is the BSDC?

- A **novel data center initiative**, developed under the ICRANet-Brasil framework, with support of CBPF and the AEB.

- Built from start within the concept proposed by the **Open Universe** Initiative

What is the BSDC?

- A **novel data center initiative**, developed under the ICRANet-Brasil framework, with support of CBPF and the AEB.

- Built from start within the concept proposed by the **Open Universe** Initiative

- In the long run, it aims to respond to the **necessity of the Brazilian community** for a regional data center for Astrophysics in the Country.

What is the BSDC?

- A **novel data center initiative**, developed under the ICRANet-Brasil framework, with support of CBPF and the AEB.

- Built from start within the concept proposed by the **Open Universe** Initiative

- In the long run, it aims to respond to the **necessity of the Brazilian community** for a regional data center for Astrophysics in the Country.

- Conceived as a **virtual infrastructure**, that works as a point of integration to data and software tools from multiple providers.

What is the BSDC?

- A **novel data center initiative**, developed under the ICRANet-Brasil framework, with support of CBPF and the AEB.

- Built from start within the concept proposed by the **Open Universe** Initiative

- In the long run, it aims to respond to the **necessity of the Brazilian community** for a regional data center for Astrophysics in the Country.

- Conceived as a **virtual infrastructure**, that works as a point of integration to data and software tools from multiple providers.

A cost-effective model for developing countries

Requirements for a Science Data Center in the context of the ‘Open Universe’ initiative:

Transparency

Resurfacing Data

Broad User-base

Requirements for a Science Data Center in the context of the ‘Open Universe’ initiative:

Transparency

- Easy of access and usage
- Powerful interface : focus on data mining software
- Collaborate with the IVOA community to enhance VO experience

Resurfacing Data

Broad User-base

Requirements for a Science Data Center in the context of the ‘Open Universe’ initiative:

Transparency

- Easy of access and usage
- Powerful interface : focus on data mining software
- Collaborate with the IVOA community to enhance VO experience

Resurfacing Data

- Focus on data integration : VO as a basis
- Interoperability with other existing tools : “the ship in a fleet”
- Development of meaningful interface for specific science (CR, v, Polar.)

Broad User-base

Requirements for a Science Data Center in the context of the ‘Open Universe’ initiative:

Transparency

- Easy of access and usage
- Powerful interface : focus on data mining software
- Collaborate with the IVOA community to enhance VO experience

Resurfacing Data

- Focus on data integration : VO as a basis
- Interoperability with other existing tools : “the ship in a fleet”
- Development of meaningful interface for specific science (CR, v, Polar.)

Broad User-base

- The SDC as a data mining tool, as opposed to a repository
- Strong interaction with the community of users and providers
- A platform open for other people’s data and software tools / ideas.

Motivations for an SDC in Brazil

- **Space science for development.** Transparent access to data + flexible and meaningful tools for data handling have the potential to impact in education and formation of human resources.

- **Scientific information in the right context.** Data mining and integration is crucial for driving knowledge and potential new research avenues, specially in the era of big data.

Motivations for an SDC in Brazil

- **Space science for development.** Transparent access to data + flexible and meaningful tools for data handling have the potential to impact in education and formation of human resources.

- **Scientific information in the right context.** Data mining and integration is crucial for driving knowledge and potential new research avenues, specially in the era of big data.

Not a repository, not a set of protocols.
A data mining platform.

Motivations for an SDC in Brazil

- **Space science for development.** Transparent access to data + flexible and meaningful tools for data handling have the potential to impact in education and formation of human resources.

- **Scientific information in the right context.** Data mining and integration is crucial for driving knowledge and potential new research avenues, specially in the era of big data.

… but also

- **Collaboration with other VO services and Open Universe.** The BSDC is built for a context of cooperation, **concerned with delivering products and productivity** — a model we consider interesting for other emerging countries.

Motivations for an SDC in Brazil

- **Space science for development.** Transparent access to data + flexible and meaningful tools for data handling have the potential to impact in education and formation of human resources.

