



Leibniz-Institut für
Astrophysik Potsdam

Erfahrungen des AIP bei der Datenkuratorierung

Jochen Klar

Leibniz-Institut für Astrophysik Potsdam (AIP)

jklar@aip.de / [@jochenklar](https://twitter.com/jochenklar)

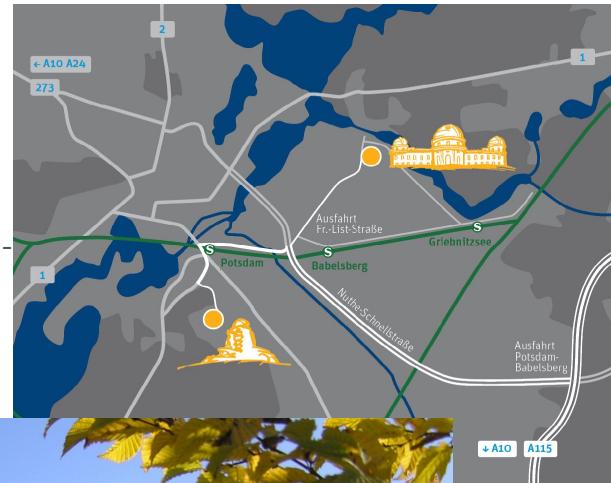
Leibniz-Institut für Astrophysik Potsdam (AIP)

Forschung

- Kosmische Magnetfelder
- Extragalaktische Astrophysik
- Entwicklung von Forschungsinfrastruktur und -

