



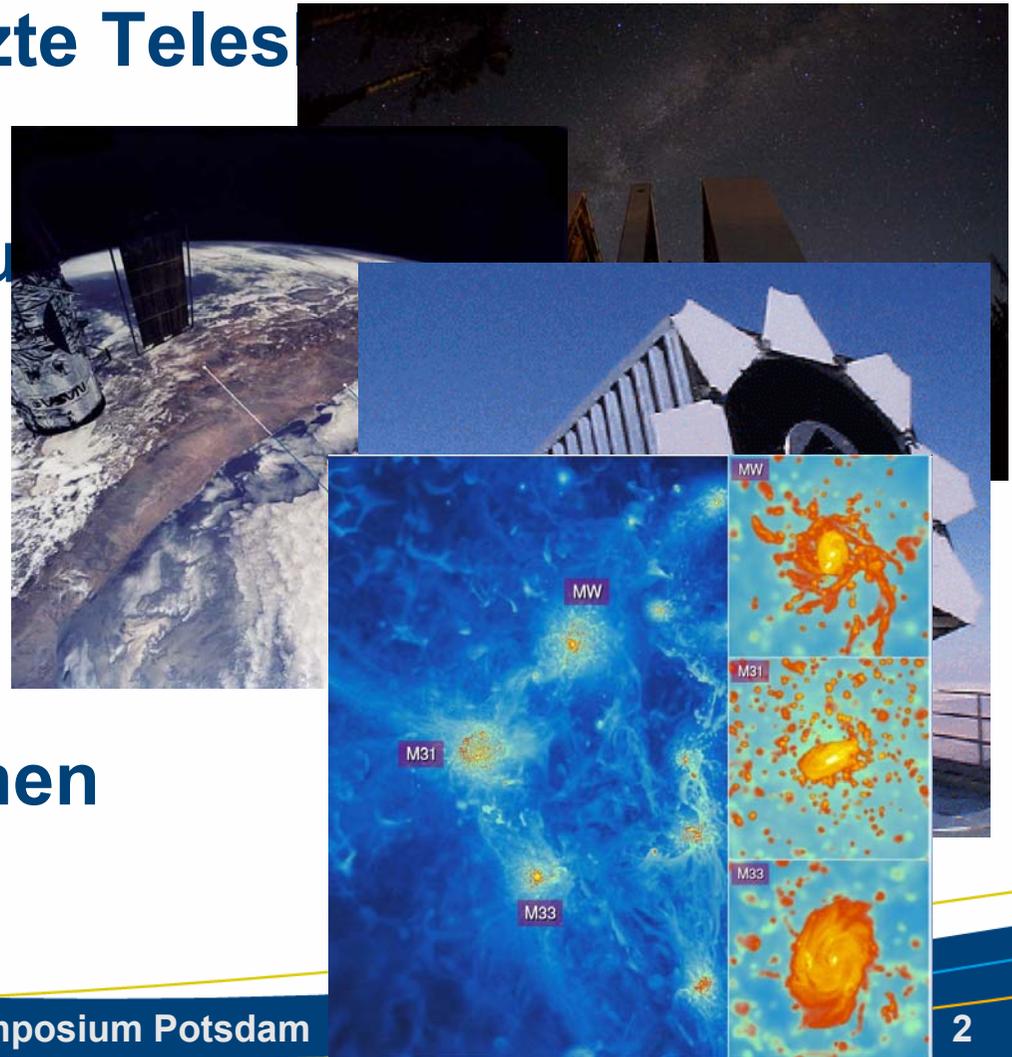
# Grids und virtuelles Observatorium: mehr als ein Google für Sterne!



Matthias Steinmetz  
Leibniz-Institut für Astrophysik Potsdam (AIP)

# Datenquellen in der Astronomie

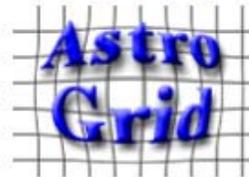
- **Große bodengestützte Teles**
- **Satelliten im Weltraum  
(keine Atmosphäre)**
- **Durchmusterungen**
- **Computersimulationen**

