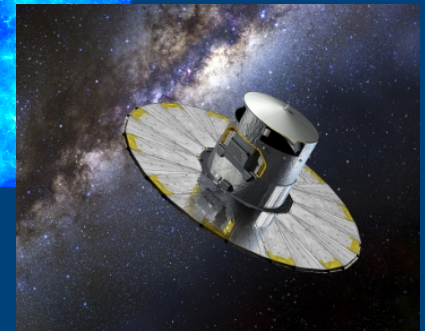
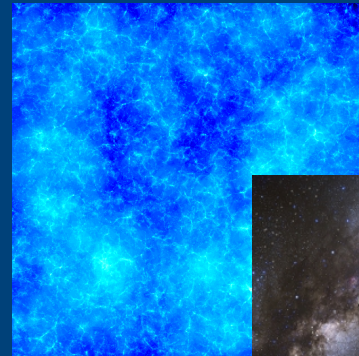




Leibniz-Institut für  
Astrophysik Potsdam

# Daiquiri - Python based framework for the publication of scientific databases

Anastasia Galkin  
Jochen Klar  
Gal Matijević  
**Kristin Riebe**  
Harry Enke



ADASS 2019

# Scientific Databases hosted @AIP

The screenshot shows the homepage of the CosmoSim website. At the top, there is a navigation bar with links: CosmoSim, Blog, Documentation, Database, Files, Query, Contact, and Login. The main heading is "CosmoSim" in a large, bold, white font. Below the heading, a paragraph states: "The CosmoSim database provides results from cosmological simulations performed within different projects: MultiDark and Bolshoi, CLUES, and Galaxies." There are three main content boxes. The first box is titled "MultiDark Bolshoi" and describes the Spanish MultiDark Consolider project. It lists simulation types: MDR1, SMDPL, MDPL, MDPL2, BigMDPL, Bolshoi, BolshoiP, and MDPL2. The second box is titled "Galaxies" and describes the MDPL2 simulation. It lists simulation types: MDPL2 Galacticus, MDPL2 SAG, and MDPL2 SAGE. The third box is titled "CLUES" and describes the CLUES project. It lists simulation types: Clues3\_LGDM and Clues3\_LGGas. To the right of these boxes is a sidebar with a "Register to CosmoSim" button, the AIP logo, and text stating "CosmoSim.org is hosted and maintained by the Leibniz-Institute for Astrophysics Potsdam (AIP)." Below this is the GAVO logo and text stating "It is a contribution to the German Astrophysical Virtual Observatory." At the bottom of the sidebar, it mentions "The MultiDark and Bolshoi simulations were run on the NASA's Pleiades supercomputer at the NASA Ames Research Center." At the bottom of the main content area, there is a "Database access" section with text: "The database can be queried by entering SQL statements directly into the Query Form or via Scripted access. If you haven't done so, please register first via the Registration Form to get your own private database where..."

CosmoSim

Blog Documentation Database Files Query Contact Login

## CosmoSim

The CosmoSim database provides results from cosmological simulations performed within different projects: MultiDark and Bolshoi, CLUES, and Galaxies.

### MultiDark Bolshoi

The Spanish MultiDark Consolider project supports efforts to identify and detect matter, including dark matter simulations of the universe.

MDR1 BigMDPL  
SMDPL Bolshoi  
MDPL BolshoiP  
MDPL2

### Galaxies

Available now for the MDPL2 simulation - galaxy catalogs contain galaxy properties from different semi-analytical codes.

MDPL2 Galacticus  
MDPL2 SAG  
MDPL2 SAGE

### CLUES

The CLUES project produces constrained simulations of the local universe, partially with gas and star formation.

Clues3\_LGDM  
Clues3\_LGGas

Please visit the linked sites for more information about the projects and about the appreciated form of acknowledgment, if the data is used in a scientific publication or proposal.

Check out the [Documentation](#) and the [Simulations](#) section for more information or the [CosmoSim blog](#) for latest news, additional materials, tutorials and much more.

### Database access

The database can be queried by entering SQL statements directly into the [Query Form](#) or via [Scripted access](#). If you haven't done so, please register first via the [Registration Form](#) to get your own private database where

[Register to CosmoSim](#)

AIP

CosmoSim.org is hosted and maintained by the Leibniz-Institute for Astrophysics Potsdam (AIP).

GAVO  
German Astrophysical Virtual Observatory

It is a contribution to the German Astrophysical Virtual Observatory.