E-Science

- Infrastruktur
- Services
- Softwareentwicklung
- Support
- Community



E-Science und Forschungsdatenprojekte

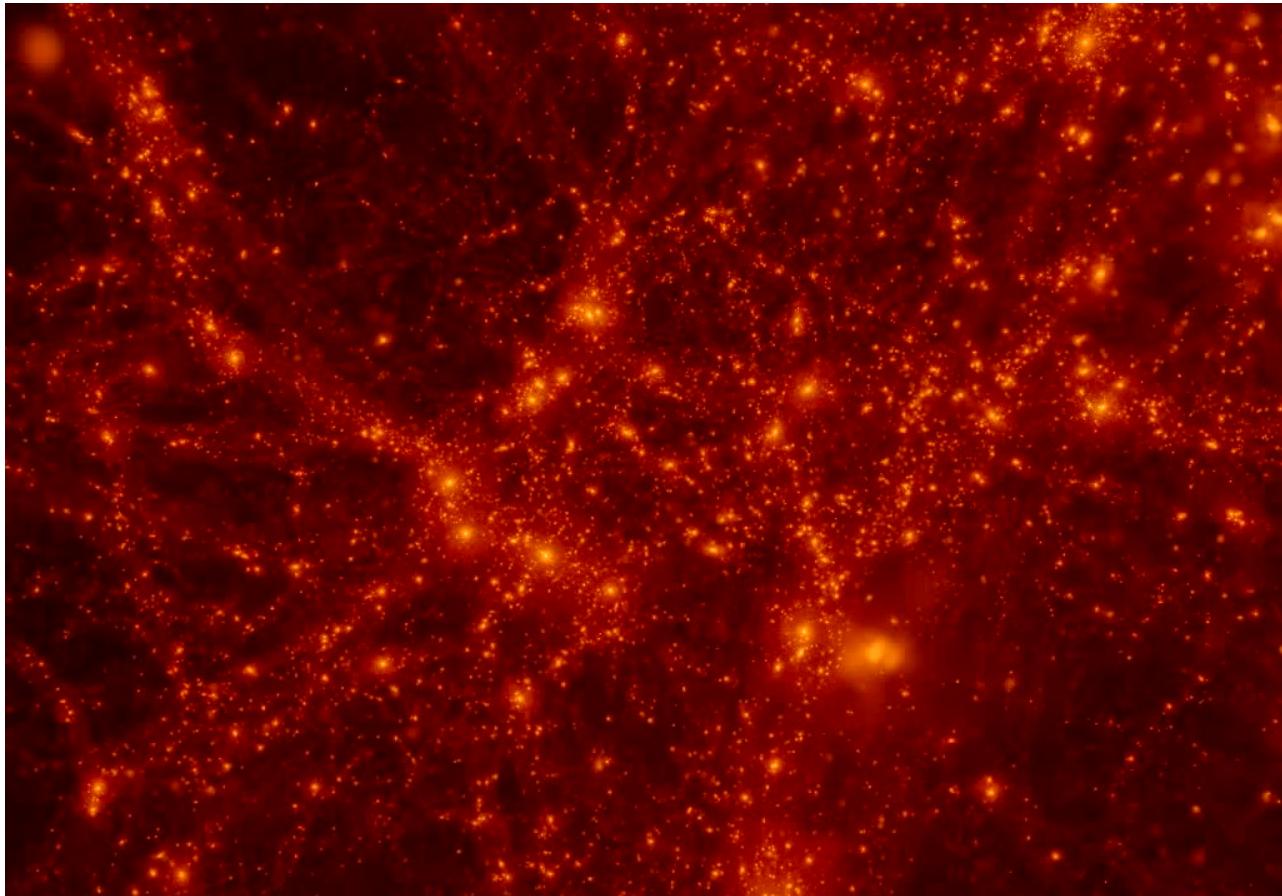
Astromie und Astrophysik

- seit 2003 [German Astrophysical Virtual Observatory \(GAVO\)](#)
- 2005 - 2009 [AstroGrid-D](#)
- seit 2017 [e-inf-astro: Interoperable e-Science-Infrastruktur](#)

interdisziplinär

- 2009-2012 [WissGrid](#)
- 2011-2013 [Rahmenbedingungen einer disziplin-übergreifenden Forschungsdaten-Infrastruktur \(Radieschen\)](#)
- 2013-2014 [Erfolgskriterien für den Aufbau und nachhaltigen Betrieb Virtueller Forschungsumgebungen](#)
- seit 2015 [Research Data Management Organiser \(RDMO\)](#)

Kosmologische Simulationen

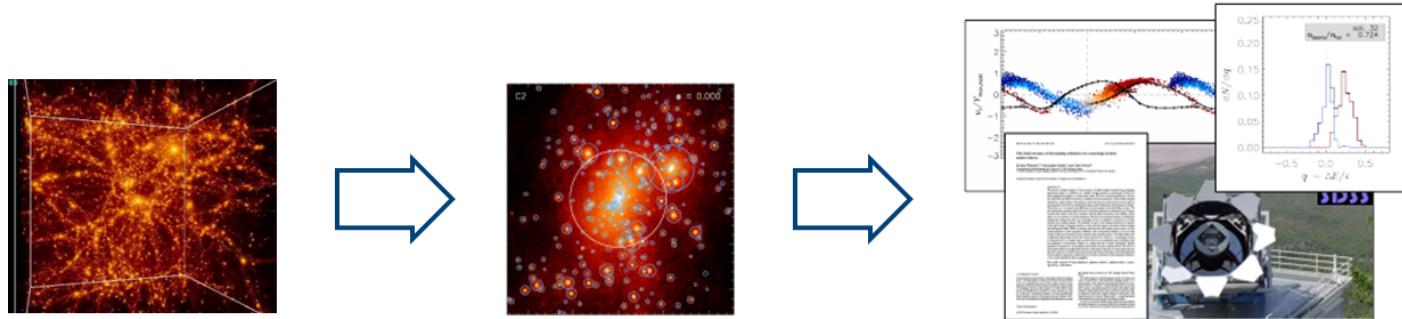


0:00



K. Riebe and the CLUES team, www.clues-project.org

EREBOS CRE



CLUES: Constrained Local UniversE Simulations

- ~ 70 Forschende aus Potsdam, Jerusalem, Madrid, Las Cruces, Santa Cruz, Lyon, Hawaii, Bogota, ...
- große Datenmengen aus Simulationen (~ 1 Pb)
- Rohdaten aus internationalen Rechenzentren (LRZ, NASA Ames, MareNostrum)
- *Post processing* zu weiteren Datenprodukten
- Wissenschaftliche Programme, Compiler, Bibliotheken
- Login per Secure Shell (SSH)

EREBOS CRE

User



Internet

erebos.aip.de
64 cores / 512 Gb

geras.aip.de
48 cores / 512 Gb

theia.aip.de
48 cores / 2 Tb

Infiniband network



Lustre file systems
~900 Tb disk space



Archive
(read-only)

Backup

Publication on
CosmoSim

CosmoSim

- Online unter <https://www.cosmosim.org>
- Abfrage von (Teil-)Datensätzen mit SQL, *share the query not the data*
- ~ 100 TB Daten, CCO

```
SELECT * FROM MDR1.FOF WHERE snapnum=8 ORDER BY mass DESC LIMIT 10
```

CosmoSim Blog Simulations Documentation Query Admin Contact My Account Logout

Query interface

DATABASE STATUS
There is no job in the queue.
You are using 15.6 kB of your quota of 100.0 GB.