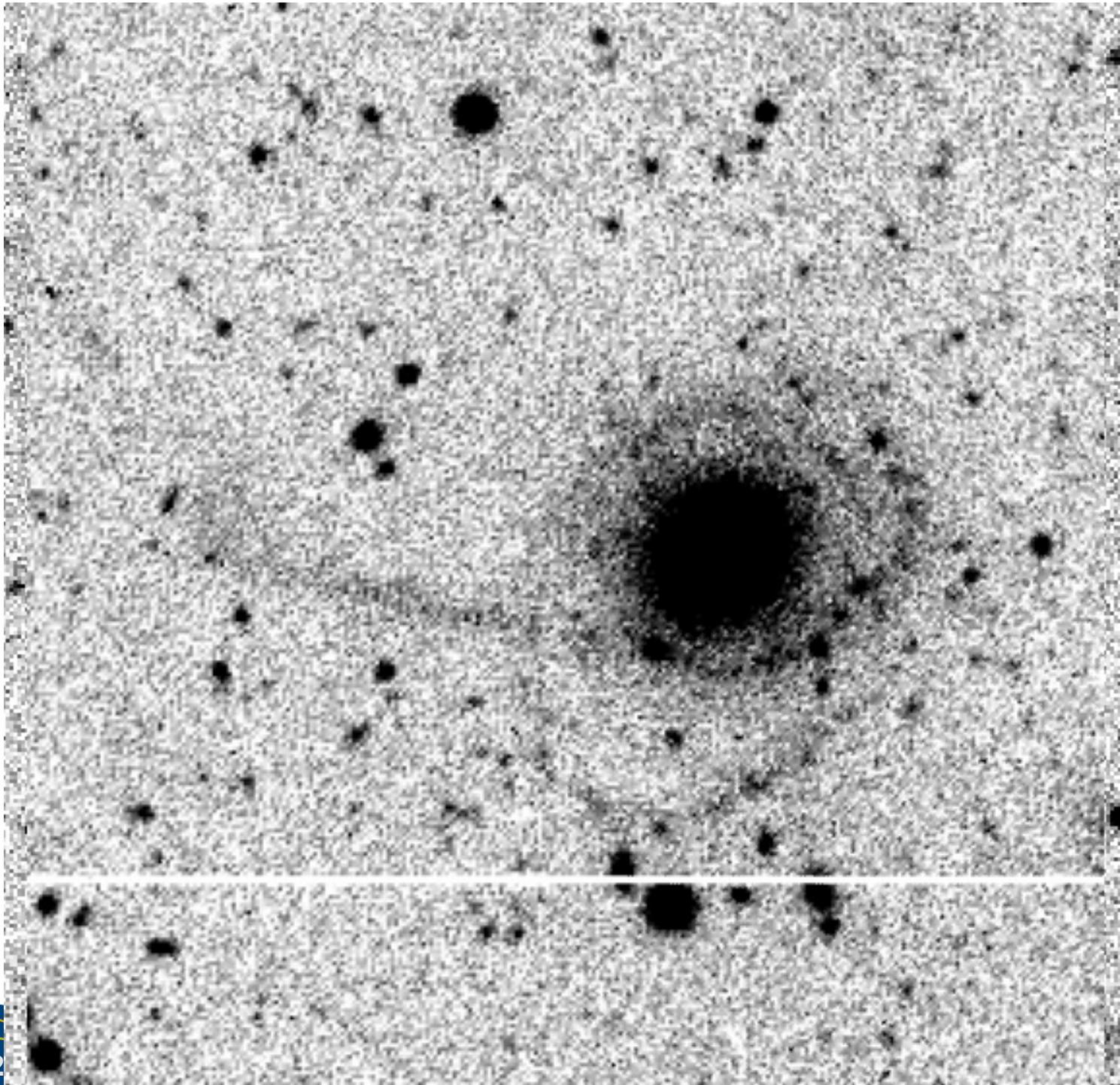






# International Virtual Observatory Alliance





Astronomie: Google für Sterne | Wissen | ZEIT ONLINE - Mozilla Firefox

http://www.zeit.de/2003/37/N-Virt\_Observatorium

Meistbesuchte Seiten: BEOLINGUS: Wörterb..., AIP - Telephone list, W Hauptseite - Wikipedia, Astrophysical Institute ...

ZEIT-Abo | Print-Archiv | Audio/E-Paper | Spiele | Blogs | Schlagzeilen | ZEITmagazin | ZEITCampus | ZEITGeschichte | KinderZEIT | ZEITWissen

**ZEIT ONLINE | WISSEN**

Partnersuche | Stellenmarkt | Immobilien | Autosuche | ZEIT Shop

STARTSEITE POLITIK WIRTSCHAFT MEINUNG GESELLSCHAFT KULTUR **WISSEN** DIGITAL STUDIUM KARRIERE LEBENSART

REISEN AUTO SPORT Gesundheit | Umwelt | Geschichte Anmelden | Registrieren

---

**ASTRONOMIE**

## Google für Sterne

In den Datenbanken der Astronomen schlummern viele neue Entdeckungen. „Virtuelle Observatorien“ sollen helfen, diese Schätze zu heben

Matthias Steinmetz vom Astrophysikalischen Institut Potsdam schwärmt von einer nahen Zukunft, die seiner Zunft ein hübsches neues Werkzeug an die Hand gibt. „Praktisch jedes Land hat jetzt eine Initiative gestartet“, sagt Steinmetz. In Kanada und den USA, Australien, China, Indien, Japan und Südkorea genauso wie in Deutschland, Frankreich, Großbritannien, Italien und Russland. Rund um den Planeten Erde herrscht Euphorie unter den Himmelsguckern. Sie alle freuen sich auf ein riesiges Archiv mit angeschlossener Suchmaschine, eine Art Sternen-Google.

**VON** Stefan Schmitt

**DATUM**

**QUELLE** (c) DIE ZEIT  
04.09.2003 Nr.37

**★ EMPFEHLEN** E-Mail  
verschicken | Bookmarks

**ARTIKEL DRUCKEN**  
Druckversion | PDF

**SCHLAGWORTE**  
Naturwissenschaft | Forschung  
| Forschung | Physik |  
★ Astronomie |||||

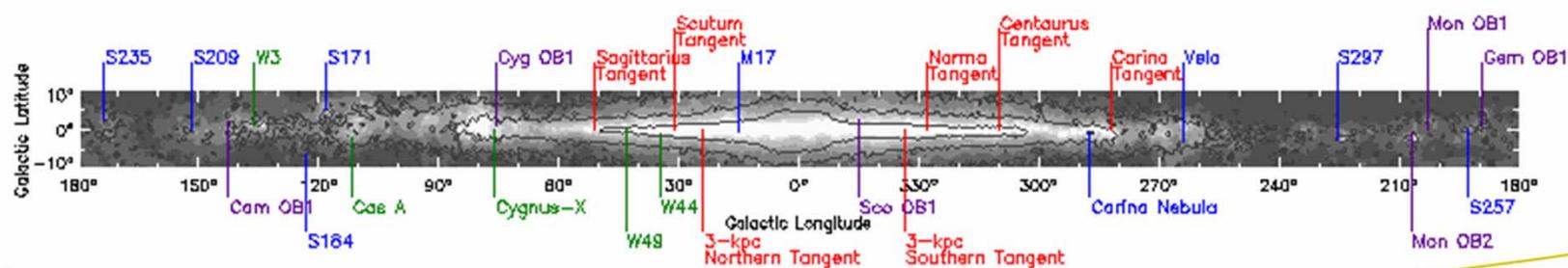
Artikel-Tools präsentiert von:

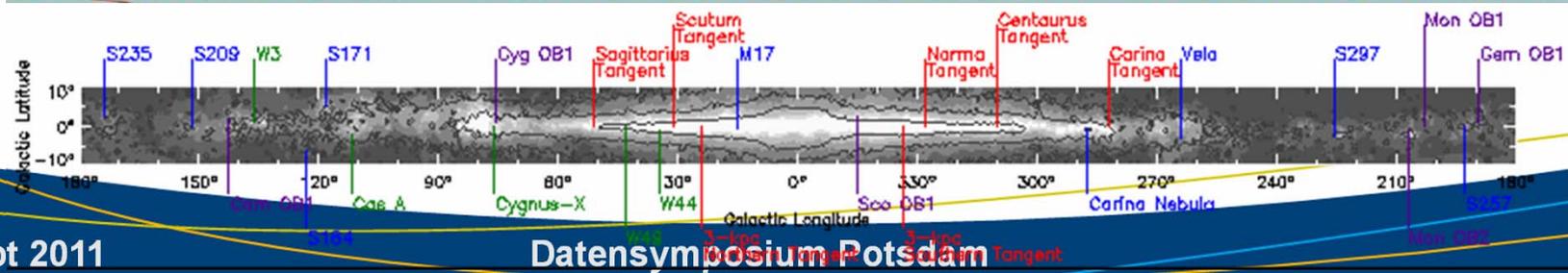
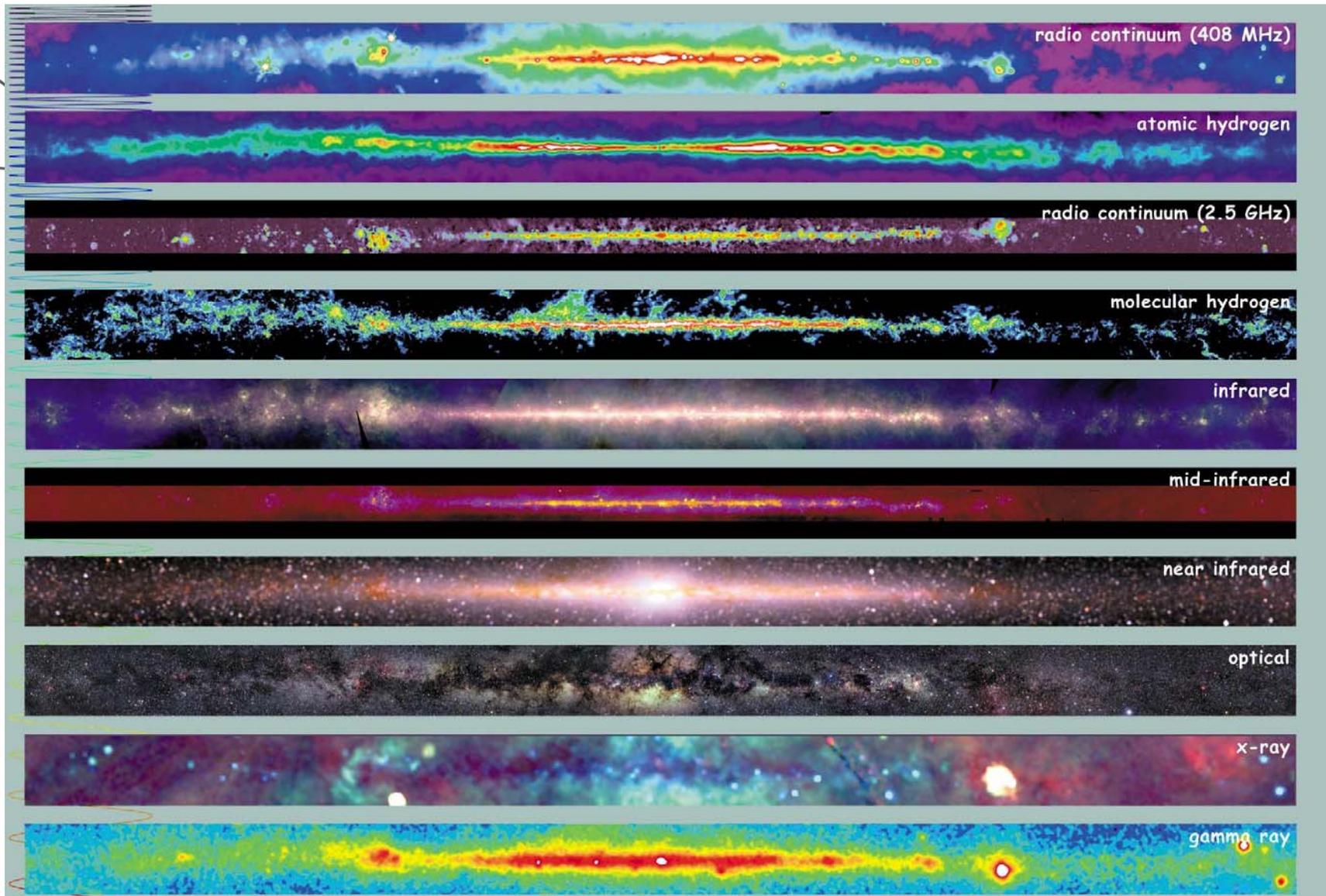
**NEU IM RESSORT**

1. **ADENAUERZEIT (TEIL 2)** Die  
Volksgemeinschaft wird

# Ein panchromatischer Blick auf die Galaxis

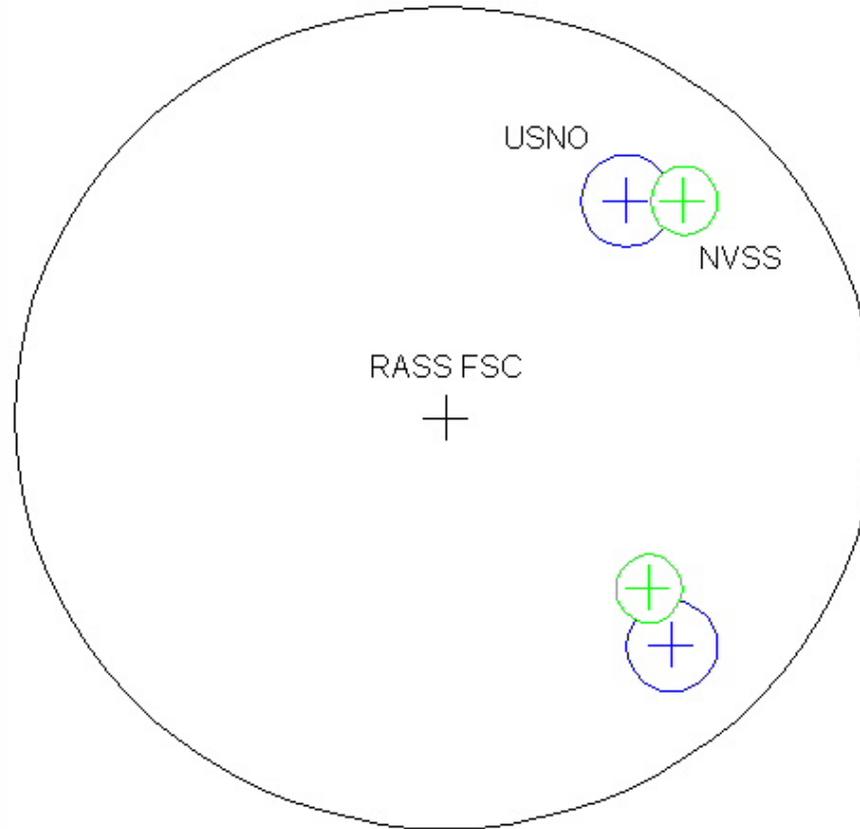
- **Eine digitale Galaxie**
  - Sternkataloge
  - Interstellares Medium und Staub
  - Vergleich mit theoretischen Modellen (Simulationen)