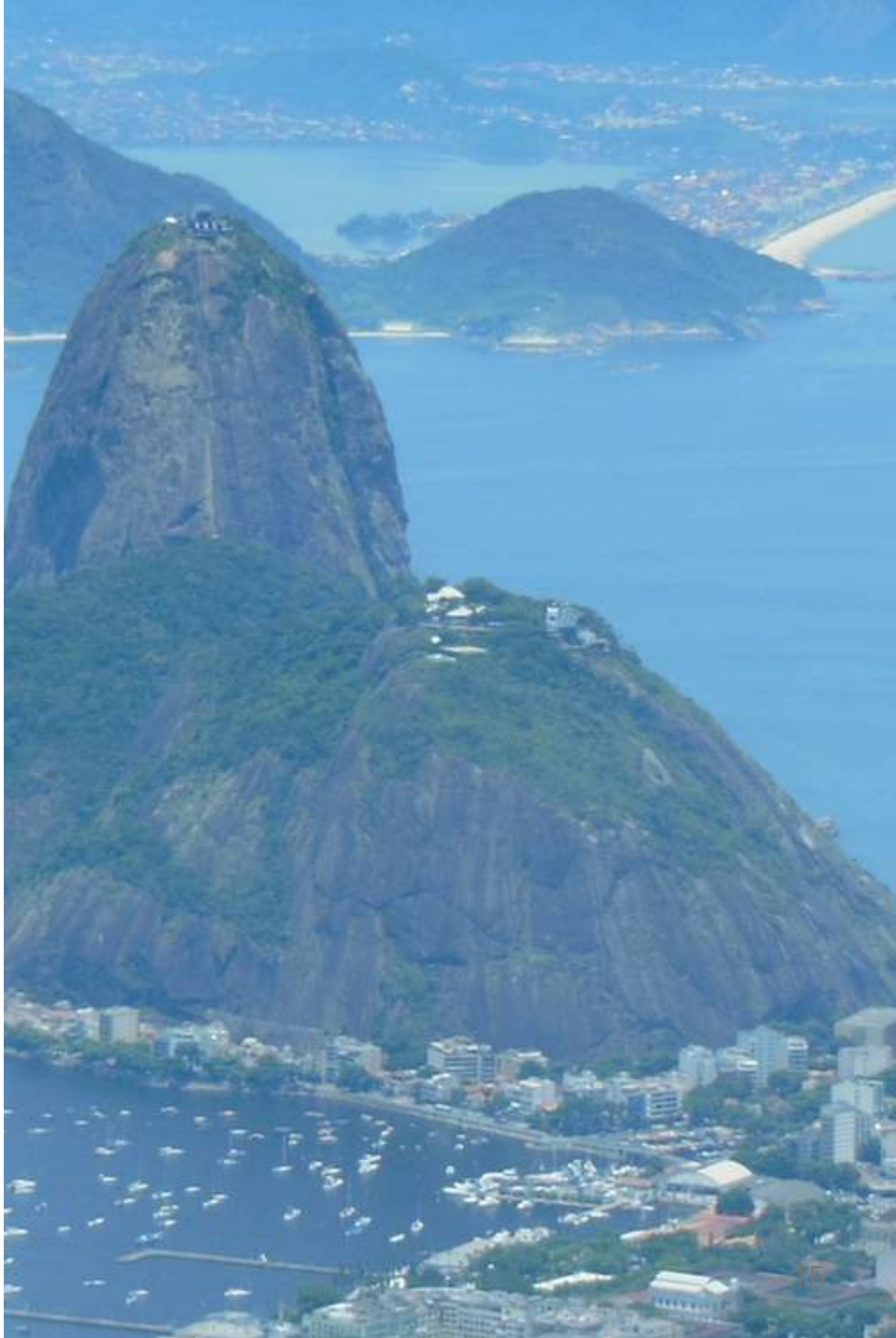
- **Scientific information in the right context.** Data mining and integration is crucial for driving knowledge and potential new research avenues, specially in the era of big data.

… but also

ASTRONOMICAL DATA AND COMPUTATION

Activities of the BSDC

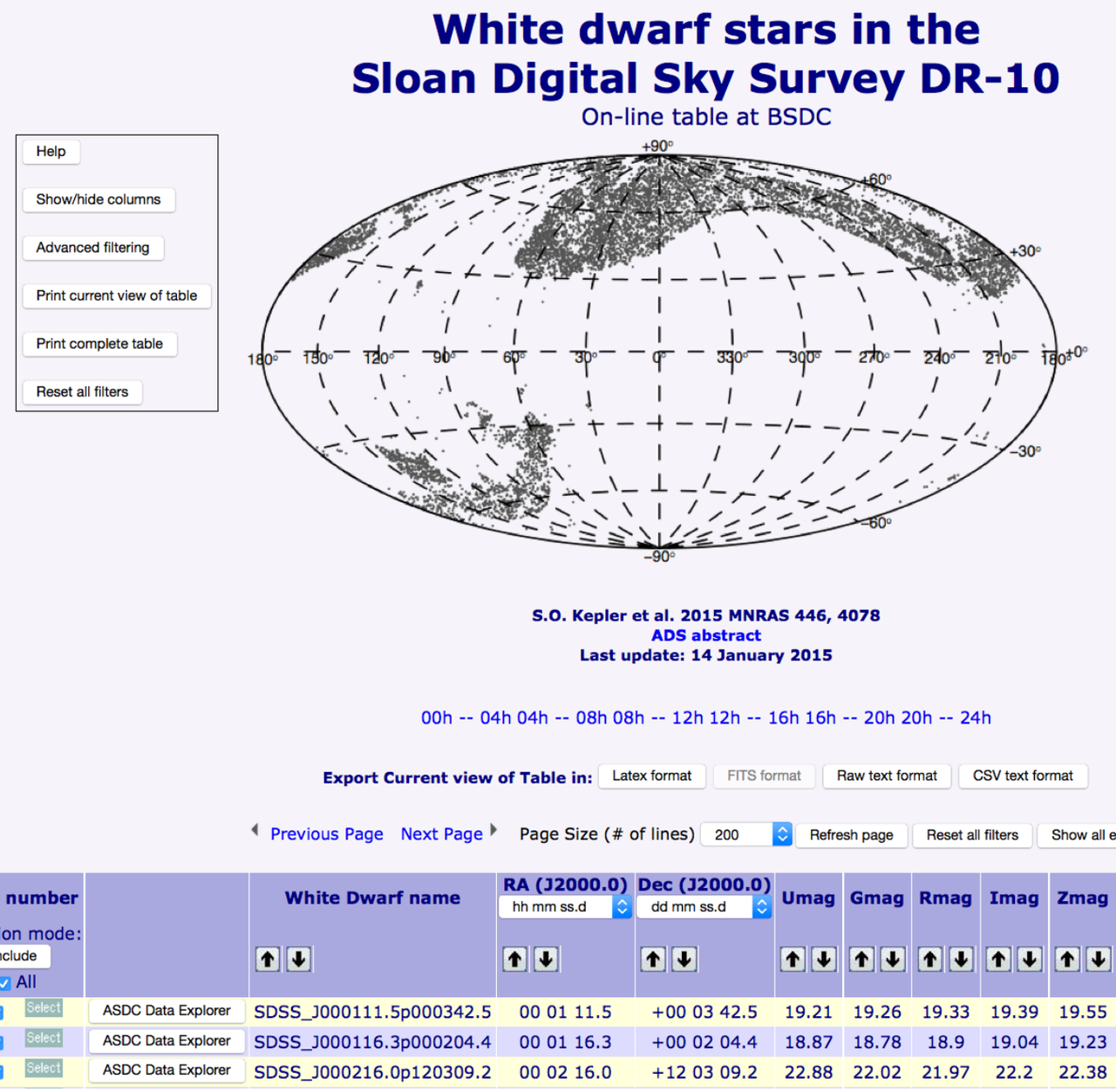
The BSDC is in development at CBPF (Rio) since 2016. First, as a database experiment in the context of scientific collaborations between Brazil and ICRANet.