The MultiDark and Bolshoi simulations were run on the NASA's Pleiades supercomputer at the NASA Ames Research Center.

- CosmoSim <https://www.cosmosim.org>

# Scientific Databases hosted @AIP



- CosmoSim <https://www.cosmosim.org>
- RAVE <https://www.rave-survey.org>

## RAVE DR6

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est. Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est. Lorem ipsum dolor sit amet.



## RAVE DR5 and Gaia DR2

Comparison of radial velocities from Gaia DR2 and RAVE DR5 is looking good. The

CosmoSim
Blog
Documentation
Database
Files
Query
Contact
Login

# CosmoSim



The CosmoSim database project supports efforts to detect matter, including matter simulations of the

**Multidark Bolshoi**

The Spanish Multidark project supports efforts to detect matter, including matter simulations of the

MDR1  
SMDPL  
MDPL  
MDPL2

Please visit the linked acknowledgment, if the latest news, additional

**Database access**

The database can be queried if you haven't done so,

About Documentation Database tables Query Contact

## RAVE Survey

### The Radial Velocity Experiment

RAVE is to date one of the largest spectroscopic surveys in the community. It enables stellar spectroscopic observations giving positions and proper stellar atmospheric parameters and therefore, to fully define Of all the spectroscopic survey astrometric solution catalog

- 2003-2013: 574,630 stars
- accuracy of velocity
- derived stellar parameters (WRIH)
- accuracy of distance
- accuracy of elemental
- Number of data release

Gaia@AIP
Query
Documentation
Database tables
Blog
FAQ

## Gaia@AIP Services

hosted by the Leibniz-Institute for Astrophysics Potsdam (AIP)

Welcome to Gaia@AIP Services

Launched in December 2015, Gaia@AIP is dedicated to provide the most accurate map yet of the Milky Way. By means of accurate measurements of the positions and motions of stars in the Milky Way, it will answer questions about the origin and evolution of our galaxy.

The first data release (DR1) contains three-dimensional positions and two-dimensional motions of a subset of two million stars. The second data release (DR2) will increase that number to over 1.4 billion. Gaia's measurements are as precise as planned, paving the way to a better understanding of our galaxy and its neighborhood.

The AIP hosts the Gaia data as one of the external data centers along with the main data archive maintained by ESA.

- RAVE DR6

Lorem ipsum dolor sit amet  
 tempor invidunt ut labore e  
 eos et accusam et justo duo  
 takimata sanctus est. Lorem  
 consetetur sadipscing elitr, :  
 dolore magna aliquyam erat  
 dolores et ea rebum. Stet cl  
 ipsum dolor sit amet.



In preparation for the upcoming second data release we updated **Gaia@AIP** to the new version of our Daiquiri framework. We also switched from MariaDB to PostgreSQL for the underlying database system.

We migrated your personal accounts and the metadata of all your jobs to the new system, but due to the changes to the database system we only restored the result tables querying GDR1. More information about the migration is given [here](#).

Gaia second data release (GDR2)

The second data release of the Gaia mission was published on 25 April 2018. As with the first release, the data was simultaneously released through all **Partner Data Centres** including **Gaia@AIP**. More information is available at the **ESA Gaia archive**.

Gaia first data release (GDR1)

The first release catalog includes positions and G-band magnitudes for more than a billion sources. The five-parameter astrometric solution (positions, parallaxes and proper motions) is available for over 2 million stars that are shared between Gaia and Tycho-2 and Hipparcos catalogs. More information about the release can be found [here](#).

## Getting started

Username

Password

☐ Remember Me☐ Remember Me  

If you have not created an account yet then please [sign up](#) first.

If you forgot your password and want



# Scientific Databases hosted @AIP

The image shows a collage of three scientific database websites. The top website is CosmoSim, with a dark blue background and the text 'CosmoSim' in large white letters. The middle website is RAVE Survey, with a dark blue background and the text 'RAVE Survey The Radial Velocity Experiment'. The bottom website is Gaia@AIP Services, with a dark blue background and the text 'Gaia@AIP Services hosted by the Leibniz-Institute for Astrophysics Potsdam (AIP)'. The Gaia@AIP Services page also features a 'Welcome to the MUSE-Wide survey project pages!' section and a 'Recent News' section.