# Das Identifikationsproblem



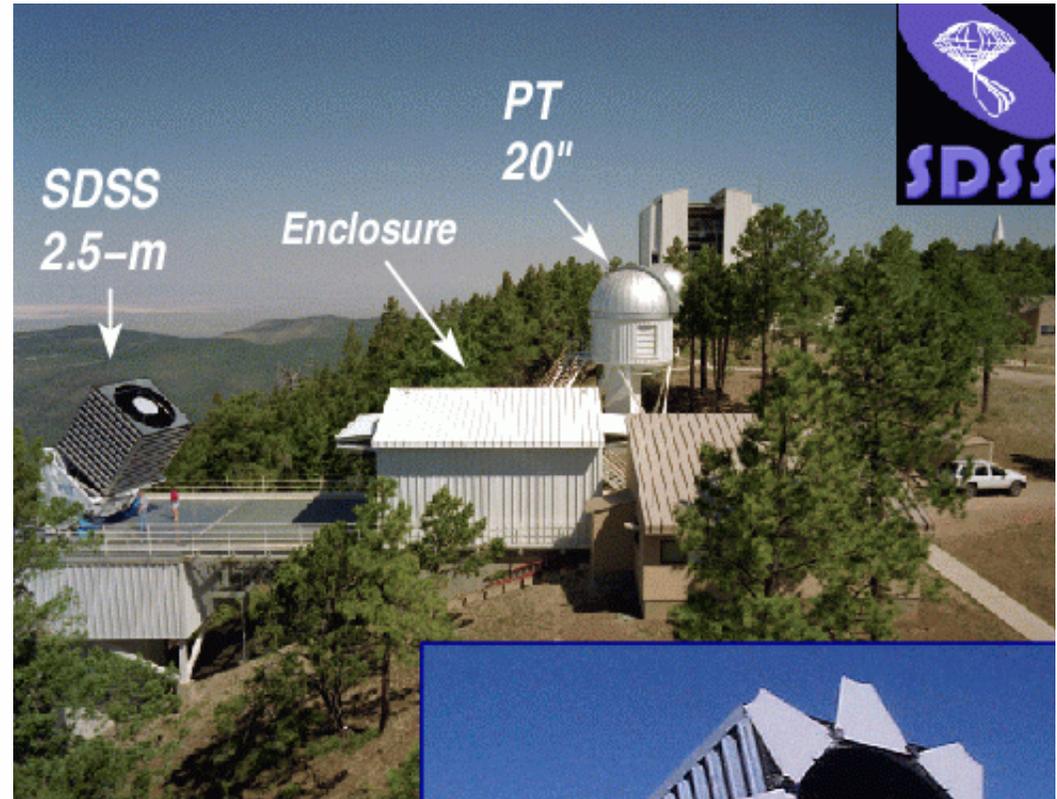


# Beispiel: Sloan Digital Sky Survey

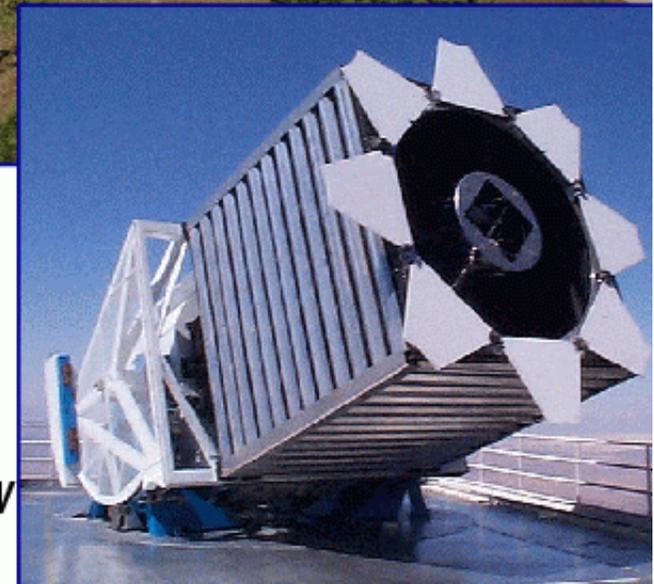
SDSS benutzt ein speziell dafür entwickeltes 2.5-m Telescope

Aufgaben:

- Abbilden
- Spektroskopie



**Apache Point  
Observatory**  
New Mexico, USA  
Lat.  $32^{\circ}46'49.3''$  N  
Long.  $105^{\circ}49'13.5''$  W  
Elev. 2788 m