Activities of the BSDC

The BSDC is in development at CBPF (Rio) since 2016. First, as a database experiment in the context of scientific collaborations between Brazil and ICRANet.

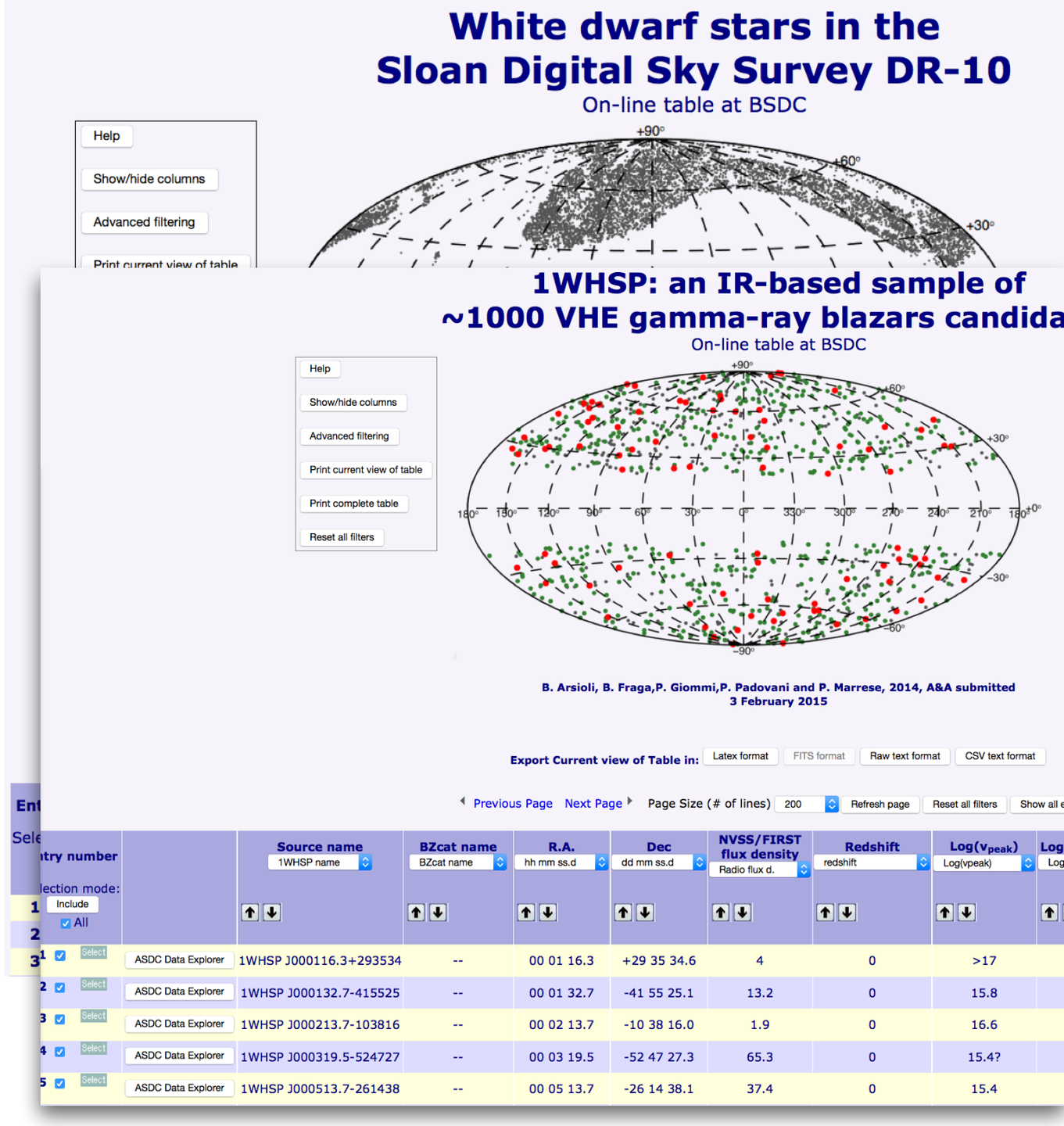
- SDSS White Dwarf catalog;