- CosmoSim <https://www.cosmosim.org>
- RAVE <https://www.rave-survey.org>
- Gaia@AIP <https://gaia.aip.de>
- MuseWIDE <https://musewide.aip.de>

**RAVE DR6**

Lorem ipsum dolor sit amet tempor invidunt ut labore e eos et accusam et justo duo takimata sanctus est. Lorem consetetur sadipscing elitr, dolore magna aliquyam erat dolores et ea rebum. Stet cl ipsum dolor sit amet.

**Two-component fit**

0.35  
0.30

**Welcome to the r**

In preparation for the u also switched from Mari We migrated your perso system we only restored

**Gaia second data**

The second data release released through all Par

**Gaia first data re**

The first release catalog astrometric solution (pro and Tycho-2 and Hippar

**Getting started**

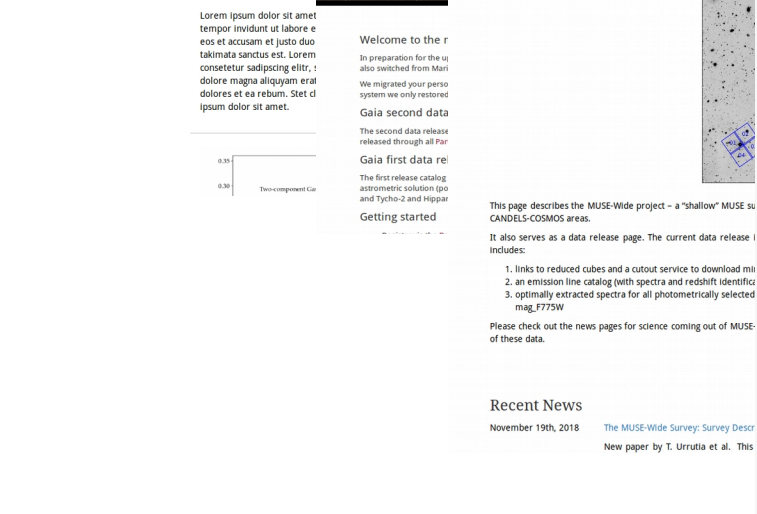
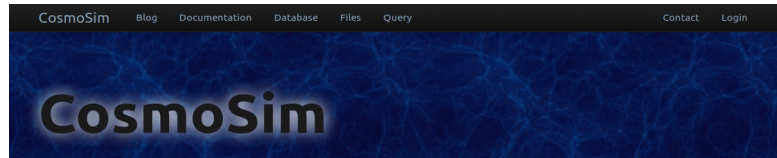
**Recent News**

November 19th, 2018

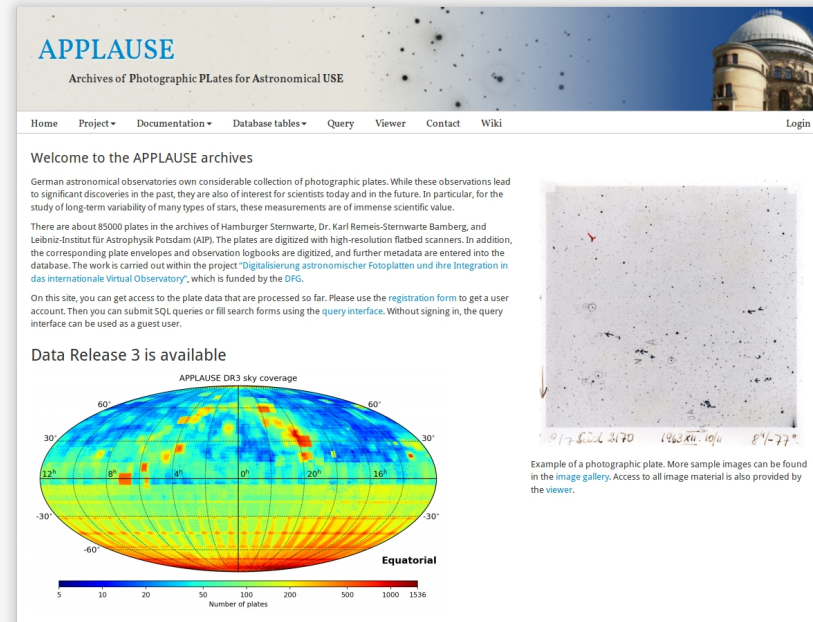
**The MUSE-Wide Survey: Survey Description and First Data Release**

New paper by T. Urrutia et al. This is the first data release paper that describes the MUSE-Wide survey and the contents of this webpage. We

# Scientific Databases hosted @AIP



- CosmoSim <https://www.cosmosim.org>
- RAVE <https://www.rave-survey.org>
- Gaia@AIP <https://gaia.aip.de>
- MuseWIDE <https://musewide.aip.de>
- APPLAUSE <https://www.plate-archive.org>



## Query interface

### DATABASE STATUS

There is no job in the queue.

