DOI of ESA science datasets: status and ways forward

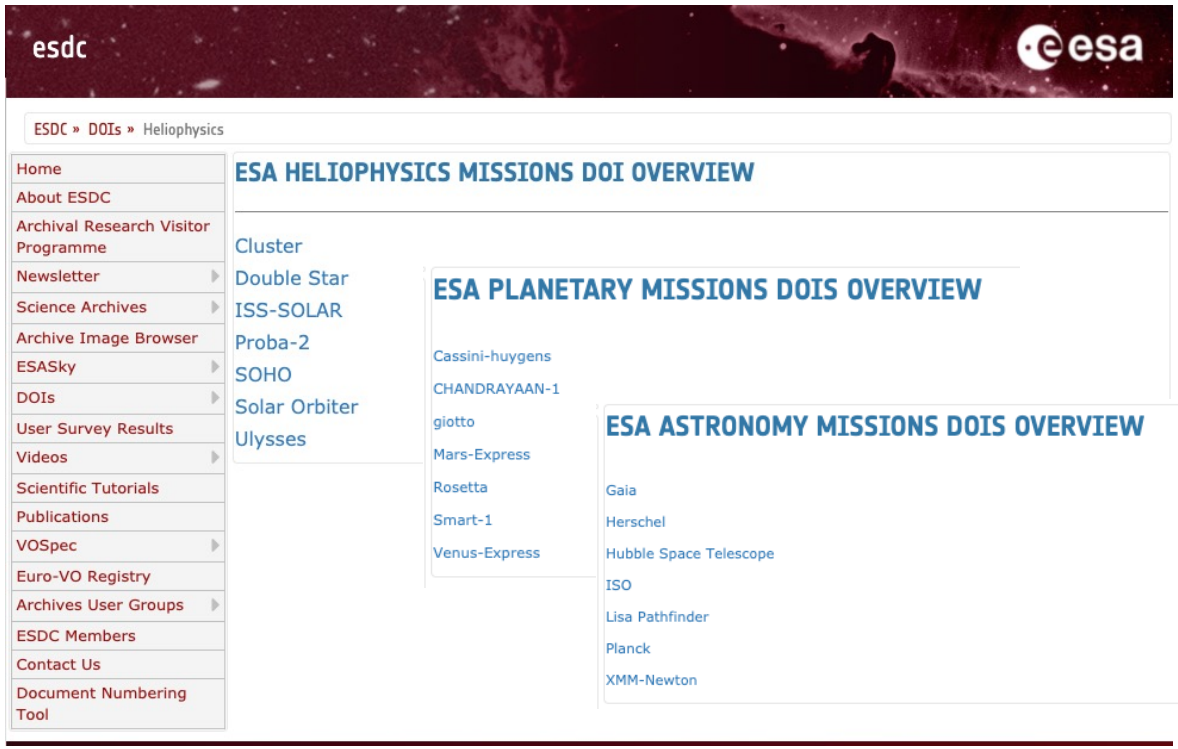
Arnaud Masson
Technical Manager, scientific support to ESA archives (Heliophysics, Planetary and Astronomy)
ESAC Science Data Centre, ESA, Madrid

IVOA interop
19 October 2022

Outline

1. 2021 activites
2. Recent progress in 2022
3. Way forward

25,000+ DOI minted for ESA science directorate missions' datasets



https://cosmos.esa.int/web/esdc/doi/

A dataset provided by the European Space Agency

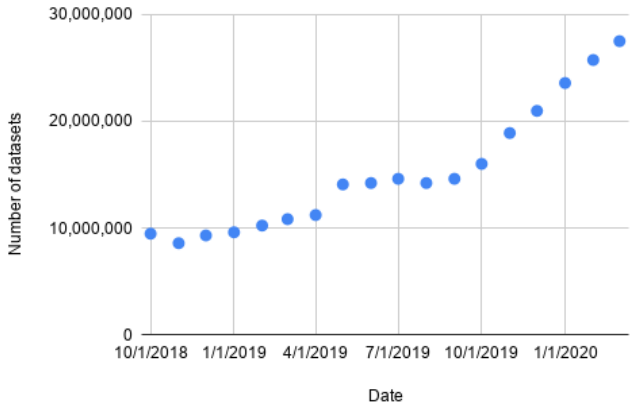
Name	MAG, MAGnetometer
Mission	Solar Orbiter
URL	http://soar.esac.esa.int/
DOI	10.5270/esa-
Abstract	The Solar Orbiter magnetometer is a conventional dual fluxgate design. Two sensors are accommodated on the spacecraft boom: MAG-IBS and MAG-OBS. A dual sensor configuration provides redundancy and, since they are at different distances from the spacecraft body (approx. 1 m for IBS and 3 m for OBS), also allows gradiometer magnetometer characterisation of spacecraft signals in flight. The instrument noise floor has been successfully tested at 10 pT Hz−1/2 at 1 Hz.
Description	Calibrated magnetic field data in RTN coordinates and in the spacecraft reference frame. Field vector components are given units of nanoteslas and in RTN coordinates, where R is the Sun-spacecraft axis, T is the cross product of the solar rotation a and R, and N is the cross product of R and T. During cruise phase, MAG is operating at 1 vector/s cadence. After the cruise phase, MAG is expected to operate continuously at 16 vectors/s cadence (normal mode) except during 1 hour per day at 128 Hz or during 2 hours at 64 Hz (Burst mode). Alternatively, burst modes will be triggered in coordination with other in-situ instruments' burst modes.
Publication	Horbury, A., et al., The Solar Orbiter magnetometer, Astron. Astrophys., 2020; DOI: doi.org/10.1051/0004-6361/201937257
Temporal Coverage	2020-05-01 - present
Mission Description	Solar Orbiter is a mission of international collaboration between ESA and NASA. It explores the Sun and the heliosphere fron close up and out of the ecliptic plane. Launched on 10 February 2020, it aims to address the overarching science question: how does the Sun create and control the Heliosphere – and why does solar activity change with time? To answer it, the Solar Orbiter spacecraft is cruising to a unique orbit around the Sun, eventually reaching a minimum perihelion of 0.28 AU, and performing measurements out of the ecliptic plane: reaching 18° heliographic latitude during its nominal mission phase, and above 30° during its extended mission phase. It carries six remote sensing instruments to observe the Sun and the solar corona, and four in-situ instruments to measure the solar wind, its thermal and energetic particles, and electromagnetic fields Müller, D., O.C.St. Cyr, I. Zouganelis, et al., Astron. Astrophys., 2020; DOI: doi.org/10.1051/0004-6361/202038467 Müller, D., Marsden, R.G., St. Cyr, O.C. et al., Solar Orbiter, Sol. Phys., 285, 25–70 (2013); doi.org/10.1007/s11207-012-008
Creator Contact	Prof. T. Horbury, Principal Investigator, Imperial College, United Kingdom, t.horbury@imperial.ac.uk
Publisher And Registrant	European Space Agency