Activities of the BSDC

The BSDC is in development at CBPF (Rio) since 2016. First, as a database experiment in the context of scientific collaborations between Brazil and ICRANet.

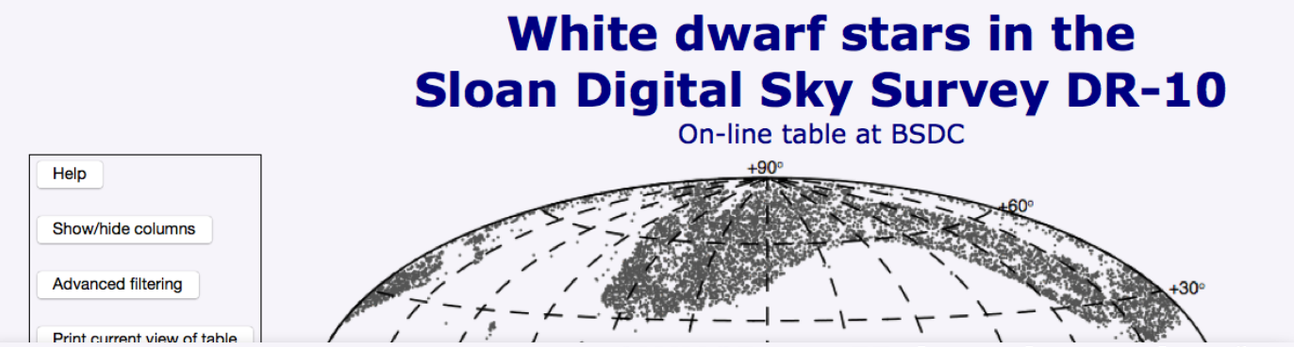
- SDSS White Dwarf catalog;
- WISE Blazar Candidates Catalogues;