You are using 15.6 kB of your quota of 100.0 GB.

### NEW QUERY

#### SQL query

[Mass function query](#)

### JOB LIST

2017-08-26-10-23-26-4438	✓
foo1	✓
2016-05-11-15-21-28-7558	✓
2016-04-08-14-41-43-3598	✓
foo	✓

[Edit jobs and groups](#)

[New Query](#)

### SQL query

Place your SQL statement directly in the text area below and submit your request using the button.

[Database browser](#)

[Function browser](#)

[Examples](#)

```
1 SELECT * FROM MDR1.FOF
2 WHERE snapnum=85
3 ORDER BY mass desc
4 LIMIT 10
```

Name of the new table (optional)

[Submit new SQL Query](#)

[Clear input window](#)

[Short queue](#)

[Long queue](#)



## Query interface

### Database status

You are using the guest user. For a personal account, please sign up [here](#).

The guest user is using 22.2 MB of its quota of 1.0 TB.

The query jobs of the guest user will be automatically archived after one week.

### New query job

#### SQL query

#### Cone search

#### Upload VOTable

### Job list

test6	⚠
test4	✓
test3	✓
test2	⚠
test	✓
UBC44_2	✓
NGC_6735	✓
NGC_6755_1	✓
NGC559_1	✓

## SQL query

Place your SQL statement directly in the text area below and submit your request using the button. You can use the dropdown menus to get information about the database, to query external services or to access example queries.

Optionally, you can specify the name of the resulting database table and/or assign a run id, to group associated queries in the job list. Please choose a queue with a suitable timeout for your query.

Database ▾
Columns ▾
Simbad ▾
VizieR ▾
Examples ▾

SCHEMAS	TABLES	COLUMNS
gdr2	gaia_source	solution_id
gdr2_contrib	sso_observation	designation
gdr1	sso_source	source_id
catalogs	vari_cepheld	random_index
TAP_SCHEMA	vari_classifier_class_definition	ref_epoch

A double click will paste the schema/table/column into the query field.

### SQL query

```

1 SELECT gmag * 0.1 AS gmag_bin, COUNT(gmag) AS number
2 FROM
3 (
4     SELECT FLOOR(phot_g_mean_mag * 10) AS gmag
5     FROM gdr1.tgas_source
6 ) AS gmag_tab
7 GROUP BY gmag;
```

### Table name

### Run id

### Query language

### Queue



## Query interface

### Database status

You are using the guest user. For a personal account, please sign up [here](#).

The guest user is using 12.1 MB of its quota of 1.0 TB.

The query jobs of the guest user will be automatically archived after one week.

### New query job

SQL query

Cone search

Upload VOTable

### Job list

2019-10-01-10-57-48-746481	▶
2019-09-30-17-05-20-566810	✓
2019-09-30-05-55-22-758059	✓
2019-09-29-09-20-38-683789	⚠
shorse	⚠
2019-09-28-12-53-05-101300	✓

## Cone search

Please specify a coordinate in right ascension (RA) and declination (DEC) and a cone radius. The query will result in objects that have coordinates in the search cone.

Optionally, you can specify the name of the resulting database table and/or assign a run id, to group associated queries in the job list. Please choose a queue with a suitable timeout for your query.

RA (degrees)

DEC (degrees)

Radius (arcsec)

Table name

Run id

Queue

Submit new cone search

## Query interface

### Database status

You are using the guest user. For a personal account, please sign up [here](#).

The guest user is using 12.1 MB of its quota of 1.0 TB.

The query jobs of the guest user will be automatically archived after one week.

### New query job

[SQL query](#)[Cone search](#)[Upload VOTable](#)

### Job list

2019-10-01-11-12-32-085469	✓
2019-10-01-10-57-48-746481	✗
2019-09-30-17-05-20-566810	✓
2019-09-30-05-55-22-758059	✓
2019-09-29-09-20-38-683789	⚠
shorse	⚠

## Upload VOTable

Please specify a local VOTable to upload and ingest as a new job into your user database. The job's table can be used in together with other tables in subsequent queries using JOIN and similar SQL commands. VOTables can be created using [TOPCAT](#) or [astropy](#).