2. DOI at ESDC

Google Dataset Search (GDS) is a search engine from Google, launched in January 2020.
https://datasetsearch.research.google.com

All 25,000+ datasets are searchable with GDS

Masson et al., Google dataset search and DOI for data in the ESA space science archives, Adv. Space Res., 67, 8, 2504-2516, 2021
https://doi.org/10.1016/j.asr.2021.01.035



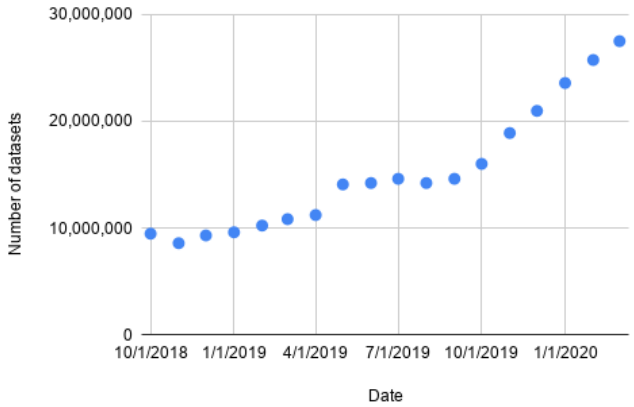
Google Dataset Search

How to get indexed?

Google Dataset Search (GDS) is a search engine from Google, launched in January 2020.
https://datasetsearch.research.google.com

To get indexed, GDS requires to include a **JSON script** with at least **two compulsory properties**: name and description

```
<script type=application/ld+json>
{
@context: http://schema.org/
@type: Dataset
name: GOLF
description: Global Oscillations at Low Frequencies
}
</script>
```