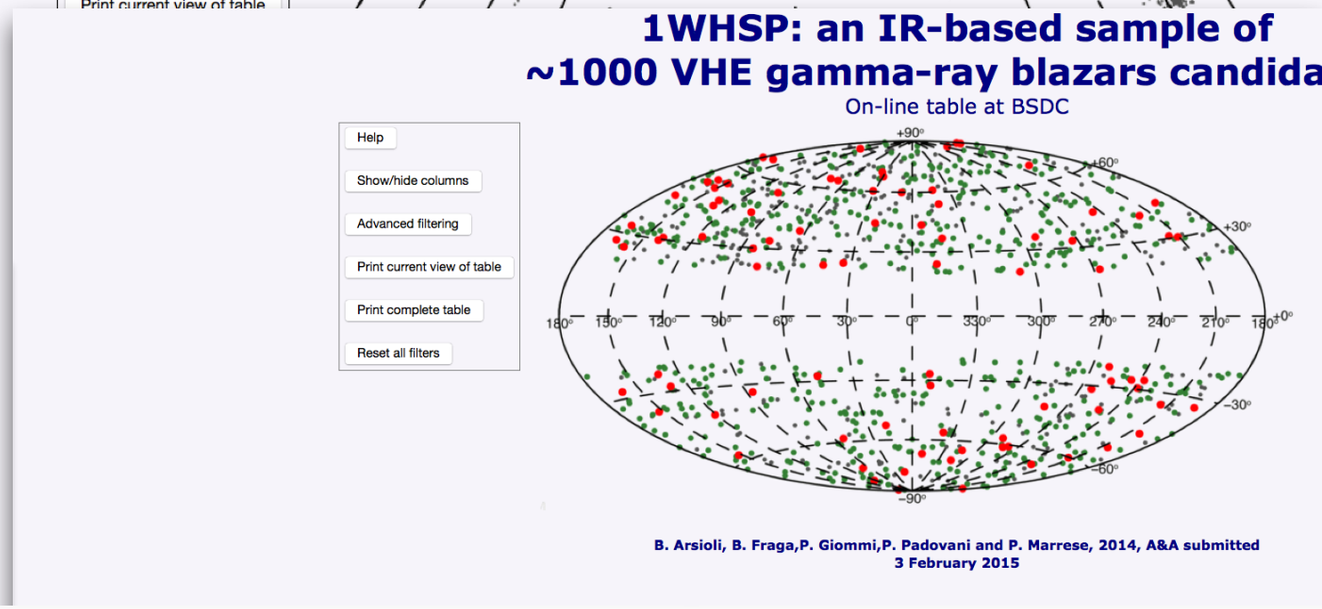


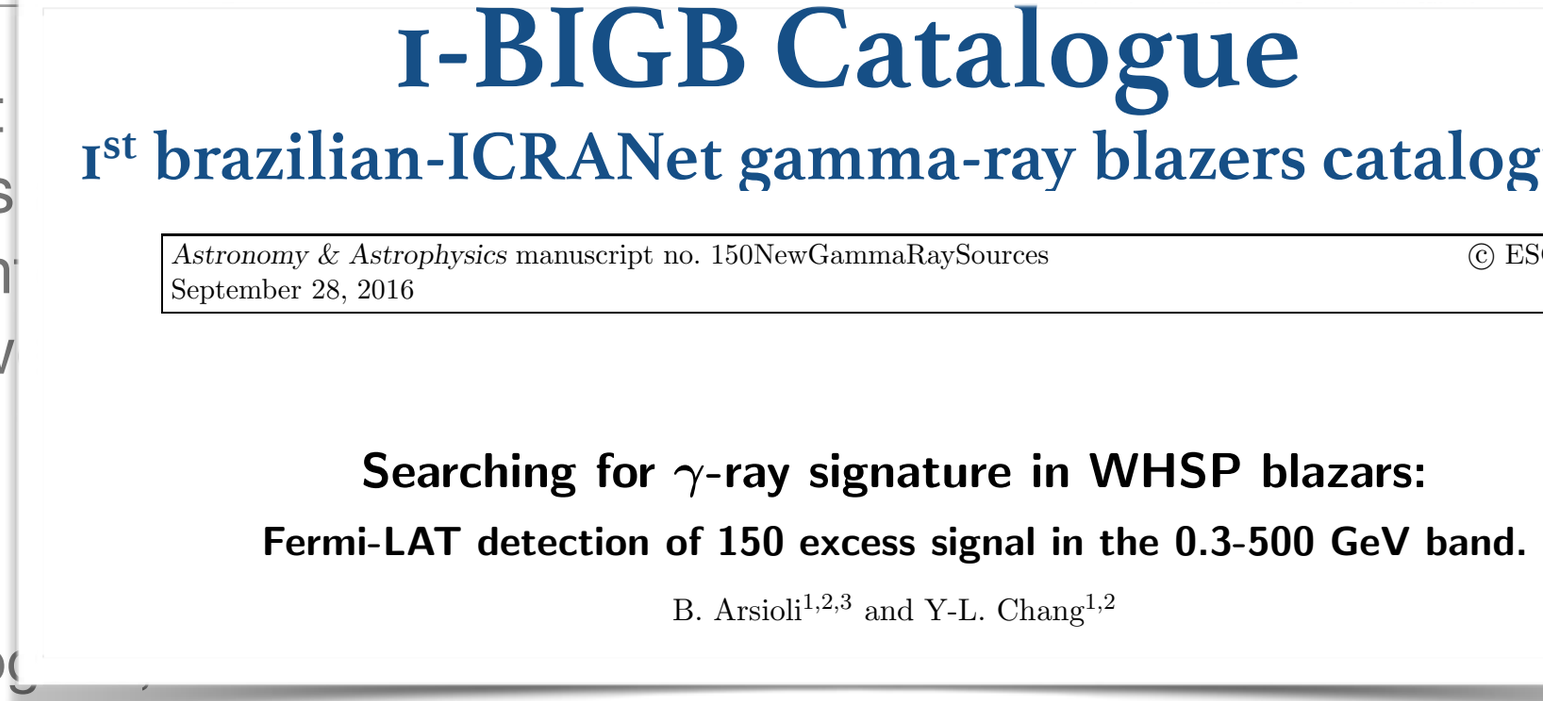
Activities of the BSDC

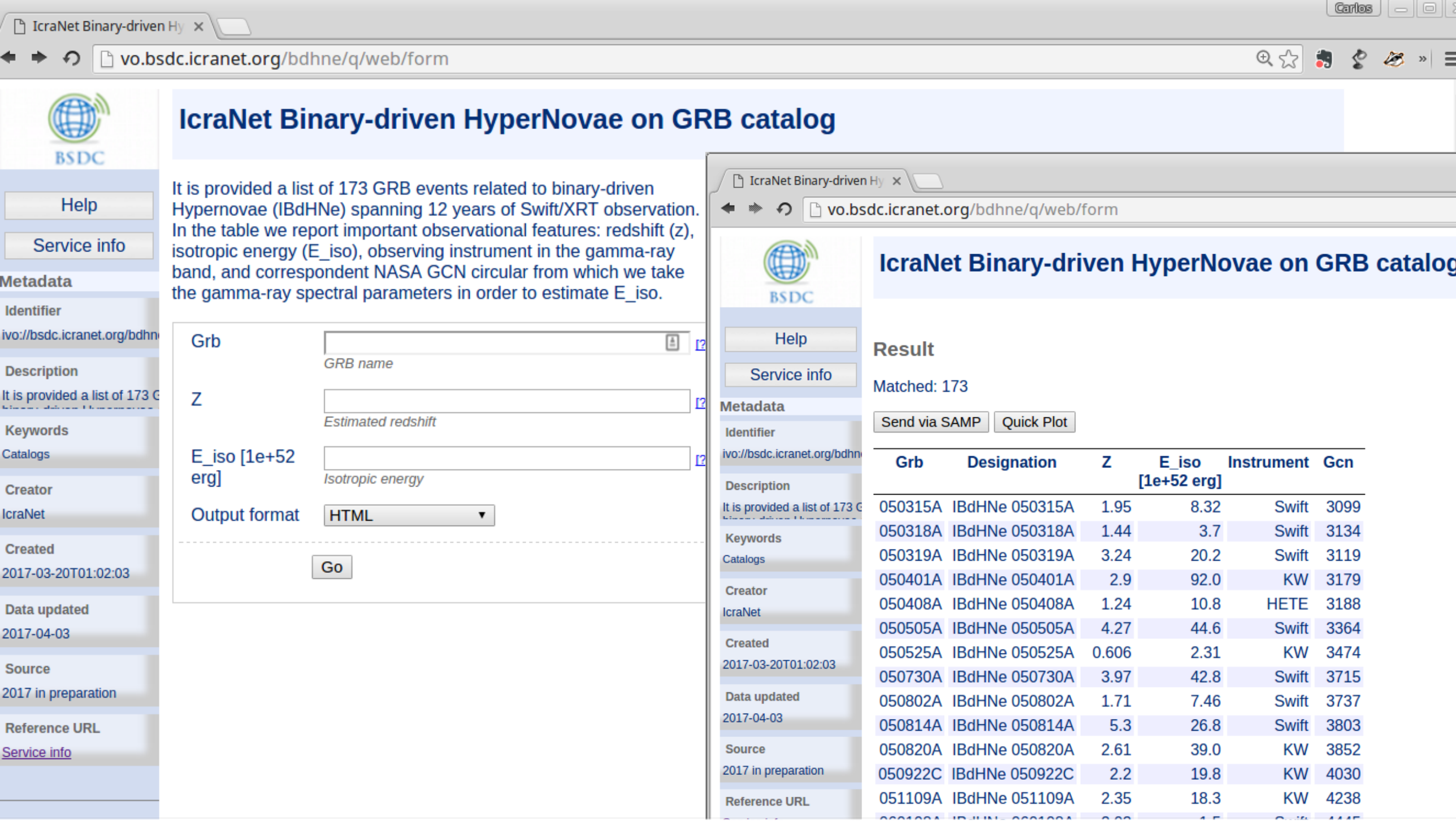
The BSDC is in development at CBPF (Rio) since 2016. First, as database experiment in the con of scientific collaborations betw Brazil and ICRANet.

