

Beste Daten für alle: Das Virtuelle Observatorium

Markus Demleitner

AK Astronomie der Universität Stuttgart, 8. Januar 2025

Das Virtuelle Observatorium...

...ist *nicht*

- Google Sky oder das World Wide Telescope
- Ein Programm
- Eine Webseite
- Fernsteuern von Teleskopen



– all diese Dinge haben aber mit dem VO zu tun, und ESA Sky z.B. *ist* VO.

Das Virtuelle Observatorium ist...

Ein **umfassendes** Netz von
Daten und Diensten,
mit Relevanz für die **Astronomie**,
die mit **Standardmethoden**
von überallher genutzt werden können
und die Ergebnisse der digitalen Astronomie
erhält.

Trotzdem: Das VO ist beileibe nicht nur für Profis und Nerds.

Warum die Mühe?

Zwei große Triebkräfte bestimmen die Entwicklung des VO:

- Größe und
- Diversität von Daten.

Noch in den 1990ern

Der Stapel Bücher nebenan ist der Hipparcos-Sternkatalog mit genauen Positionen, Eigenbewegungen und ähnlichen Daten von 118218 Sternen.

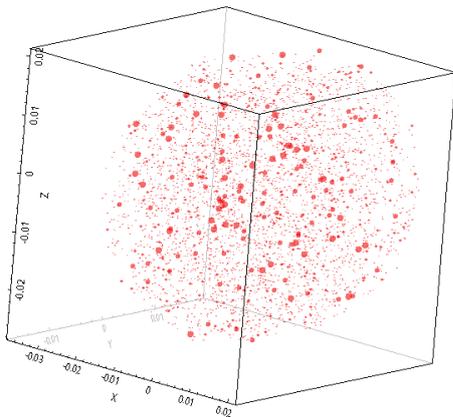
Für heutige Verhältnisse ist das ein kleiner Katalog.



Exkurs: Kataloge?

Listen von astronomischen Objekten, ihren Positionen und Eigenschaften heißen in der Astronomie „Katalog“.

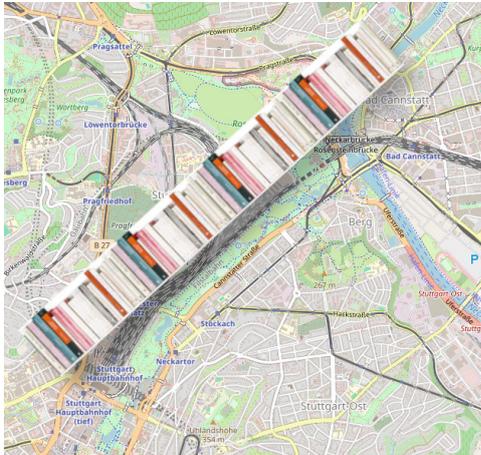
AstronomInnen machen und brauchen Kataloge, um herauszufinden, wie unsere Galaxis aussieht, wie Sterne sterben, wo es nach dunkler Materie aussieht – oder eigentlich immer.



Das Bild zeigt die 6300 Sterne, die derzeit innerhalb von rund 100 Lichtjahren um die Sonne bekannt sind, in einer quasiräumlichen Darstellung. Wer selbst mit den Daten spielen will, kann nach [cns5](#) suchen.

Zwei Milliarden Objekte

Moderne Kataloge wie der des Astrometriesatelliten Gaia enthalten Milliarden von Objekten. Gaia DR3 entspricht rund 20'000 Hipparcos-Katalogen. Gedruckt: Ein Bücherbrett von 4 km Länge.



Dabei ist das Problem nicht einmal so sehr, solche Datenmengen zu transportieren oder zu speichern. Das würde schon noch gehen. Aber um etwas Sinnvolles damit anzustellen, müssen solche Daten indiziert oder anderweitig zugänglich gemacht werden.

Überschlagsrechnung: Wer 1000 Objekte pro Sekunde verarbeitet (und schon das ist auch auf modernen Systemen ggf. eine Herausforderung), braucht für zwei Milliarden Objekte zwei Millionen Sekunden. Ein Jahr sind rund $\pi \times 10^7$ Sekunden, also reden wir von einem runden Fünfzehnteljahr oder einem Monat.