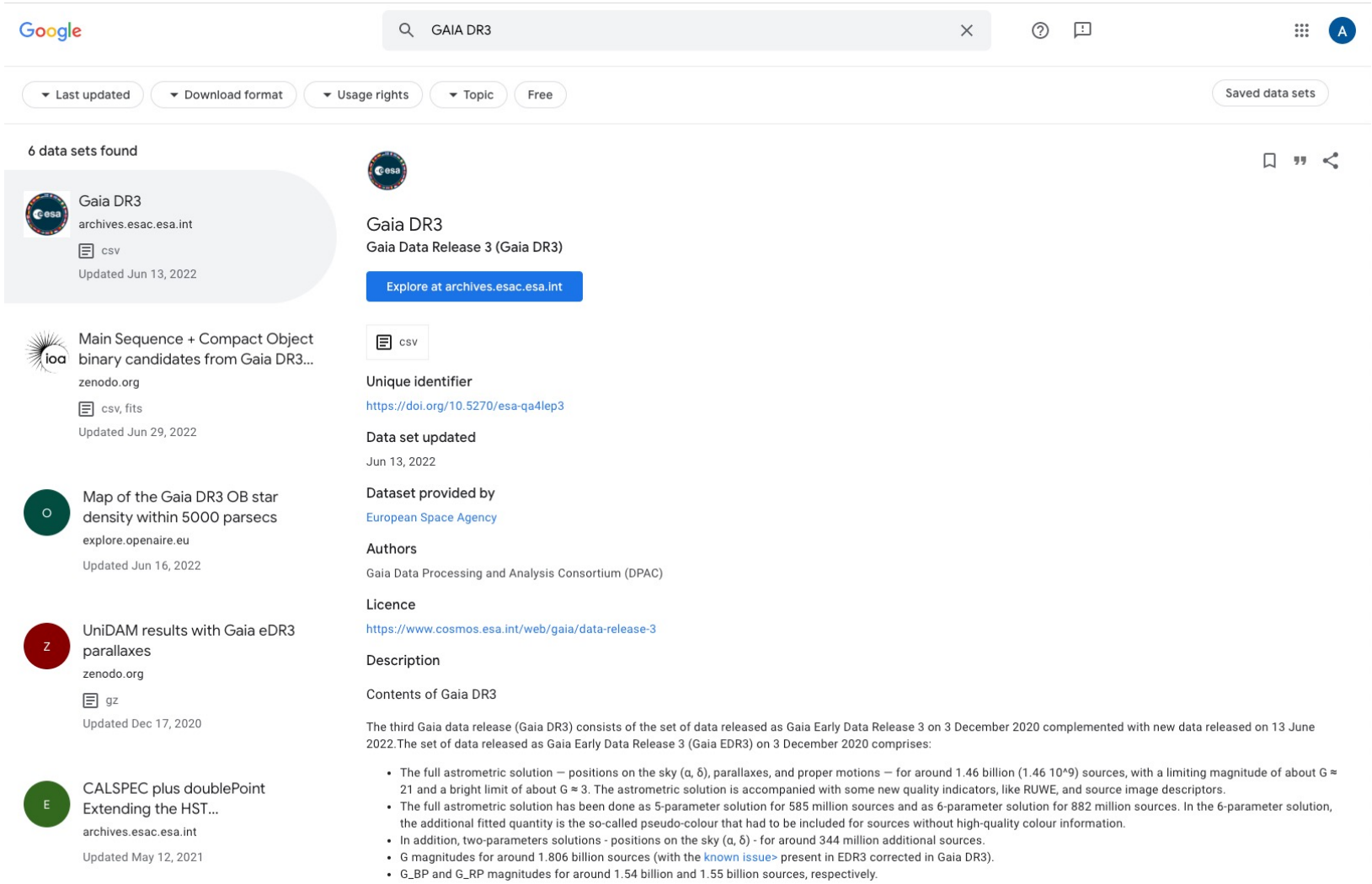
Google Dataset Search

How to get indexed?

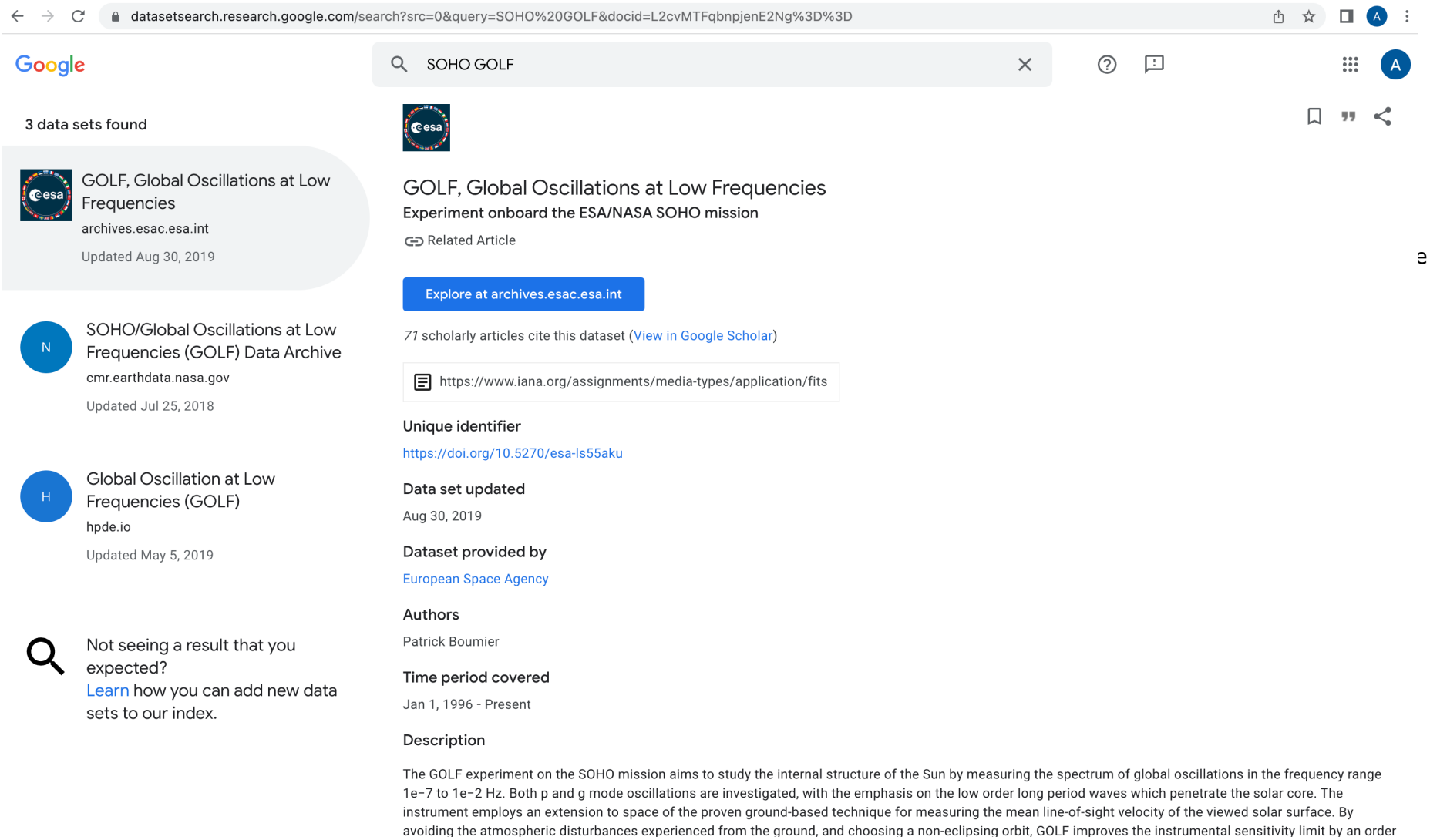
Schema.org property used in Heliophysics DOI landing pages JSON script	Google Dataset Search compliance	Remark
@context	Compliant	
@type	Compliant	
name	Compliant	
alternate name	Compliant	This has been used to specify the mission, as sometimes the same type of experiment with the same name has been used on multiple missions like FGM, PEACE or ASPOC. It could also be used to specify the names of the different instruments of an experiment.
citation	Compliant	Key scientific paper listed, those agreed with PS/PI. At the moment, only one reference is linked (May 2020)
identifier	Compliant	this should be the DOI address
creator	Compliant	The creator or author of a dataset. To uniquely identify individuals, Google advises to use ORCID ID as the value of the `sameAs` property of the `Person` type. To uniquely identify institutions and organizations, use ROR ID. Implemented when available.
publisher	Compliant	Contains web address and new ESA logo
distribution	Compliant	Direct link to data download. Added when relevant, like for mission long files for SOHO.
temporalCoverage	Compliant	Correct format implemented
dateModified	Compliant	Added when dataset was updated recently, or last update is known
description	Compliant	Abstract content added before dataset description content, double quotes and HTML tags for hyperlinks removed. Link to the ESDC archive added.
keywords	Compliant	Keywords agreed with PS/PI
audience	Compliant	Manually edited per mission to fit scientific communities