# Der Sloan Digital Sky Survey



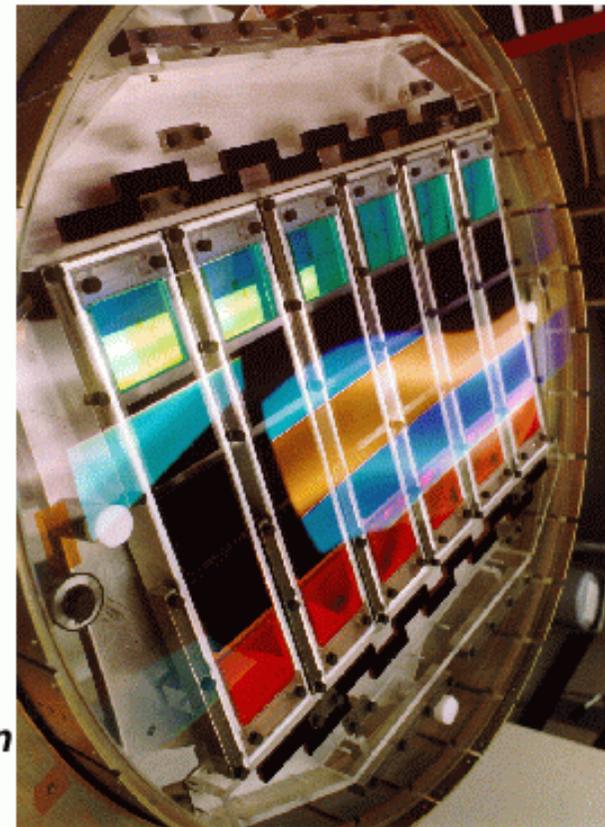
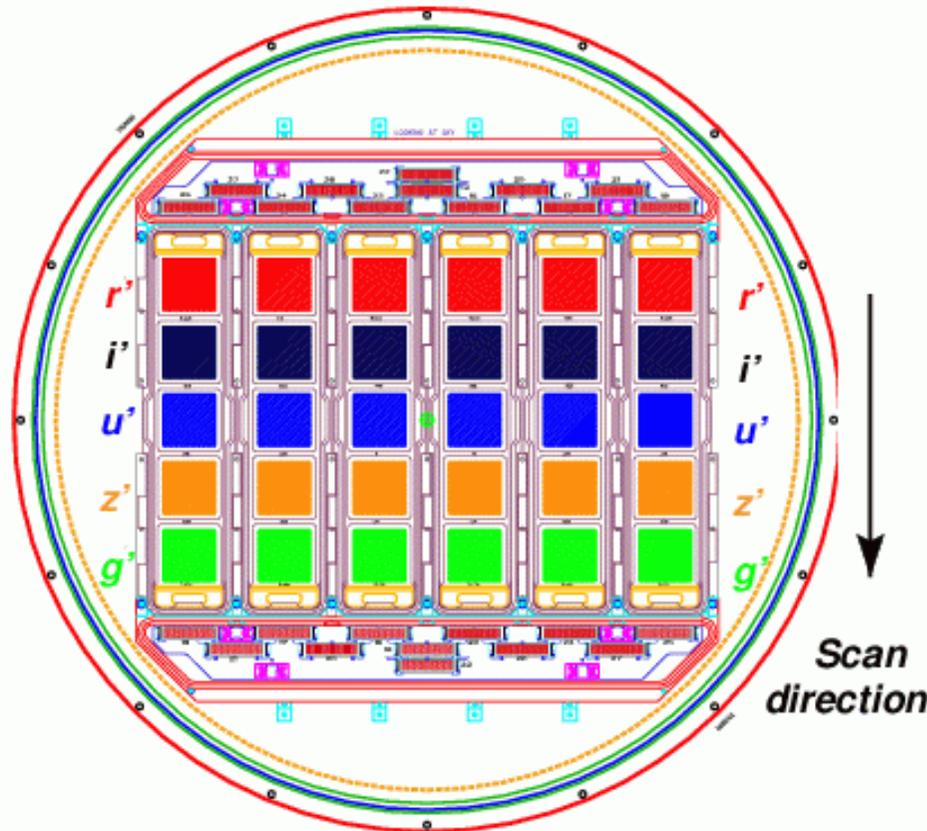
## SDSS Camera

6x5 2K Tek/SITe chips

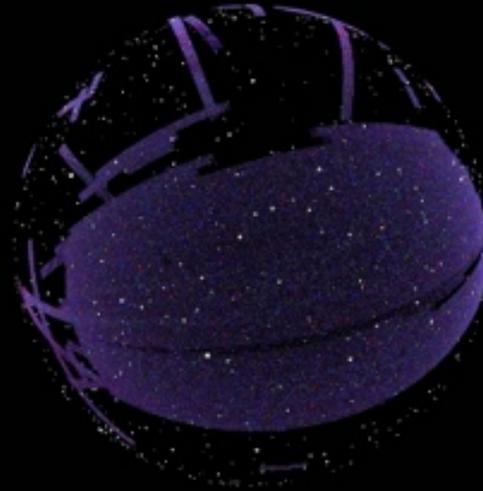
Sidereal scanning rate

15' / sec = 54 sec exposure time

- 30 2048x2048 photometric chips
  - 22 2048x400 astrometric chips
  - 2 2048x400 focus monitor chips
- 0.4" / pixel