Und dann die Zeitachse

Der Himmel ist nicht statisch. Wenn wir die zeitliche Entwicklung dokumentieren, explodieren die Datenmengen weiter.

Das Rubin Observatory wird ab demnächst 30 TB Daten pro Nacht produzieren.

Am Ende reden wir über Bücherstapel kosmischen Ausmaßes. Neben an Erde und Mond neben den Rubin-Bücherstapel gemalt.

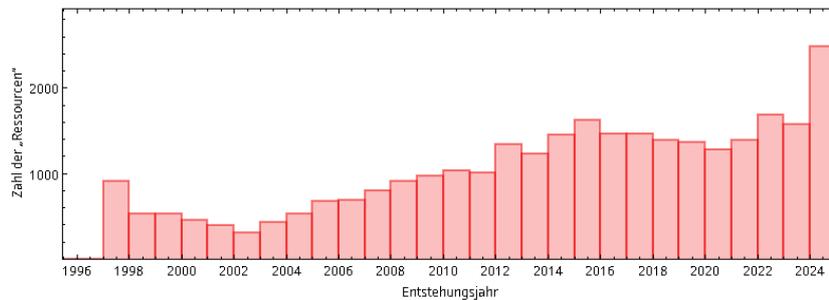


Noch Wichtiger: Wachsende Erkenntnis

Große Daten lassen sich letztlich in der Regel mit Hardware bewältigen.

Die Diversität und den Reichtum der Quellen ist aber nur nutzbar, wenn Maschinen diese automatisch auswerten können: Gute Metadaten und maschinenzugängliche Protokolle sind entscheidend.

Jedes Jahr über 1000 neue Datensammlungen im VO:



Und das VO?

- Standardisierte Metadaten in der *Registry* erlauben das systematische Finden relevanter Datensammlungen.
- Durch *typisierte Protokolle* (z.B. für Kataloge, Bilder, Spektren) können verschiedene Datensammlungen in einheitlicher Weise behandelt werden.
- Das *Table Access Protocol* erlaubt das Auslagern vieler Filteraufgaben auf entfernte Rechner mit vorbereiteten Großdaten.
- Einheitliche, maschinenlesbare Metadaten erlauben die Nutzung vieler verschiedener Daten mit überschaubarem Aufwand.
- und vieles mehr; alle Standards auf <https://ivoa.net/documents>.

(Aber natürlich sollte (fast) niemand all die Standards lesen müssen).

VO Lernen

Fast alles im VO ist öffentlich zugänglich; vieles ist sicher nur für WissenschaftlerInnen und Astro-Nerds zumutbar. Anderes ist auch für Laien durchaus interessant.

Wer den Umgang lernen will:

- <https://dc.g-vo.org/VOTT> hat Lern- und Lehrmaterial, insbesondere
- unsere Vorlesung zur Nutzung des VO <https://doi.org/10.21938/avVAXDIGOiu0Byv7NOZCsQ>.

Demo Time

Um einen Eindruck von der Arbeit mit dem VO zu geben, und gleichzeitig um zu illustrieren, was für ein dramatischer Fortschritt der Gaia-Astrometriesatellit ist, wollen wir Farbenhelligkeitsdiagramme offener Sternhaufen ansehen.

Ich verwende hier die beiden Clients [TOPCAT](#) und [Aladin](#), die jeweils eine Java VM auf dem Rechner brauchen.

Wer mit dem VO arbeiten will und Python kann, sollte sich auch [pyVO](#) ansehen.

Wir wollen zunächst den letzten tiefen All-Sky-Katalog vor Gaia nutzen, den PPMXL; auch er enthielt schon eine Milliarde Sterne.

Daten Holen

Dazu im TOPCAT VO → Cone Search wählen, nach PPMXL suchen, einen geeigneten Dienst auswählen, unter *Object Name* "Pleiades" eingeben, *Resolve* klicken und 1 Grad als Radius eingeben.