Table 1. Schema.org properties used in the Heliophysics DOI landing pages

Please search ‘GAIA DR3’ @ https://datasetsearch.research.google.com/



Please search ‘SOHO GOLF’ @ https://datasetsearch.research.google.com



Direct link to the data

A dataset provided by the European Space Agency

esa

soho

Name	GOLF, Global Oscillations at Low Frequencies
Mission	SOHO
URL	https://www.cosmos.esa.int/web/soho/mission-long-files
DOI	10.5270/esa-ls55aku
Abstract	The GOLF experiment on the SOHO mission aims to study the internal structure of the Sun by measuring the spectrum of global oscillations in the frequency range 10^{-4} to 10^{-2} Hz. Both p and g mode oscillations are investigated, with the emphasis on the low order long period waves which penetrate the solar core. The instrument employs an extension to space of the proven ground-based technique for measuring the mean line-of-sight velocity of the viewed solar surface. By avoiding the atmospheric disturbances experienced from the ground, and choosing a non-eclipsing orbit, GOLF improves the instrumental sensitivity limit by an order of magnitude to 1 mm s^{-1} over 20 days for frequencies higher than 2.10^{-4} Hz. A sodium vapour resonance cell is used in a longitudinal magnetic field to sample the two wings of the solar absorption line. The addition of a small modulating field component enables the slope of the wings to be measured. This provides not only an internal calibration of the instrument sensitivity, but also offers a further possibility to recognise, and correct for, the solar background signal produced by the effects of solar magnetically active regions. The use of an additional rotating polariser enables measurement of the mean solar line-of-sight magnetic field, as a secondary objective.
Description	GOLF key scientific data products are mission long files related to calibrated line of sight velocities based on measurements collected either through the instrument PhotoMultiplier 1 (PM1), PhotoMultiplier 2 (PM2), or a mean of the two signals (PM1+PM2). The calibration of these data is based on method described in Appourchaux et al., 2018.
Publication	Gabriel, A.H., et al., Global Oscillations at Low Frequency from the SOHO mission (GOLF), Sol. Phys., 162, 61–99 (1995); https://doi.org/10.1007/BF00733427 Appourchaux, T., et al., Searching for g modes. I. A new calibration of the GOLF instrument, A&A, 617, A108, 2018; https://doi.org/10.1051/0004-6361/201833535

SCIENCE MISSIONS | EUROPEAN SPACE AGENCY | SCIENCE &TECHNOLOGY | SIGN I

soho esa

Home | SOHO Science Archive | Command line access (TAP) | Mission Long files | Livelink (SOHO) | GDPR | Contact point

SOHO MISSION LONG FILES

This page contains a list of the SOHO mission-long file bundles generated for certain instruments thus far.
You can still use the ESA SOHO Science Archive graphical user interface or the ESA SOHO Science Archive command line access (TAP server) to search for individual observations from these instruments (e.g. images, daily files, spectra), but those might not have final calibrations applied. If no files loads below, try refreshing or switching to a different browser.

CELIAS

Mission long files	Direct download
Solar EUV flux at 15 s cadence (CELIAS Solar EUV Monitor)	Click here
Solar EUV flux at 1 day cadence (CELIAS Solar EUV Monitor)	Click here
Solar wind parameters at 5 minutes cadence (CELIAS Proton Monitor)	Click here
Solar wind parameters at 30 s cadence (CELIAS Proton Monitor)	Click here