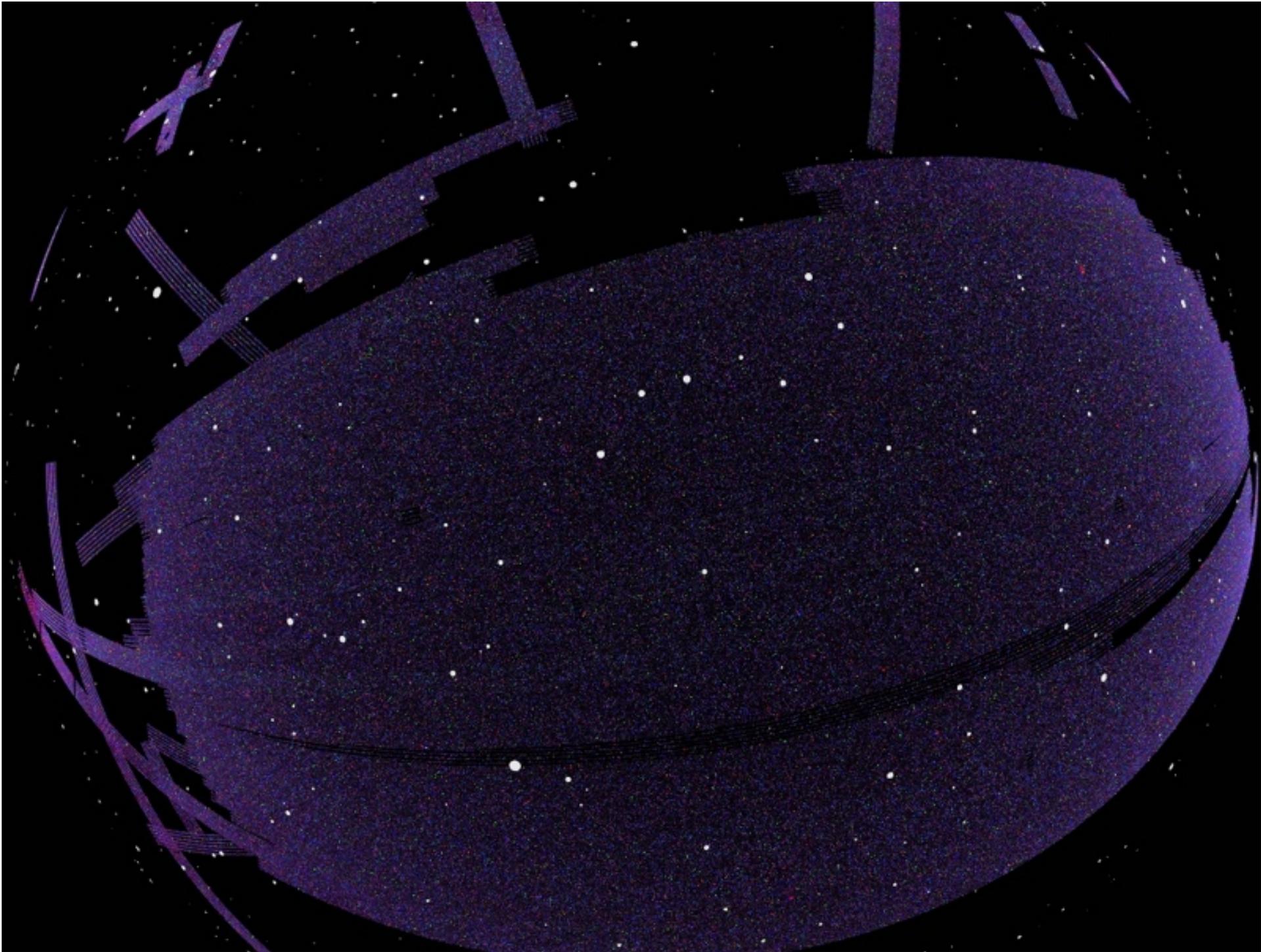


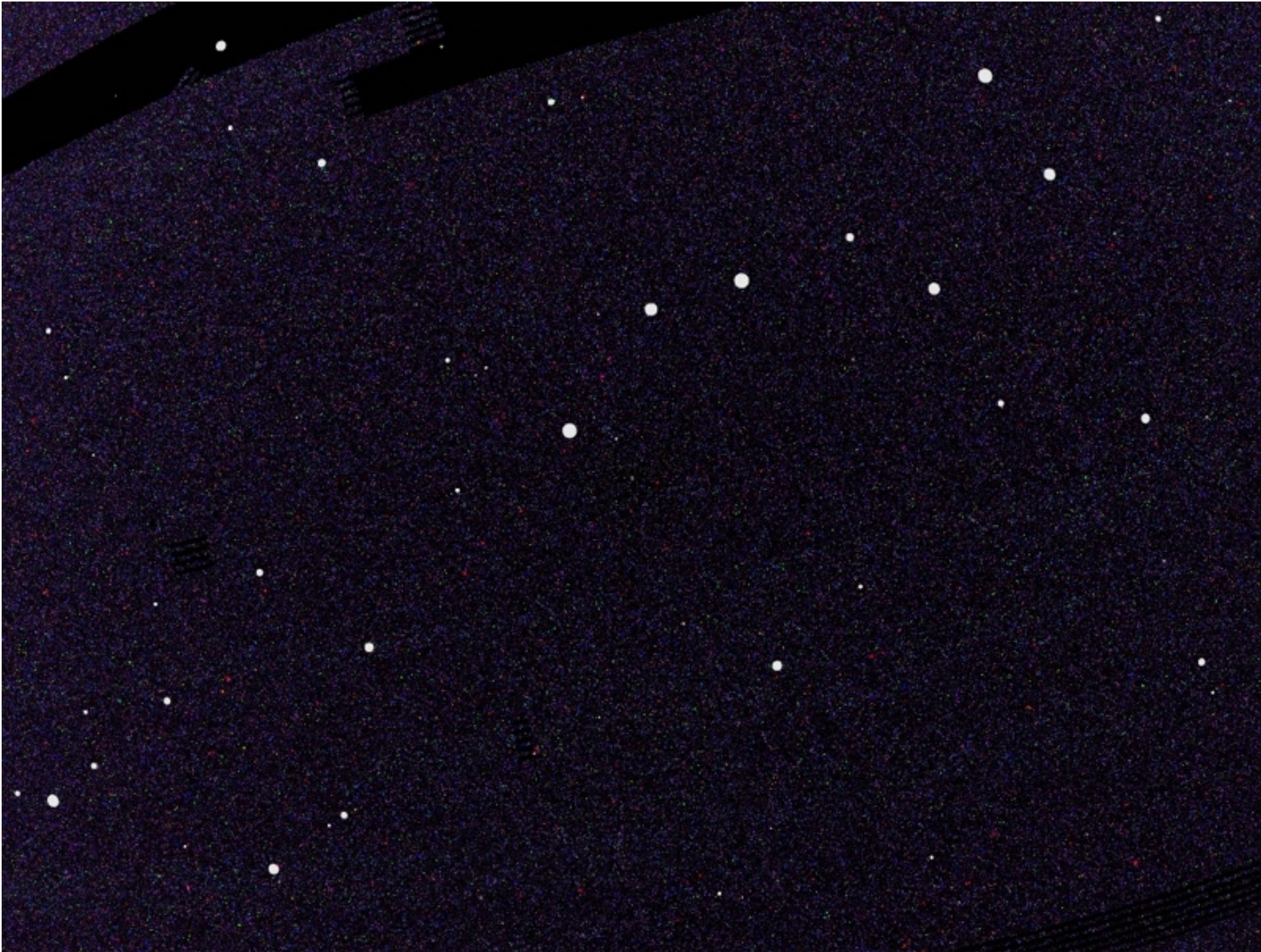
data: Sloan Digital Sky Survey  
and the Bright Star Catalog



visualization: David W. Hogg (NYU)  
with help from Blanton, Finkbeiner,  
Padmanabhan, Schlegel, Wherry

