



# Nachhaltigkeitsaspekte von Standards in Daten-Infrastrukturen: Das Beispiel Virtual Observatory

M. Demleitner

Statustreffen ErUM-Data: KI als Treiber von  
Wissenschaft und Wertschöpfung, 19.3.2026,  
Berlin

# Das Virtual Observatory (VO) ist nicht

- ein großes, **monolithisches** Institut
- eine Webseite („**Plattform**“)
- ein Satz von menschenlesbaren Webseiten
- oder **ein großes Programm**, das „Astronomie macht“.

**Standards**, die sagen, wie Daten mit Maschinenhilfe gefunden, genutzt und beschrieben werden können

plus

~ 50 **Datenzentren** in aller Welt, die sich an die Standards halten (darunter fast alle wichtigen Organisationen wie ESO, NASA, ESA)

plus

einige Freiwillige, die ein **wenig zentrale** Infrastruktur betreiben (etwa Registry-Endpunkte, Mailinglisten, Dokumentrepositorien)

plus

**Autor:innen** von Wissenschaftssoftware, Bibliotheken, Webseiten u.a., über die die Ressourcen bequem genutzt werden können.

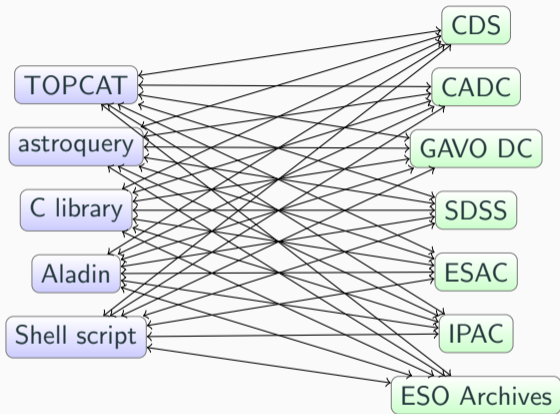
## Zwei weitere wichtige Teile förderierter Infrastrukturen

**Personen**, die die Standards ausarbeiten, aushandeln, betreuen, lehren – und die Werbung für sie machen

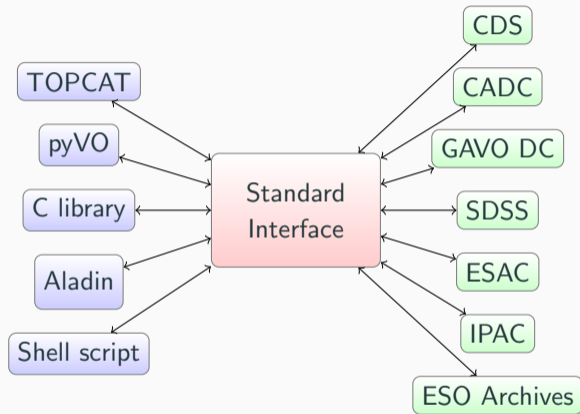
plus

**Forschungsfinanzierer**, die ihnen das ermöglichen, obwohl diese Arbeit sich nur indirekt und verzögert in Metriken spiegelt.

# Als Karikatur: $N + M$ Adapter statt $N \times M$



Ohne VO



Mit VO

Das Virtual Observatory wurde entwickelt, um Forschung zu **erleichtern** oder erst zu **ermöglichen**.

Standard-basierte Infrastrukturen haben aber **auch** erhebliche Nachhaltigkeitswirkungen:

- Bessere Daten-Nachnutzung
- Bessere Wiederverwertbarkeit von Wissenschafts-Software
- Längere Zyklen auf der Archivseite
- Föderierte Systeme sind resilienter gegen externe Interferenzen

Auch wenn Speicher und Server nicht *ganz* emissionsfrei sind:

- Archivdaten haben fast immer einen viel **kleineren Fußabdruck** als neu erhobene Daten.
- Wenn sie leicht gefunden und genutzt werden können, ergibt sich „genauso viel Wissenschaft“ bei geringerer Budget- und Umweltbelastung.

Mit standardisierten Protokollen und Metadatenmodellen können neue Instrumente oder Codes mühelos **bestehenden Code aufnehmen** und verbessern.

- Solche Entwicklungen verbessern im Nebeneffekt meist auch die Software-Infrastruktur von **Geräten im Maintenance-Modus**.
- Für Forschende ergibt sich weniger Lock-in und **mehr Nachnutzung**.

Standards bewegen sich im Konsens und damit langsam. Das ist hier gut, denn

- Dienste, die 2010 geschrieben wurden, können auch mit modernster Software interoperieren und mit **modernsten UIs** bedient werden.
- Die Entkopplung von Protokoll und UI/Wissenschaftscode erlaubt rasche Entwicklung dort, wo sie sinnvoll ist, **ohne serverseitig Aufwand zu verursachen**.
- Archive können auf lang entwickelte, eher **robuste Komponenten** zurückgreifen und können sich Eigenentwicklungen (weitgehend) sparen.

Global föderierte Datenzentren können leichter Ausfälle kompensieren und Daten vor dem Verlust bewahren.

- Ein global zentrales Archiv hängt dagegen von der Stabilität politischen Willens und technischer Einrichtungen in **einer Jurisdiktion** (und der gastgebenden Institution) ab.
- Gerade, wenn Datenzentren Software gemeinsam nutzen, ist eine **schnelle Übernahme** von Daten bei Schwierigkeiten zumindest realistisch.

## Warum macht das dann nicht jede:r?

- Eine scheinbar unproduktive Einrichtung (Datenzentrum) muss dauerhaft finanziert werden (und wird dann auch noch meist überwiegend aus dem Ausland genutzt)
- Erhebliche Anlaufinvestitionen, um eine Community zu entwickeln und Standards zu schreiben
- Geringes Interesse typischer PIs, die meist schon zufrieden sind, wenn sie irgendwie an *ihre* Daten kommen
- NIH (Ablehnung von „Not invented here“) auf der Arbeitsebene
- Allerlei Bedenken gegen allzu offene Datenpolitiken an sich

- Globale, standardbasierte Dateninfrastrukturen sind nicht nur gut für die Wissenschaft, sie fördern auch ökologische wie wissenschaftliche Nachhaltigkeit
- Da sie in der Regel orthogonal zur wissenschaftsinternen Steuerung (Metriken, Karrieremöglichkeiten u. dgl.) liegen, ist ihre Etablierung und Steuerung zumindest *auch* eine politische Aufgabe.