COSTEP

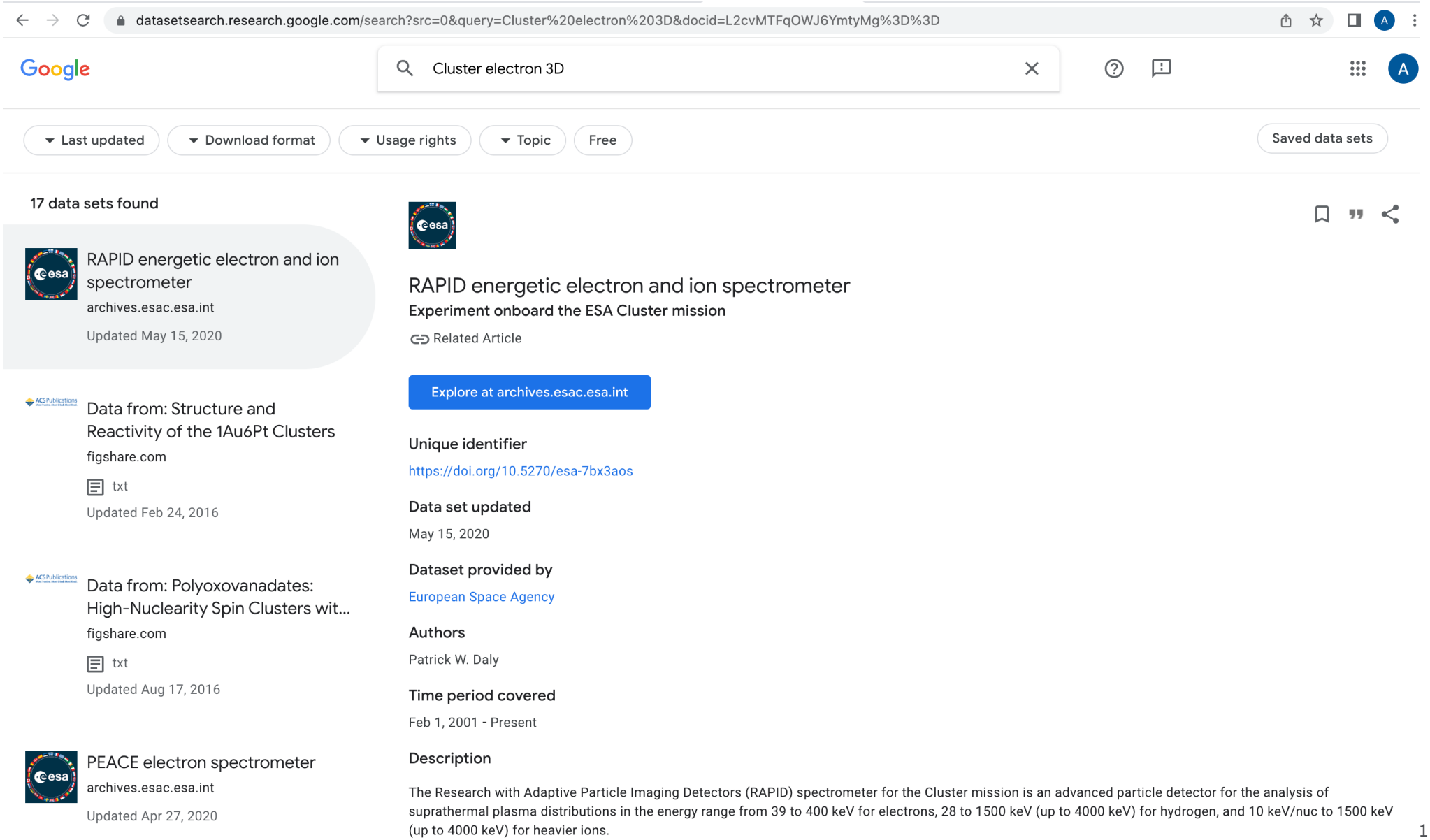
Mission long files	Direct download
Proton and He fluxes at 1 mn cadence (COSTEP EPHIN)	Click here
Proton and He fluxes at 5 mn cadence (COSTEP EPHIN)	Click here
Proton and He fluxes at 10 mn cadence (COSTEP EPHIN)	Click here
Proton and He fluxes at 30 mn cadence (COSTEP EPHIN)	Click here
Proton and He fluxes at 1 h cadence (COSTEP EPHIN)	Click here
Proton and He fluxes at 1 day cadence (COSTEP EPHIN)	Click here

GOLF

Mission Long files	Direct download
Line of sight calibrated velocity through PhotoMultiplier 1 (PM1)	Click here
Line of sight calibrated velocity through PhotoMultiplier 2 (PM2)	Click here
Line of sight calibrated velocity Mean of PM1 and PM2	Click here

cillary mission long files
LF_D4.2_SVEL_960411_121005
F_FrequencyShiftTables_960411_130407
OLF_FrequencyTables_960411_130407

Direct download of the mission long files from the ESA SOHO Archive


datasetsearch.research.google.com/search?src=0&query=Cluster%20electron%203D&docid=L2cvMTFqOWJ6YmtyMg%3D%3D
Google
Cluster electron 3D
Last updated
Download format
Usage rights
Topic
Free
Saved data sets
17 data sets found
RAPID energetic electron and ion spectrometer
archives.esac.esa.int
Updated May 15, 2020
Data from: Structure and Reactivity of the 1Au6Pt Clusters
figshare.com
txt
Updated Feb 24, 2016
Data from: Polyoxovanadates: High-Nuclearity Spin Clusters wit...
figshare.com
txt
Updated Aug 17, 2016
PEACE electron spectrometer
archives.esac.esa.int
Updated Apr 27, 2020
RAPID energetic electron and ion spectrometer
Experiment onboard the ESA Cluster mission
Related Article
Explore at archives.esac.esa.int
Unique identifier
https://doi.org/10.5270/esa-7bx3aos
Data set updated
May 15, 2020
Dataset provided by
European Space Agency
Authors
Patrick W. Daly
Time period covered
Feb 1, 2001 - Present
Description
The Research with Adaptive Particle Imaging Detectors (RAPID) spectrometer for the Cluster mission is an advanced particle detector for the analysis of suprathermal plasma distributions in the energy range from 39 to 400 keV for electrons, 28 to 1500 keV (up to 4000 keV) for hydrogen, and 10 keV/nuc to 1500 keV (up to 4000 keV) for heavier ions.

Google

Solar Orbiter MAG dataset

All Images News Videos Shopping More Tools

About 72,700 results (0.47 seconds)

https://archives.esac.esa.int › data › SolarOrbiter › MAG

Dataset provided by the European Space Agency

The **Solar Orbiter MAG** magnetometer is a high performance dual fluxgate instrument, with data rates up to 128 vectors/s from each sensor. Each sensor has a noise ...