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

Database browser Function browser Examples

NEW QUERY
SQL query
Mass function query

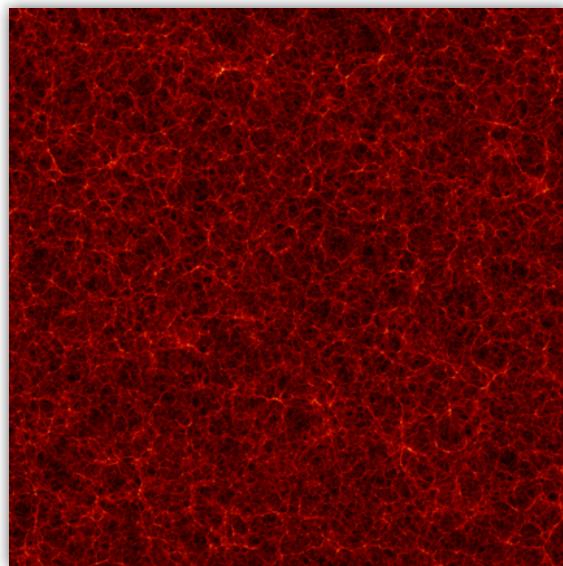
JOB LIST
2017-08-26-10-23-26-4438 ✓
foo1 ✓
2016-05-11-15-21-28-7558 ✓
2016-04-08-14-41-43-3598 ✓
foo ✓
Edit jobs and groups

```
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

CosmoSim contains data from different cosmological The MultiDark and Bolshoi simulations were run on the



Multi Unit Spectroscopic Explorer

- *First light* am 31 Januar 2014
- Besteht aus 24 Integral-field units (IFU)
- Produziert Datenkuben: 2 räumliche Dimensionen + Wellenlänge

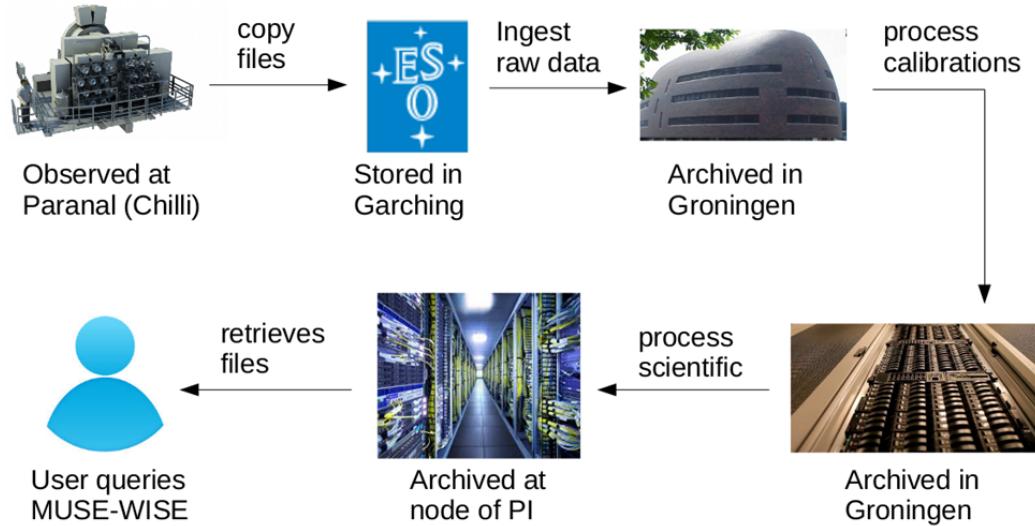