Optionally, you can specify the name of the resulting database table and/or assign a run id, as with regular queries.

### File

No file selected.

Drag and drop file or click to open a file browser

### Table name

### Run id

# Fields in MUSE-Wide DR1

Download Exposure Map: [expmap\\_dr1\\_seconds.fits](#)

Download Field Mapping: [musewide\\_field\\_mapping.fits](#)

Please click on the **field\_name** of a field for an overview, download options and the photometric catalog for this field.

Search

Q

First

Previous

Next

Last

Reset

field_name	ra	dec	datacube_header	datacube_fits
candels-cdfs-01	53.0627015	-27.8081858	DATAcube_candels-cdfs-01_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-01_v1.0_dcbgc_effnoised.fits
candels-cdfs-02	53.0686827	-27.822625	DATAcube_candels-cdfs-02_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-02_v1.0_dcbgc_effnoised.fits
candels-cdfs-03	53.0741458	-27.8371172	DATAcube_candels-cdfs-03_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-03_v1.0_dcbgc_effnoised.fits
candels-cdfs-04	53.0819606	-27.8519903	DATAcube_candels-cdfs-04_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-04_v1.0_dcbgc_effnoised.fits
candels-cdfs-05	53.086262	-27.8666155	DATAcube_candels-cdfs-05_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-05_v1.0_dcbgc_effnoised.fits
candels-cdfs-06	53.0787868	-27.80296	DATAcube_candels-cdfs-06_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-06_v1.0_dcbgc_effnoised.fits
candels-cdfs-07	53.0848577	-27.8173972	DATAcube_candels-cdfs-07_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-07_v1.0_dcbgc_effnoised.fits
candels-cdfs-08	53.0912035	-27.8320211	DATAcube_candels-cdfs-08_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-08_v1.0_dcbgc_effnoised.fits
candels-cdfs-09	53.0968875	-27.8466478	DATAcube_candels-cdfs-09_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-09_v1.0_dcbgc_effnoised.fits
candels-cdfs-10	53.1028208	-27.86133	DATAcube_candels-cdfs-10_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-10_v1.0_dcbgc_effnoised.fits
candels-cdfs-11	53.0954458	-27.7974778	DATAcube_candels-cdfs-11_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-11_v1.0_dcbgc_effnoised.fits
candels-cdfs-12	53.1014583	-27.81201	DATAcube_candels-cdfs-12_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-12_v1.0_dcbgc_effnoised.fits
candels-cdfs-13	53.1078417	-27.8267356	DATAcube_candels-cdfs-13_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-13_v1.0_dcbgc_effnoised.fits
candels-cdfs-14	53.1133833	-27.84136	DATAcube_candels-cdfs-14_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-14_v1.0_dcbgc_effnoised.fits
candels-cdfs-15	53.1194042	-27.8559689	DATAcube_candels-cdfs-15_v1.0_dcbgc_effnoised.txt	DATAcube_candels-cdfs-15_v1.0_dcbgc_effnoised.fits

Page 1 of 3 (44 rows total)

Show 15 of rows

Page 1 of 3 (44 rows total)

Show 15 of rows

## Cut-Out Service

**RA**

Central right ascension (J2000) in deg

**a**

Radius in RA in arcsec

**DEC**

Central declination (J2000) in deg

**B**

Radius in DEC in arcsec

 **$\lambda_0$** 

Start of the wavelength in Å

 **$\lambda_1$** 

End of the wavelength range in Å

[Download cut-out](#)[Reset values](#)

It can take a moment until the download starts. Please be patient.