Your page is not mobile-friendly.

https://issues.cosmos.esa.int › attachments › Solar ... PDF

Solar Orbiter magnetometer user guide - 1

Sep 29, 2020 — This document describes the publicly released data from the magnetometer experiment (**MAG**) on the. **Solar Orbiter** spacecraft.

https://www.aanda.org › articles › pdf › 2020/10

The Solar Orbiter magnetometer - Astronomy & Astrophysics

by TS Horbury · 2020 · Cited by 72 — T. S. Horbury et al.: The **Solar Orbiter** magnetometer. Table 1. **MAG** performance requirements, including measurement requirements and constraint…

https://www.aanda.org › articles › full_html › 2020/10

The Solar Orbiter magnetometer - Astronomy & Astrophysics

by TS Horbury · 2020 · Cited by 72 — The magnetometer (**MAG**) instrument will measure the structure of coronal mass ejections (CMEs) and, in conjunction with other missions, their…

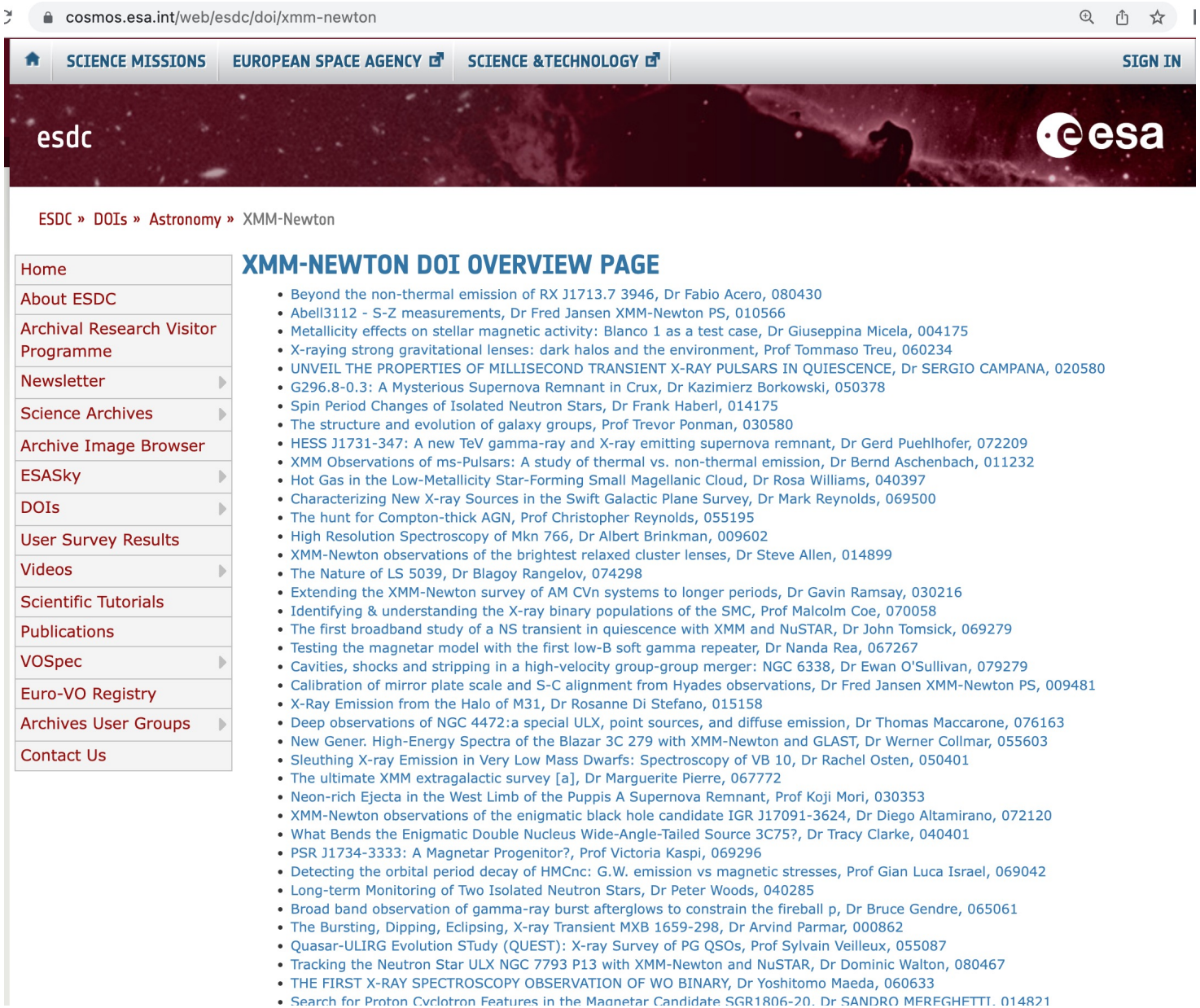
https://cdaweb.gsfc.nasa.gov › misc › NotesS