The screenshot shows the TOPCAT Cone Search interface. The 'Available Cone Services' window is open, displaying search results for 'ppmxl'. The 'Cone Parameters' window is also visible, showing the search parameters for the Pleiades cluster.

Short Name	Title	Description
GAVO PPMXL	The PPMXL Catalog	stars, surveys, astrometry, proper-motions
I/317	The PPMXL Catalog	surveys, astrometry, proper-motions, visible-astro
PPMXL	PPMXL Catalog	all sky survey
PPMXLBMD	PPM-XL Bright M Dwarfs Catalog	Star
XAO PPMXL	PPMXL Catalog Cone Search	Reference systems, Astrometry, Catalogs, Stars: P

Resource Count: 5

Cone Parameters

Cone URL: <http://dc.zah.uni-heidelberg.de/ppmxl/q/cone/scs.xml?>

Object Name: Pleiades

RA: 56.601 degrees (J2000) Accept Sky Positions

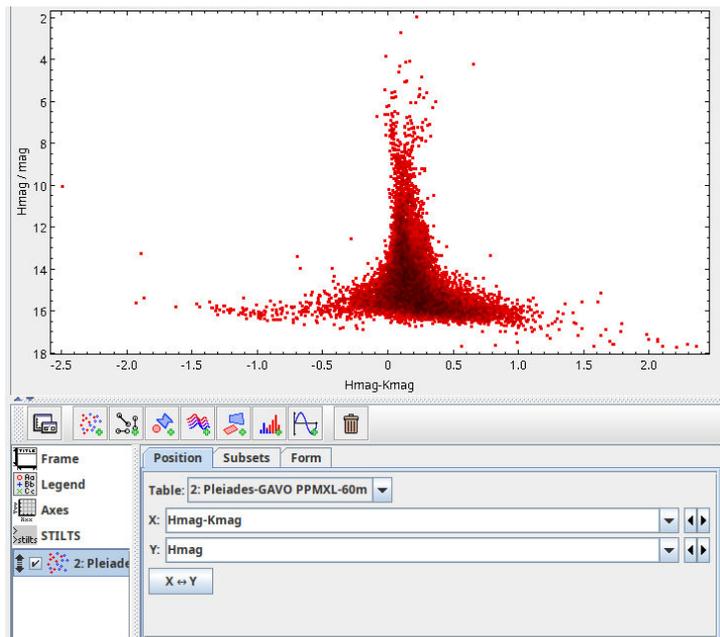
Dec: 24.114 degrees (J2000)

Radius: 1 degrees

Verboisity: 2 (normal)

Nach Ok kommt recht schnell eine Tabelle zurück. Graphics → Plane Plot, und dort Hmag-Kmag versus Hmag plotten, die Ordinate unter *Axes* flippen. Das gibt dann etwa so etwas:

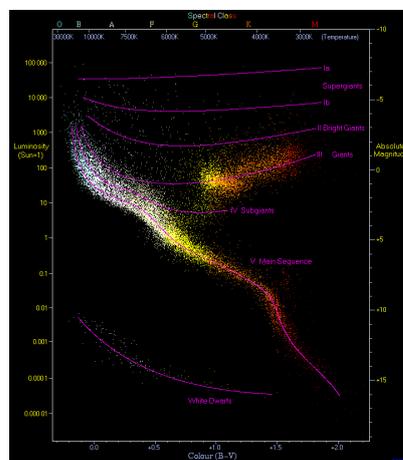
Damit: PPMXL FHD



Ich nutze hier Infrarot-Magnituden, hauptsächlich, weil die optischen Größen im PPMXL sehr haarig im Umgang sind. Wer damit spielen will, bekommt sie durch setzen von von *Verbosity* auf 3 im Cone Search-Fenster. Durch *Views* → *Column Info* sind die Spalten leicht zu finden, aber man muss dann die schlechten Werte rausfiltern, bevor irgendwas Nutzbares rauskommt. Die Infrarot-Magnituden sind deutlich besser.

Aber:

Ein Farben-Helligkeits-Diagramm sollte doch so aussehen?