- SDSS White Dwarf catalog;
- WISE Blazar Candidates Catalo
- New Catalog of Fermi-LAT BLLacs;









Fermi-LAT detection of 150 excess signal in the 0.3-500 GeV band.

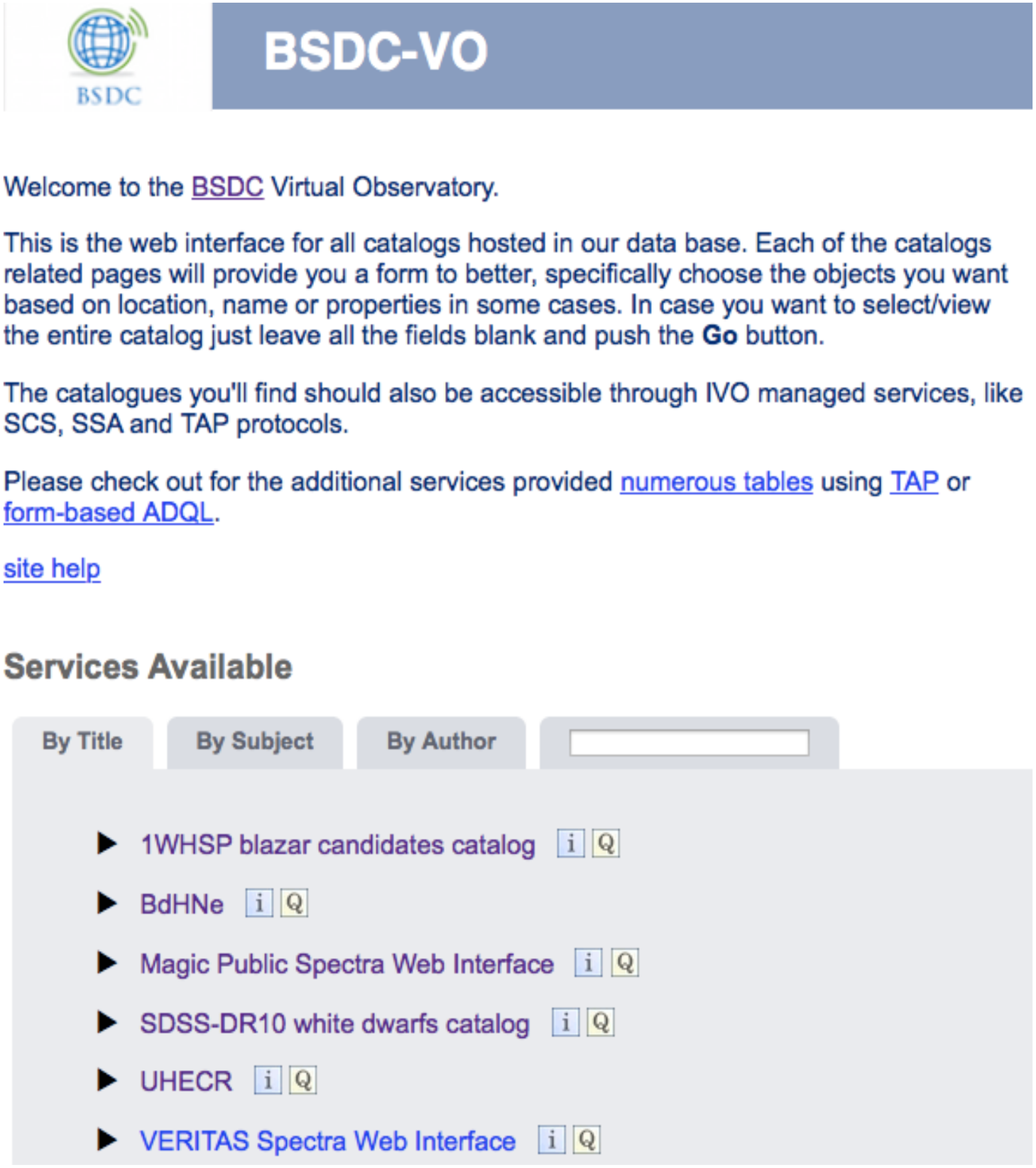
B. Arsioli[1,2,3] and Y-L. Chang[1,2]