Iztok Boncina/ESO



Eric Le Roux / Service Communication / UCBL / MUSE

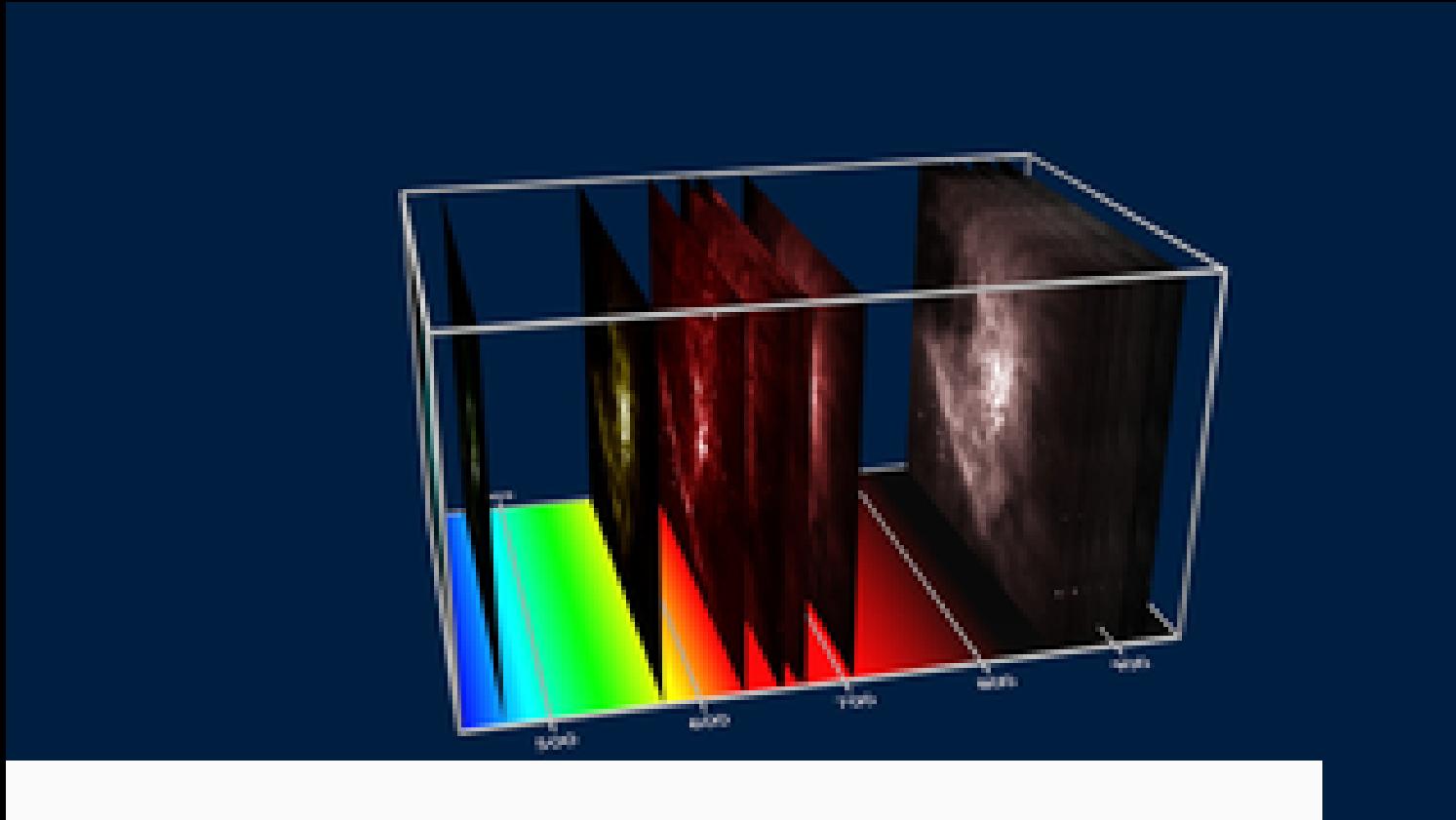
MUSE CRE



MUSE-WISE environment

- 2 Datenbankserver in Potsdam und Groningen (Metadata, Provinence)
- 7 Nodes mit Datenserver (Images, graphs, previews, logs) und Distributed Processing Units (DPU)
- Einheitlicher Software stack, Einheitliches Datenmodell

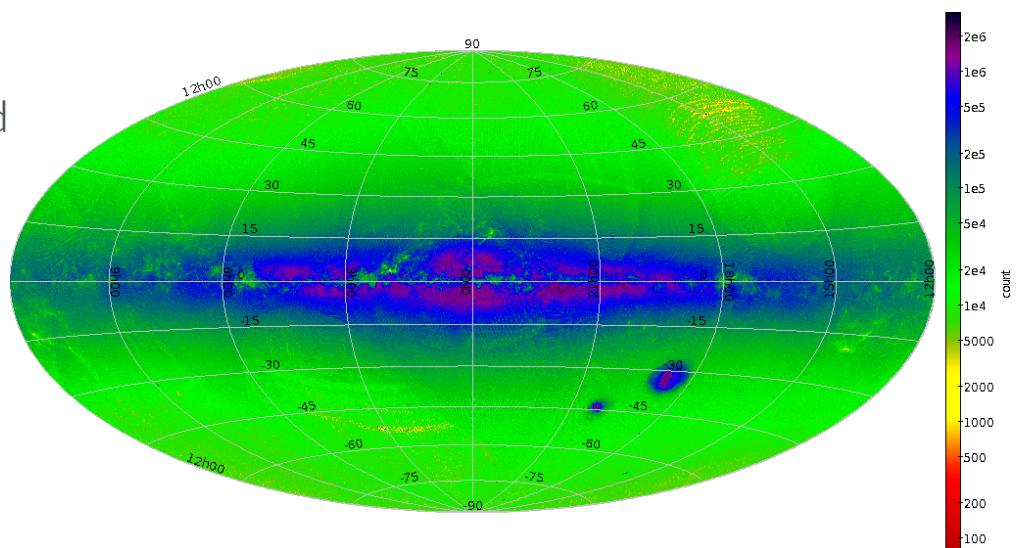
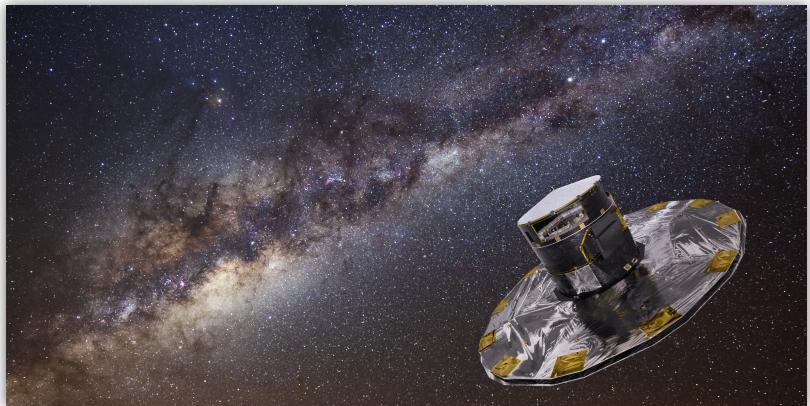
Image credit: Willem-Jan Vriend



Credit: Kristin Riebe, Peter Weilbacher, MUSE collaboration

Gaia Survey

- Satellitenmission der ESA
 - Start: 19. 12. 2013
 - Betrieb: 2014-2018
- Hochpräzise Vermessung
 - Paralaxen
 - Radialgeschwindigkeiten
 - Sterneigenschaften
- Auswertung und Publikation durch das Data Processing and Analysis Consortium (DPAC)
 - Steuerung (CU 1)
 - Auswertung (CU 2-8)
 - Archiv und Katalog (CU 9)



Gaia Archive

- Data Release 1 (GDR1) am 14. 9. 2016
- Zentrales Archiv der ESA, Mirrors in Potsdam, Heidelberg, ...

Gaia@AIP Query Documentation Database tables Blog FAQ Contact My Account Admin Logout

Query interface

DATABASE STATUS
There is no job in the queue.
You are using 1.5 GB of your quota of 100.0 GB.

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 Simbad object search Examples

NEW QUERY
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`.`gaia_source`
6 ) AS gmag_tab
7 GROUP BY gmag;
```