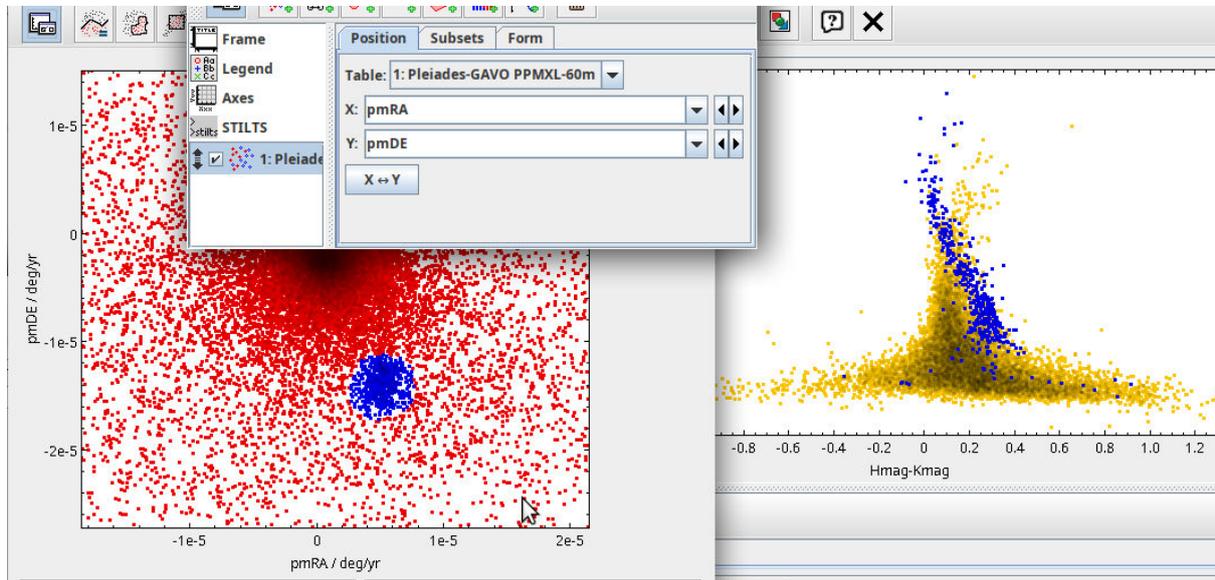
CC-BY wikipedia:Richard Powell

Nun: In unserem Diagramm sind Sterne komplett verschiedener Entfernungen durcheinander: Auf der Ordinate sind *scheinbare*, nicht absolute Helligkeiten. Der breite Fuß wiederum ist zu guten Stücken ein Artefakt großer Fehler an schwachen Quellen (auch wenn sich dort auch etliche *wirklich* komische Objekte verbergen, Quasare, weiße Zwerge, extrem rotverschobene Galaxien und dergleichen).

Wir könnten uns im Gaia-Zeitalter geometrische Abstände von vielen der Sterne besorgen, aber das ging zu PPMXL-Zeiten noch nicht – es gab fast keine Parallaxen. Aber in unseren Daten sind ja die Plejaden, die alle ungefähr den gleichen Abstand haben. Und wir können sie nach Eigenbewegungen auswählen.