- SDSS White Dwarf catalog;
- WISE Blazar Candidates Catalog;
- New Catalog of Fermi-LAT BLLacs;
- ICRANet GRB Catalog…

Activities of the BSDC

The BSDC is in development at CBPF (Rio) since 2016. First, as a database experiment in the context of scientific collaborations between Brazil and ICRANet.

Since early 2017, motivated by the Open Universe Initiative, we are building a full-scale online data platform.



vo.bsdc.icranet.org

Transparency

Resurfacing Data

Activities of the BSDC

Since early 2017, motivated by the Open Universe Initiative, we are building a full-scale online data platform. Activities are starting from research-critical topics.

- VHE Legacy Database, in collaboration with VERITAS (and now HESS).

➢For the first time, appropriate data formats for the field are being developed, in collaboration with the individual data providers to release their final data products in comprehensive way for the first time.

➢Data is released in VO format (for everyone). The BSDC helps with the release, then provides a platform for integration and adequate / specific tools to interact with the data online.

Transparency

Resurfacing Data

Activities of the BSDC

Since early 2017, motivated by the Open Universe Initiative, we are building a full-scale online data platform. Activities are starting from research-critical topics.

- VHE Legacy Database, in collaboration with VERITAS (and now HESS).

- Optical Polarimetry database.

➢For the first time, an effort to systematically release VO data products in the field and integrate data from multiple optical polarisation providers (usually medium-size observatories or networks) is being undertaken.

➢BSDC helps with data standards definition, data release with the VO, and the creation of the first dedicated interface for optical polarimetry (under construction)