JOB LIST

GDR1	📁
TGAS	✓
GUMS10	📁
GOG11	📁
UNASSIGNED	
2017-08-26-09-37-53-0297	✓
2017-06-23-10-30-09-6691	✓
100	✓
2017-03-01-12-03-05-4298	✓
2016-11-18-10-54-32-8797	✓
111	∅
TGAS3	✓
TGAS2	∅

Edit jobs and groups

Name of the new table (optional)

Submit new SQL Query Clear input window Short queue Long queue

Digitalisierung von Fotoplatten



**DFG-Projekt: Digitalisierung astronomischer Fotoplatten und ihre
Integration in das internationale Virtual Observatory**

Fotoplatten aus Bamberg, Hamburg und Potsdam aus den Jahren 1909-1976

Astrophysikalisches Observatorium

Potsdam

Instrument

R 60

Aufnahmedatum

1956 Jüli 18

Expositionzeit

11° 07' - 11° 47'

Emulsion

via X 1

Entwicklung

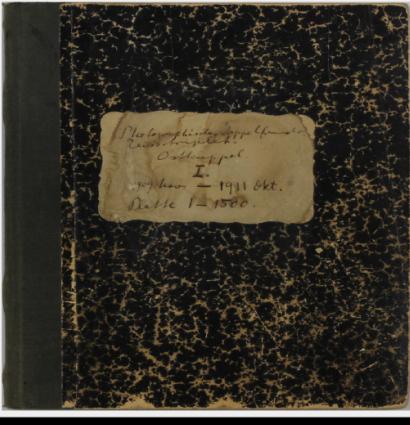
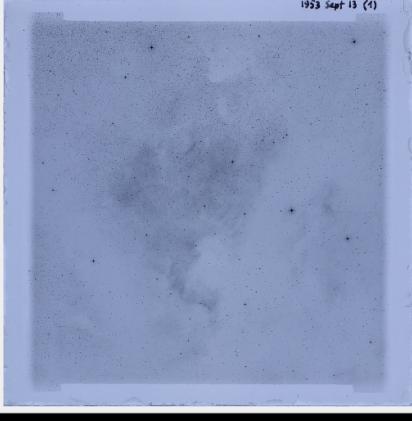
final, 10 min.

Objekt

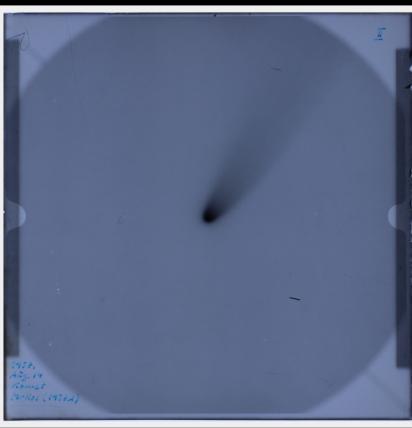
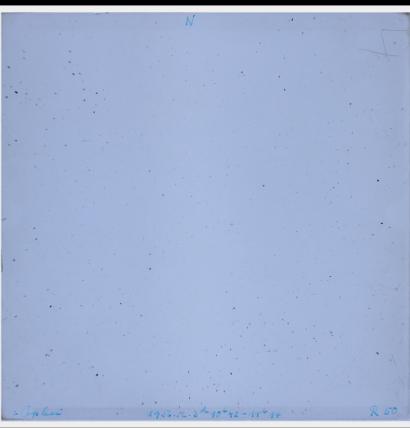
ϵ Cephei

UFG 10.256 A 3-4-26 Fe 29475

1953 Sept 13 (4)



N



Astrophysikalisches Observatorium
Potsdam

Instrument 80 cm Refraktor

Aufnahmedatum 1953 Aug 19, 20° 59' 6" - 0° 19' 6"

Expositionzeit 20 min

Emulsion Astro-Blaß

Entwicklung Radical 3:10, 6 min

Objekt Komet Honda

Besch.: Füge

UFG 10.256 A 3-4-26 Fe 29476

1953 Sept. 8 (2)

Die 1-105 sind Aufnahmen
gezogen am Wandschrank Re-
fraktor.
Klein: Platte 1, 2, 6, 11, 14 u. 16.
ca 29 als doppelt vorhanden.
Ab 22.

1912 Jan. 8.
Arctis, Span., Gestaltkreis 0°
Scl. klein 1900' ca 250
2 46 45 - 2 50 45
2 51 25 - 2 55 25+
① a 1 min
Schluss 3° 05' 00"

Blätter, Schl. klein, Park.