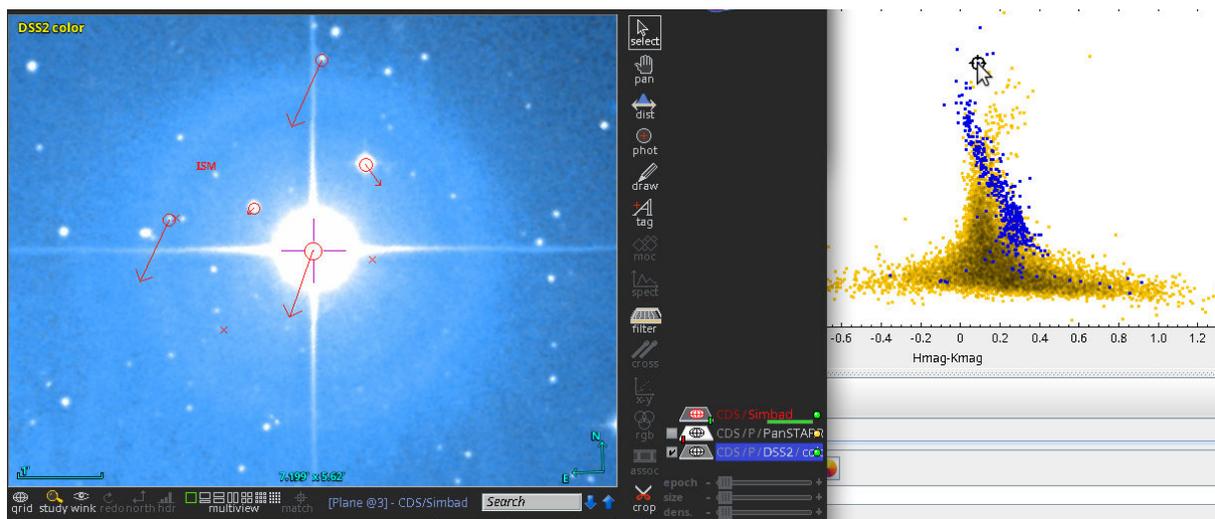
Dazu machen wir wieder einen Plane Plot, dieses Mal aber von pmRA gegen pmDE. Das dabei auftretende Viereck ist ein Artefakt der Art, wie in einem Quellkatalog des PPMXL Messungen zu Sternen kombiniert wurden. Wichtig ist der kleine Fleck neben dem großen Fleck um 0,0 herum. Den markieren wir jetzt mit *Draw Blob Subset*.

PPMXL-FHD für die Plejaden



Das sieht doch schon schön nach einer Hauptreihe aus. Wir können uns die Sterne auch ansehen. Dazu jetzt den Aladin starten und darin *DSS* auswählen, auf ein paar arcmin FoV zoomen. Im TOPCAT unter *Views* → *Activation Action Send Sky Position* abhaken. Danach können Punkte in den TOPCAT-Plots angeklickt werden, und die Quellen erscheinen im Aladin-Fenster. Es hilft, auch die Simbad-Plane im Aladin anzuschalten.