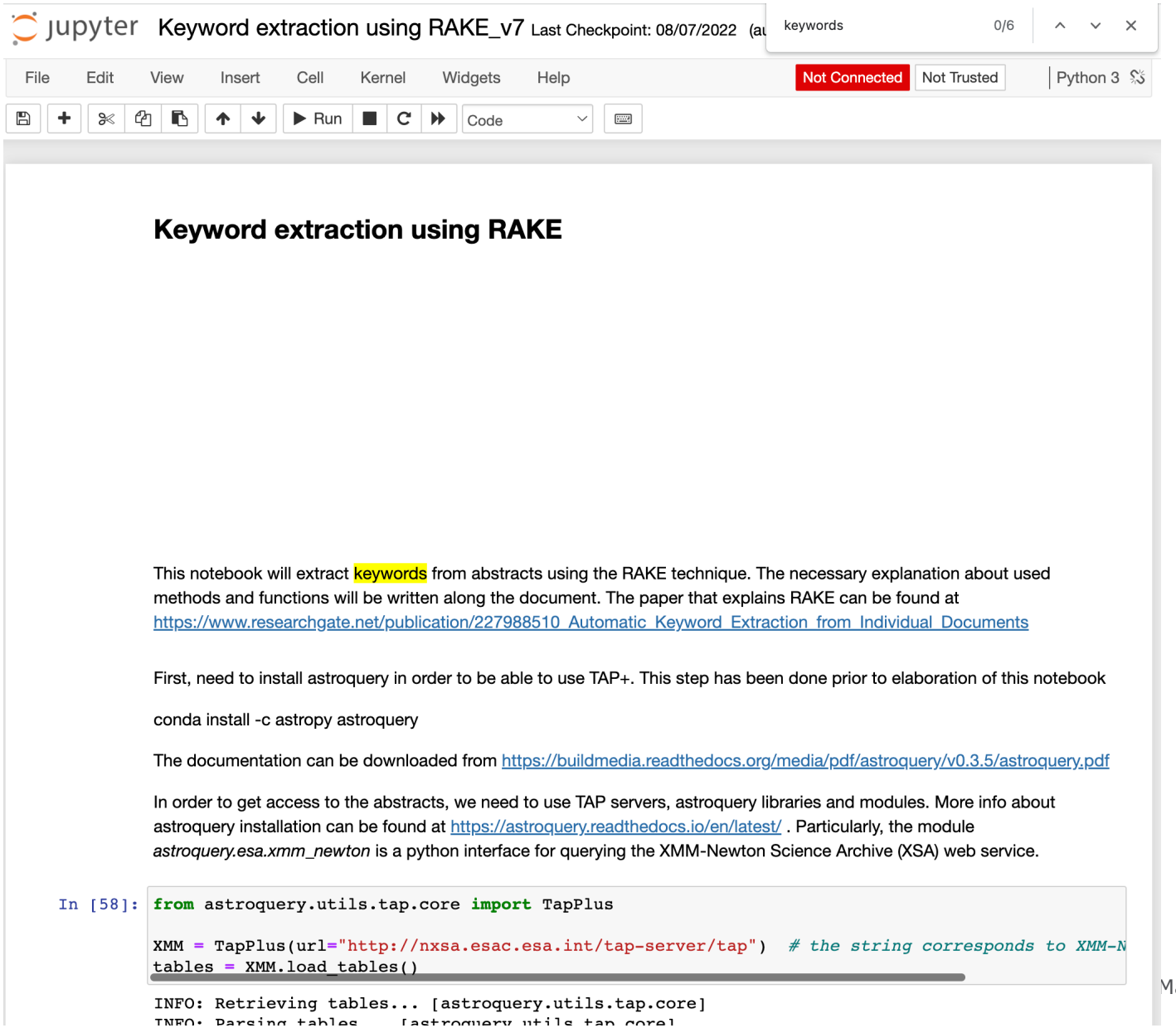
SPDF - Coordinated Data Analysis Web (CDAWeb) - NASA

CDAWeb Served Heliophysics **Datasets** Beginning with 'S' ... STEREO Behind IMPACT/**MAG** Magnetic Field and PLASTIC **Solar** Wind Plasma Data - J. Luhmann (UCB/SSL) ...

A dataset provided by the European Space Agency

Name	MAG, Solar Orbiter magnetometer
Mission	Solar Orbiter
URL	http://soar.esac.esa.int/
DOI	https://doi.org/10.5270/esa-ux7y320
Abstract	The Solar Orbiter MAG magnetometer is a high performance dual fluxgate instrument, with data rates up to 128 vectors/s from each sensor. Each sensor has a noise level of around 5pT at 1 Hz, allowing access to the ion kinetic regime and, near perihelion, access to electron kinetic phenomena at tens of Hz. Significant signals from spacecraft subsystems and other instruments can be present in the raw data. Extensive cleaning is undertaken to remove as much of the artificial signals as possible, resulting in what is normally a very high quality data set, but some artifacts are present on occasion.
Description	Level 2 data files contain the best estimates of the magnetic field and have signals from the spacecraft and other instrument removed; flags within the files indicate the quality of the data as a function of time. Normal mode data coverage is nearly complete. Burst mode coverage varies from a few minutes a day to complete, depending on available telemetry. Level 0 and 1 data files include data from both inboard and outboard sensors.
Publication	Horbury, T.S., et al., The Solar Orbiter magnetometer , A&A, 642, A9, https://doi.org/10.1051/0004-6361/201937257