The prototype BSDC Interface

science data center

BSDC **Brazilian Science Data Center**

Open UNiverse

CBPF

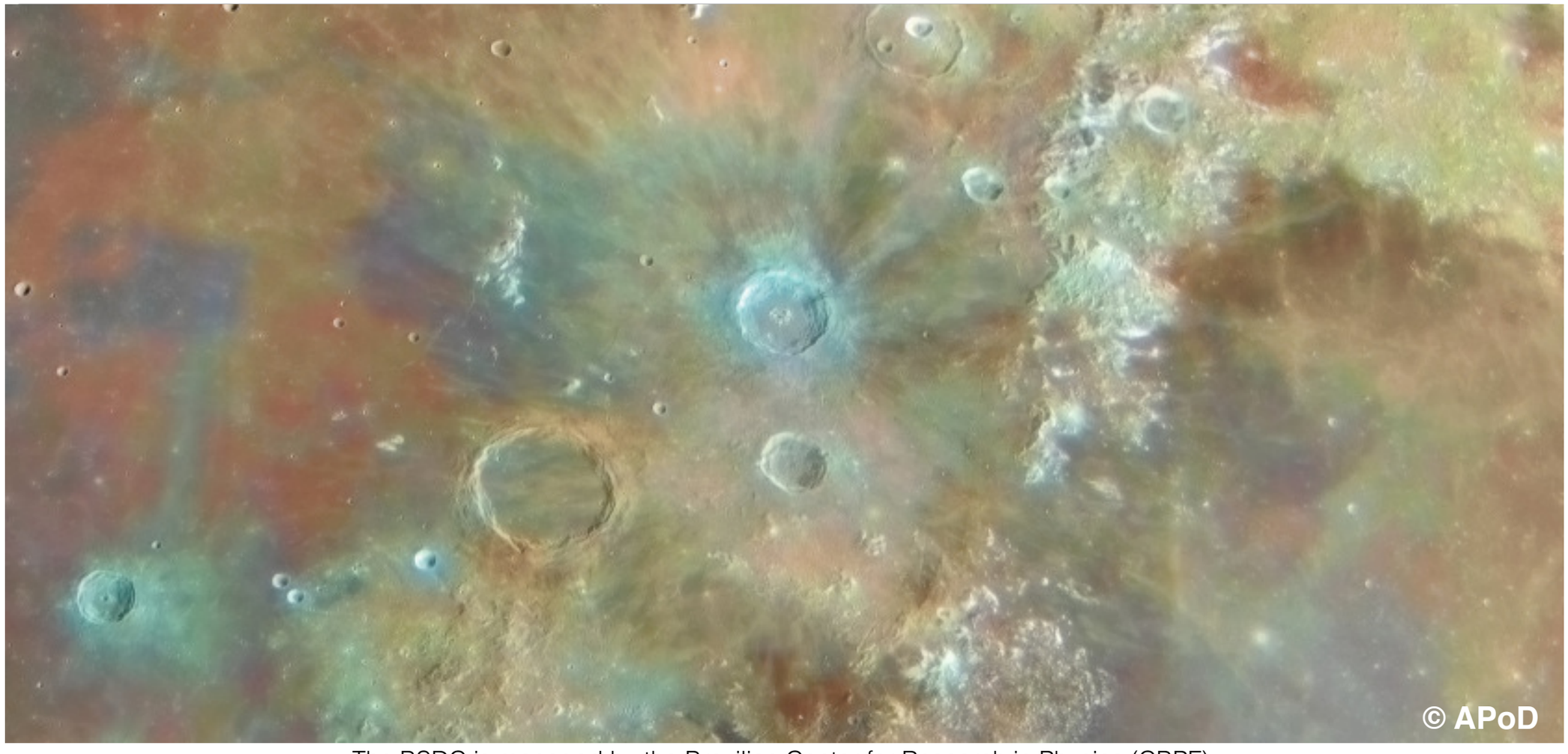
ICRANet

e.g., Mkn 421; aa.aaaa, dd.dddd; hh mm ss, dd mm ss; polarisation

Enter object's name, coordinates or keyword

- About
- Events
- Sponsors
- Contact us
- Documents

Powered by the Virtual Observatory



The BSDC is managed by the Brazilian Center for Research in Physics (CBPF).
We would appreciate that the use of BSDC for research be acknowledged in your publications.

Transparency

Resurfacing Data

Activities of the BSDC

Since early 2017, motivated by the Open Universe Initiative, we are building a full-scale online data platform. Activities are starting from research- critical topics.

- VHE Legacy Database, in collaboration with VERITAS (and now HESS).
- Optical Polarimetry database.
- Activities involving new data (Resurfacing) and increased usability (Transparency) necessarily require the development of new tools with the potential to **contribute to, and enhance the VO experience**.

Transparency

Resurfacing Data

Activities of the BSDC

Since early 2017, motivated by the Open Universe Initiative, we are building a full-scale online data platform. Activities are starting from research- critical topics.

- VHE Legacy Database, in collaboration with VERITAS (and now HESS).
- Optical Polarimetry database.
- Activities involving new data (Resurfacing) and increased usability (Transparency) necessarily require the development of new tools with the potential to **contribute to, and enhance the VO experience**.

Release of the tools and interface is expected for the end of the year.

Thank you very much!

ulisses@cbpf.br

The BSDC Team:

Bruno Arsioli, UNICAMP
Ulisses Barres, CBPF
Benno Bodmann, BSDC
Carlos Brandt, USapienza
Bernardo Fraga, CBPF
Paolo Giommi, ASI
Marko Petek, BSDC
Damiano Verzulli, ICRANet

arXiv:1702.06828v1 22 Feb 2017

The Brazilian Science Data Center (BSDC)

Ulisses Barres de Almeida and Benno Bodmann
Centro Brasileiro de Pesquisas Fsicas (CBPF), Rua Dr. Xavier Sigaud 150
Rio de Janeiro, RJ 22610-010, Brazil
ulisses@cbpf.br; benno.bodmann@gmx.de

Paolo Giommi
Agenzia Spaziale Italiana (ASDC), Via del Politecnico snc, Rome, 00133, Italy
paolo.giommi@asdc.asi.it

Carlos H. Brandt
Università di Roma "La Sapienza" (ICRANet), P.le Aldo Moro 5 Rome, 00185, Italy
carlos.brandt@asdc.asi.it

Astrophysics and Space Science are becoming increasingly characterised by what is now known as "big data", the bottlenecks for progress partly shifting from data acquisition to "data mining". Truth is that the amount and rate of data accumulation in many fields already surpasses the local capabilities for its processing and exploitation, and the efficient conversion of scientific data into knowledge is everywhere a challenge. The result is that, to a large extent, isolated data archives risk being progressively likened to "data