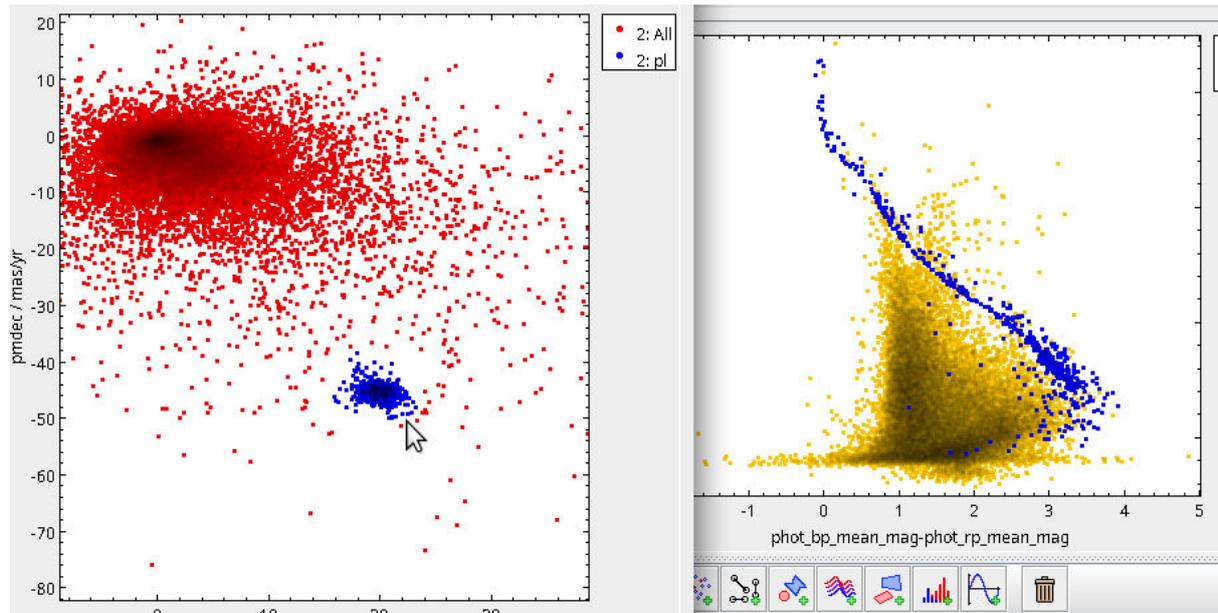
Sterne sehen mit Aladin



Die gleiche Prozedur können wir jetzt mit Gaia DR3 machen; die Spaltennamen ändern sich, aber sogar das wäre (über Metadaten namens "UCD" zu guten Stücken automatisierbar). Das einfache Springen zwischen Daten ist einer der wesentlichsten Vorteile des VO.

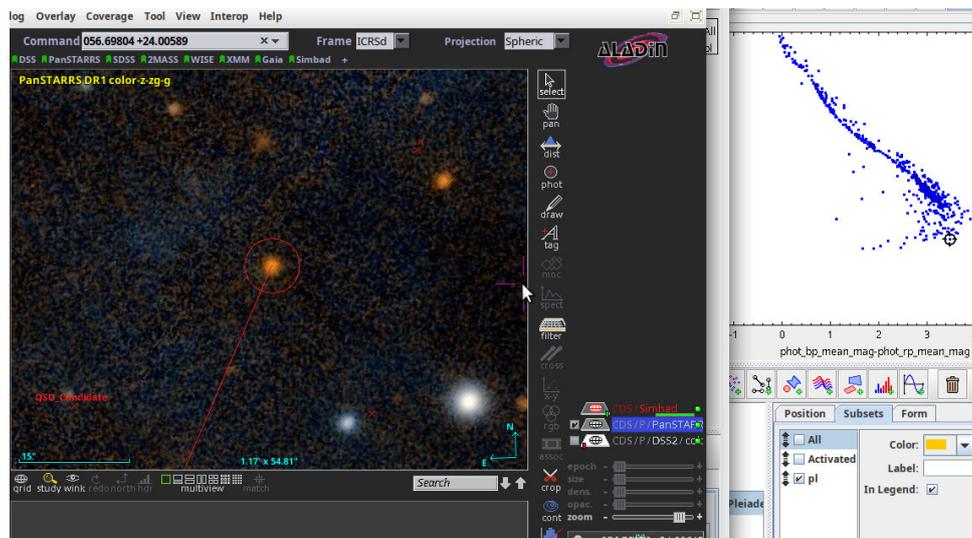
Die Gaia-Daten sind erkennbar *viel* "besser" als der PPMXL; tatsächlich sind die Plejaden hier auch ohne Eigenbewegungs-Kriterien sofort im FHD sichtbar. Die Hauptreihe wird beeindruckend schmal (die Plejaden-Sterne haben alle in etwa das gleiche Alter und die gleiche Chemie).

Ein Plejaden-FHD aus Gaia DR3



Die Farben kommen auf den moderneren PanSTARRS-Daten im Aladin besser heraus (aber hellere Sterne sind auf diesen horrös überbelichtet). Jetzt kann man zum Beispiel sehen, was die Sterne sind, die aus der Hauptreihe tanzen – oder schwache, rote Objekte ansehen.

Und wieder: Wie sieht es aus?



So könnte man zum Beispiel Braune Zwerge finden, doch sind die in den Plejaden zu schwach für Gaia. Tipp: die Hyaden sind nah genug, und innerhalb von drei Grad findet man auch mindestens einen.

Viel Spaß dabei!

Vielen Dank für Ihre Aufmerksamkeit!



Vortragsskript bei <https://docs.g-vo.org/akastro-2025.pdf>

Die Demo basiert im Wesentlichen auf <https://www.g-vo.org/tutorials/pleiades.pdf>; die Beschreibung dort ist zwar auf Englisch, aber wahrscheinlich besser nachvollziehbar.