cosmos.esa.int/web/esdc/doi/xmm-newton
SCIENCE MISSIONS
EUROPEAN SPACE AGENCY
SCIENCE &TECHNOLOGY
SIGN IN
esdc
esa
ESDC » DOIs » Astronomy » XMM-Newton
Home
About ESDC
Archival Research Visitor Programme
Newsletter
Science Archives
Archive Image Browser
ESASky
DOIs
User Survey Results
Videos
Scientific Tutorials
Publications
VOSpec
Euro-VO Registry
Archives User Groups
Contact Us
XMM-NEWTON DOI OVERVIEW PAGE
• Beyond the non-thermal emission of RX J1713.7 3946, Dr Fabio Acero, 080430
• Abell3112 - S-Z measurements, Dr Fred Jansen XMM-Newton PS, 010566
• Metallicity effects on stellar magnetic activity: Blanco 1 as a test case, Dr Giuseppina Micela, 004175
• X-raying strong gravitational lenses: dark halos and the environment, Prof Tommaso Treu, 060234
• UNVEIL THE PROPERTIES OF MILLISECOND TRANSIENT X-RAY PULSARS IN QUIESCENCE, Dr SERGIO CAMPANA, 020580
• G296.8-0.3: A Mysterious Supernova Remnant in Crux, Dr Kazimierz Borkowski, 050378
• Spin Period Changes of Isolated Neutron Stars, Dr Frank Haberl, 014175
• The structure and evolution of galaxy groups, Prof Trevor Ponman, 030580
• HESS J1731-347: A new TeV gamma-ray and X-ray emitting supernova remnant, Dr Gerd Puehlhofer, 072209
• XMM Observations of ms-Pulsars: A study of thermal vs. non-thermal emission, Dr Bernd Aschenbach, 011232
• Hot Gas in the Low-Metallicity Star-Forming Small Magellanic Cloud, Dr Rosa Williams, 040397
• Characterizing New X-ray Sources in the Swift Galactic Plane Survey, Dr Mark Reynolds, 069500
• The hunt for Compton-thick AGN, Prof Christopher Reynolds, 055195
• High Resolution Spectroscopy of Mkn 766, Dr Albert Brinkman, 009602
• XMM-Newton observations of the brightest relaxed cluster lenses, Dr Steve Allen, 014899
• The Nature of LS 5039, Dr Blagoy Rangelov, 074298
• Extending the XMM-Newton survey of AM CVn systems to longer periods, Dr Gavin Ramsay, 030216
• Identifying & understanding the X-ray binary populations of the SMC, Prof Malcolm Coe, 070058
• The first broadband study of a NS transient in quiescence with XMM and NuSTAR, Dr John Tomsick, 069279
• Testing the magnetar model with the first low-B soft gamma repeater, Dr Nanda Rea, 067267
• Cavities, shocks and stripping in a high-velocity group-group merger: NGC 6338, Dr Ewan O'Sullivan, 079279
• Calibration of mirror plate scale and S-C alignment from Hyades observations, Dr Fred Jansen XMM-Newton PS, 009481
• X-Ray Emission from the Halo of M31, Dr Rosanne Di Stefano, 015158
• Deep observations of NGC 4472:a special ULX, point sources, and diffuse emission, Dr Thomas Maccarone, 076163
• New Gener. High-Energy Spectra of the Blazar 3C 279 with XMM-Newton and GLAST, Dr Werner Collmar, 055603
• Sleuthing X-ray Emission in Very Low Mass Dwarfs: Spectroscopy of VB 10, Dr Rachel Osten, 050401
• The ultimate XMM extragalactic survey [a], Dr Marguerite Pierre, 067772
• Neon-rich Ejecta in the West Limb of the Puppis A Supernova Remnant, Prof Koji Mori, 030353
• XMM-Newton observations of the enigmatic black hole candidate IGR J17091-3624, Dr Diego Altamirano, 072120
• What Bends the Enigmatic Double Nucleus Wide-Angle-Tailed Source 3C75?, Dr Tracy Clarke, 040401
• PSR J1734-3333: A Magnetar Progenitor?, Prof Victoria Kaspi, 069296
• Detecting the orbital period decay of HMCnc: G.W. emission vs magnetic stresses, Prof Gian Luca Israel, 069042
• Long-term Monitoring of Two Isolated Neutron Stars, Dr Peter Woods, 040285
• Broad band observation of gamma-ray burst afterglows to constrain the fireball p, Dr Bruce Gendre, 065061
• The Bursting, Dipping, Eclipsing, X-ray Transient MXB 1659-298, Dr Arvind Parmar, 000862
• Quasar-ULIRG Evolution STudy (QUEST): X-ray Survey of PG QSOs, Prof Sylvain Veilleux, 055087
• Tracking the Neutron Star ULX NGC 7793 P13 with XMM-Newton and NuSTAR, Dr Dominic Walton, 080467
• THE FIRST X-RAY SPECTROSCOPY OBSERVATION OF WO BINARY, Dr Yoshitomo Maeda, 060633
• Search for Proton Cyclotron Features in the Magnetar Candidate SGR1806-20, Dr SANDRO MEREGHETTI, 014821



Work has been performed with mission specialists to define typical keywords/patterns in their community

Added long list of stop words to be neglected

Eliminating plurals and redundant keywords

Title : The X-ray Nova CI CAM (XTE J0421+560) in Quiescence
Abstract : We propose a 30 ks XMM observation of the X-ray transient XTE J0421+56 (CI Cam) in quiescence. During part of a 10 day outburst in 1998 April, this source exhibited a relativistic radio jet and moving X-ray emission lines. Green Bank monitoring indicates that CI Cam continues to be a radio source. XTE J0421+566 was clearly detected in quiescence by BeppoSAX in a 45 ks observation, some 150 days after the outburst. The count rate was too low to obtain a useful spectrum. The proposed XMM observation will provide this, and allow a sensitive search for the intense oxygen feature seen during the latter part of the outburst by BeppoSAX.

Keywords auto-extracted with the open source natural language RAKE MIT python algorithm

XTE J0421+560 10
BeppoSAX 10
XMM 10
relativistic radio jet 8.5
ci cam 8.0
quiescence 3.0

DOI @ ESA: achievements and way forward

- 25,000+ DOI minted in 2021 for ESA science datasets
- All searchable on Google Dataset Search and Google
- 5000+ XMM DOI landing pages updated in September 2022 to improve findability per observation proposal by adding keywords in their JSON script
- Same will be implemented to other observatory missions (CHEOPS, Herschel…)
- Mid-tem future: DOI on the fly to improve FAIR principles, reproducibility and findability