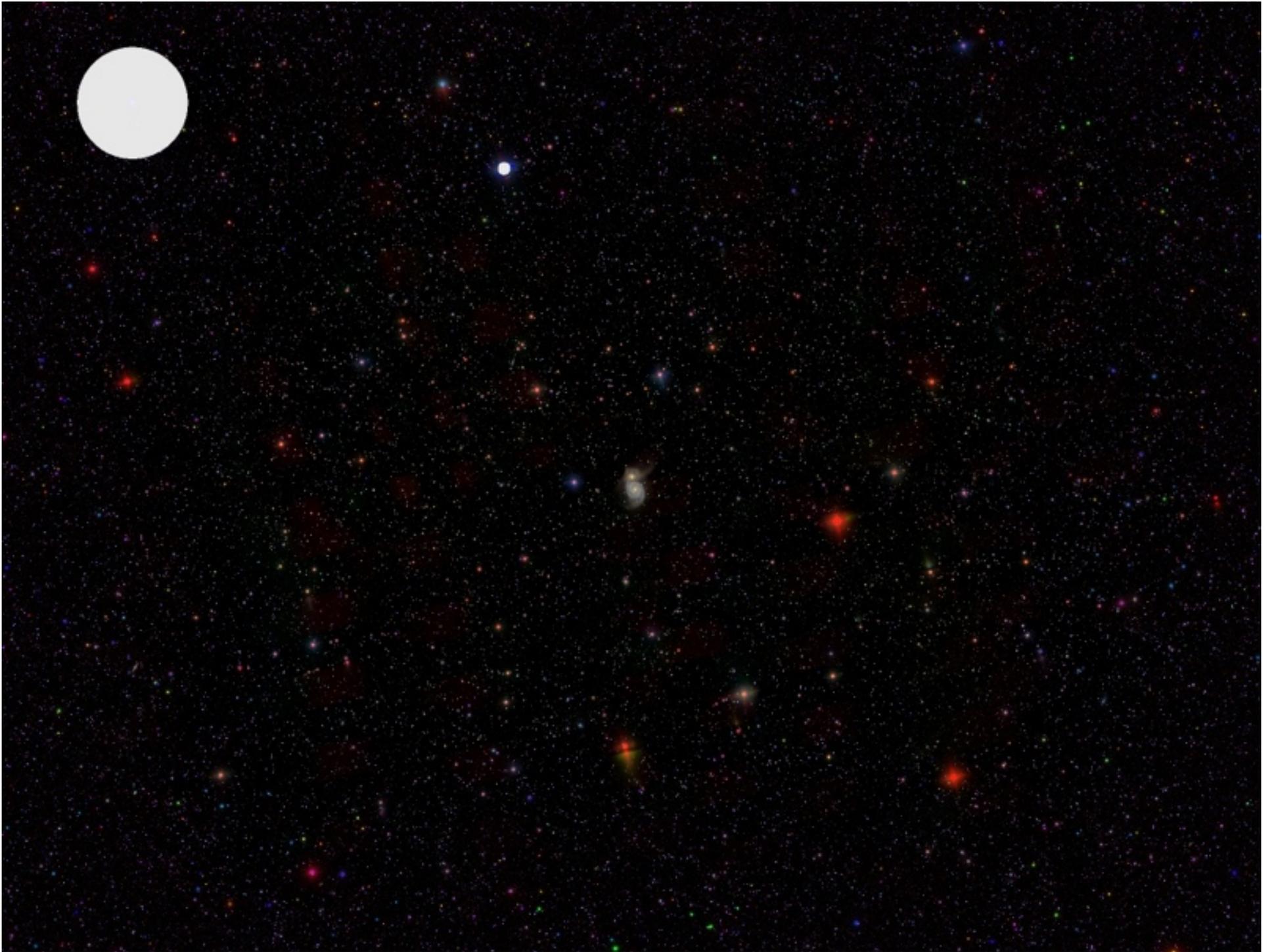


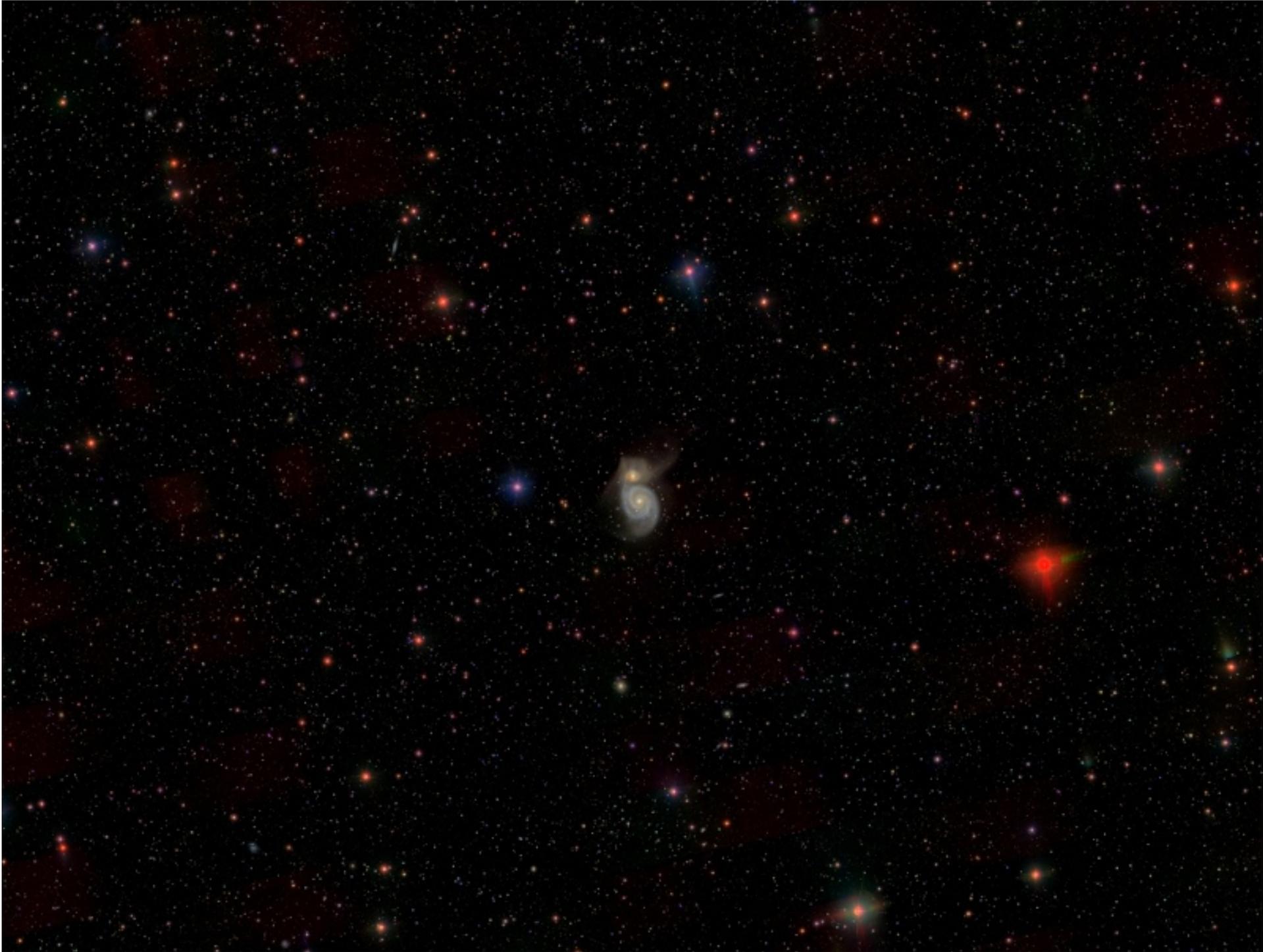