This datacube can also be directly downloaded using the following url:

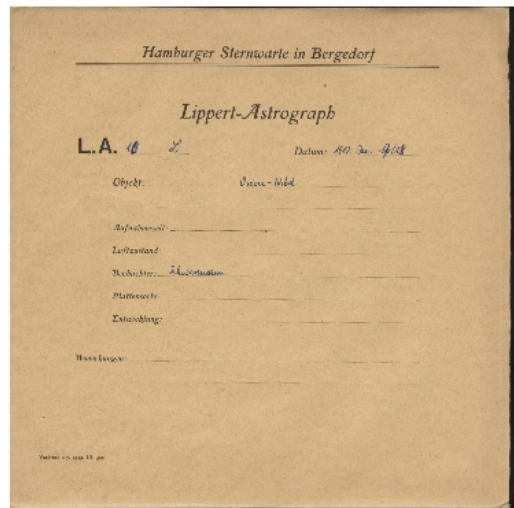
<https://musewide.alp.de/cutout/api/datacubes/?A=6&B=6&DEC=-27.81351&L0=4750&L1=9350&RA=53.0602886>





## Envelope from Lippert-Astrograph (Hamburg)

### Envelope



[Download logpage](#)

Applause ID: dr.3/envelopes/101\_8092

Archive: Lippert-Astrograph (Hamburg)

Page number:

Page order:

Image size: 1496 × 1476 px

Image created: April 21, 2010, 9:13 a.m.

File format: JPEG

File: [DR3/covers/HAM-LA/LA00010\\_cover.jpg](#)

### Digital object identifier

You can use the following DOI to cite this envelope in a publication:

[https://doi.org/10.17876/plate/dr.3/envelopes/101\\_8092](https://doi.org/10.17876/plate/dr.3/envelopes/101_8092)

### License



This envelope is published under the [Attribution 4.0 International \(CC BY 4.0\)](#).

## Navigation

[Next envelope in this archive →](#)

[Back to overview](#)

# Daiquiri – from a user perspective

## A framework for the publication of scientific databases

- Allows for highly customizable web applications
- Based on a common easily maintainable code base
- Separated into an **app** and the **daiquiri** library
- Features:
  - SQL web interface to relational databases
  - User management and user registration work flow
  - Metadata management and access control
  - First impression in-browser plotting
  - Table download
  - Archive file service
  - VOTable upload
- Employs VO protocols and standards



# Daiquiri features for the provider

- **SQL query interface** (with examples, job list, plotting) - full SQL syntax of PostgreSQL – pgSphere integration, ADQL support
- Customizable data query forms
- **Contact messages** to the support staff
- **DOI integration** and landing pages for databases and tables
- **OAI-PMH** interface for harvesting services
- File service and zip-archive creation
- Registration and log-in using OAuth2 (twitter, GitHub, Google, ORCID)
- **WordPress integration** for project presentation and documentation
- **VO protocols**: Table Access Protocol (TAP), Cone search, UWS
- File access, filtering and download for observatory archives (MUSE WIDE)
- Cut-out service for images and datacubes

open source and available on GitHub and PyPy

Web browser



Daiquiri server

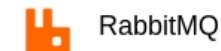


Webserver



Celery

Asynchronous tasks



Application DB



Archive DB





# Language and framework

## Python and Django

- Python: today's preferred scripting language, widely used in astronomy
- Django: full MVC framework with everything included, huge community
- Django REST framework: de-facto standard for REST interfaces in Django
- django-allauth: local and social authentication, registration work flows
- astropy: community python library for Astronomy
- Queryparser and ANTLR for translation between ADQL and PostgreSQL
- ...

## Front end: AngularJS and Bootstrap

- AngularJS 1: awesome since Daiquiri v1
- Bootstrap 3: responsive layout, mobile friendly

## Back end:

- PostgreSQL, MariaDB
- Asynchronous tasks with Celery, Redis, RabbitMQ
- Systemd, linux pipes

# Asynchronous jobs

## Celery and RabbitMQ

- Celery asynchronous task queue in Python, widely adopted
- RabbitMQ: message broker in Erlang
- redis: in-memory data structure store for task results, can also be used for caching
- systemd: new init system for Linux, make it easy to deploy daemons

## Downloading tables

### Celery and RabbitMQ and old-school unix pipes

```
mysqldump database_name table_name | some_magic > table_name.csv
```

# Try django-daiquiri!

## **as a user**

Production version of the Gaia@AIP Services: [gaia.aip.de](http://gaia.aip.de) or [plate-archive.org](http://plate-archive.org)

## **as a provider**

<https://github.com/django-daiquiri/daiquiri>

## **as docker**

<https://github.com/django-daiquiri/daiquiri-docker-compose>



Leibniz-Institut für  
Astrophysik Potsdam

# Questions?

Anastasia Galkin

[agalkin@aip.de](mailto:agalkin@aip.de)

[github.com/django-daiquiri](https://github.com/django-daiquiri)

[escience.aip.de](https://escience.aip.de)