②
4 14 40 -- 4 21 04
und 4 22 10 -- ca 4 31 00 abgela.
-15° im Wandschrank
-17° Wandschrank

1912 Jan. 10

③ Arctis, Scl. klein, Span., Achs.
von ca 140° bis ca 210°

in der Mitte des Plättl angefangen

④ neue Blätter, Breitkreis 4°

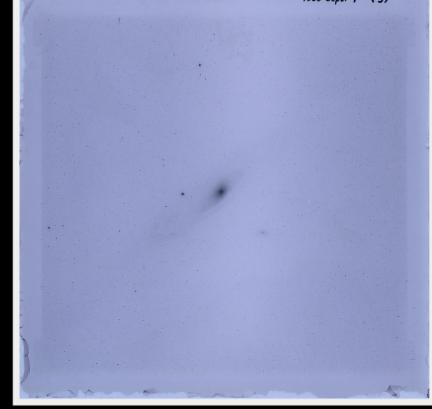
von ca 210° - 260°

Achse fehlend abgelaufen

2 26 -- 2 30 45 Jan. 10

Letzte Spur des zweiten Kreises lang in der Plättl

1953 Sept. 7 (3)



1927 Sept. 3. Waffelova 60. 1927.

4652.

Im Mittelpunkt des Kreises die Nova
Zwei Aufnahmen von gleicher Expositionzeit
stellen eine direkte Vergleichung von Beobachtungs-
aufnahmen der Nova.
gute Parallelisierung. Eindeutig in einer

14 / 19

APPLAUSE Datenbank

- Online unter <https://www.plate-archive.org>
- 51.517 Scans von 42.789 Fotoplatten
- 58.115 digitalisierte Plattenhüllen
- Logbuch-Einträge aus 149 Logbüchern
- ~ 30 TB an Daten, CCO, DOI, Europeana

APPLAUSE
Archives of Photographic PLates for Astronomical USE

Project Archives Documentation Query Contact Wiki Admin Update Profile Change Password Logout

Query interface

DATABASE STATUS
There are 40 jobs in the queue.
You are using 248.8 kB of your quota of 1.5 GB.

SQL query

Plate cone search DR2 light curve by star ID

JOB LIST

M31
2017-01-31-19-10-46-4208 ✓
2017-01-31-19-08-58-6741 ✓
2017-01-30-19-08-37-8611 ✓
2017-01-30-19-05-46-9511 ✓
2017-01-30-19-05-09-5105 ✓
count_distinct
count
mail
1943-1945
archives

New Query

SQL query

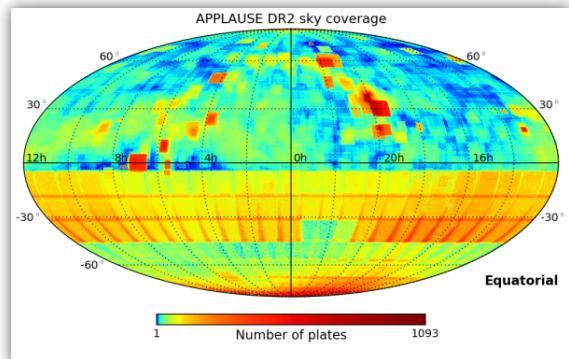
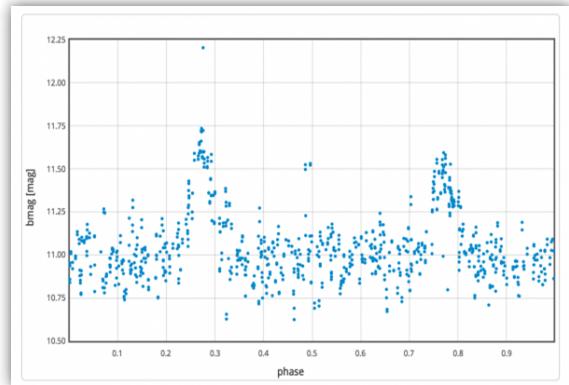
Place your SQL statement directly in the text area below and submit your request using the button.

Database browser Function browser CDS search Column search Examples

```
1 SELECT MOD(jd_mid-1.0,1.39156629)/1.39156629 AS phase, bmag, vmag
2 FROM APPLAUSE_DR2.lightcurve
3 WHERE tycho2_id='2673-02051-1'
4 AND ut_start NOT LIKE '%00:00:00'
```

Name of the new table (optional)

Submit new SQL Query Clear input window Short queue Long queue



Datenpublikationen am AIP

- RAVE database query interface for DR3 ([Siebert et al., 2011](#))
- MultiDark Database ([Riebe et al., 2011](#))
- RAVE (2013), [CosmoSim](#) (2014), [APPLAUSE](#) (2015), [Gaia@AIP](#) (2016)

The screenshot shows the RAVE website. At the top, there's a blue header bar with the RAVE logo, navigation links (Home, About Rave, Documentation, Query, Downloads, Contact, Intranet, My Account, Admin, Logout), and the Leibniz-Institut für Astrophysik (AIP) logo. Below the header, the main content area has a dark blue background with the title "RAVE – the Radial Velocity Experiment". It features a search bar, a "RAVE on facebook" button, and a sidebar with links like "About Rave", "Introduction to RAVE", "Project description", "RAVE Publications", "Standard Acknowledgment", and "Further Publications using RAVE data". The central part of the page displays a list of data releases from 2003-2013, including details like the number of spectra (574,630), velocity determination (~1.5 km/s), stellar parameters (Teff, logg, metallicity), and abundance ratios.

The screenshot shows the CosmoSim website. The header includes links for CosmoSim, Blog, Simulations, Documentation, Query, Admin, Contact, My Account, and Logout. The main title "CosmoSim" is prominently displayed. Below it, a text block describes the database's purpose: "The CosmoSim database provides results from cosmological simulations performed within different projects: the MultiDark project, the BolshoiP project, and the CLUES project." There are three project cards: "MULTIDARK" (Multi-wavelength Approach for Dark Matter Detection), "BolshoiP" (Cosmological Simulations), and "CLUES" (Constrained Local Universe Simulations). A "Register to CosmoSim" button is located on the right. The footer contains the AIP logo and a copyright notice: "Copyright © Leibniz-Institut für Astrophysik Potsdam e.V.".

The screenshot shows the APPLAUSE website. The header includes links for Project, Documentation, Query, Contact, Wiki, and Login. The main title "APPLAUSE" is at the top, followed by the subtitle "Archives of Photographic PLates for Astronomical USE". Below this, a large image of a telescope dome is shown. The main content area is titled "Welcome to the APPLAUSE archives". It discusses the collection of photographic plates from German astronomical observatories and the work done by the "Digitalisierung astronomischer Fotoplatten und ihre Integration in das internationale Virtual Observatory" project, funded by the DFG. A thumbnail image of a photographic plate is displayed.

The screenshot shows the Gaia database interface. The header includes links for Gaia@AIP, Documentation, Query, Contact, My Account, Admin, and Logout. The main title "Gaia database interface" is displayed against a background of a star-filled galaxy. Below it, a sub-header states "hosted by the Leibniz-Institut für Astrophysik Potsdam (AIP)". A text block explains the Gaia mission's goal: "The Gaia space observatory will collect unprecedented positional and radial velocity measurements of about one billion stars in our Galaxy and throughout the Local Group. This catalog will be freely available for the scientific community. The Gaia archive will comprise of a central archive hosted and maintained by ESAC and several specialized community archives." Another text block mentions the "GUMSS-10 simulation" and its role in the archive. On the right side, there's an image of the Gaia satellite in space.

Virtual Observatory



- Standards (Formate, Metadaten, Protokolle)
- Interoperable Tools und Anwendungen
- Verteilte Datenzentren, *VO registry*

Datenkuratorierung am AIP

Datenmanagement

- Virtuelle Forschungsumgebungen zum kollaborativen Arbeiten
 - EREBOS CRE
 - MUSE CRE
 - GREGOR CRE
- Unterstützung von Surveys (z.B. RAVE) oder eigenen Observatorien (z.B. Stella)
- Nachträgliche Digitalisierung von Fotoplatten (Digital Heritage)

Datenpublikation

- Zugriff über Webseiten auf SQL Datenbanken
- Nutzung von DOI für alle neueren Datenpublikationen
- Standardisierung und Formate durch das Virtual Observatory



Leibniz-Institut für
Astrophysik Potsdam

Danke für Ihre Aufmerksamkeit!

Jochen Klar

jklar@aip.de / [@jochenklar](https://twitter.com/jochenklar)

escience.aip.de

escience.aip.de/vis

www.aip.de