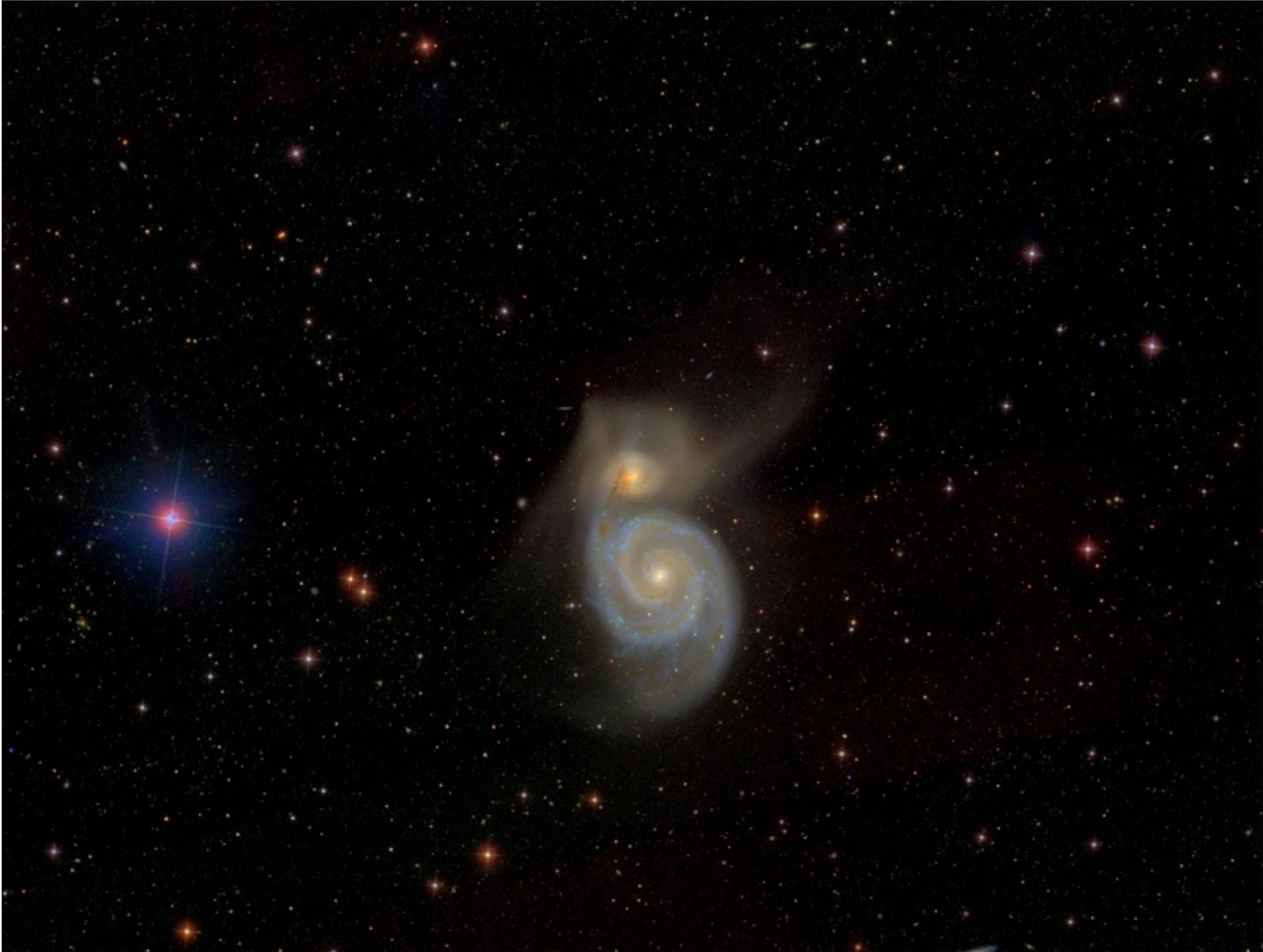




















# Beispiel: Sloan Digital Sky Survey



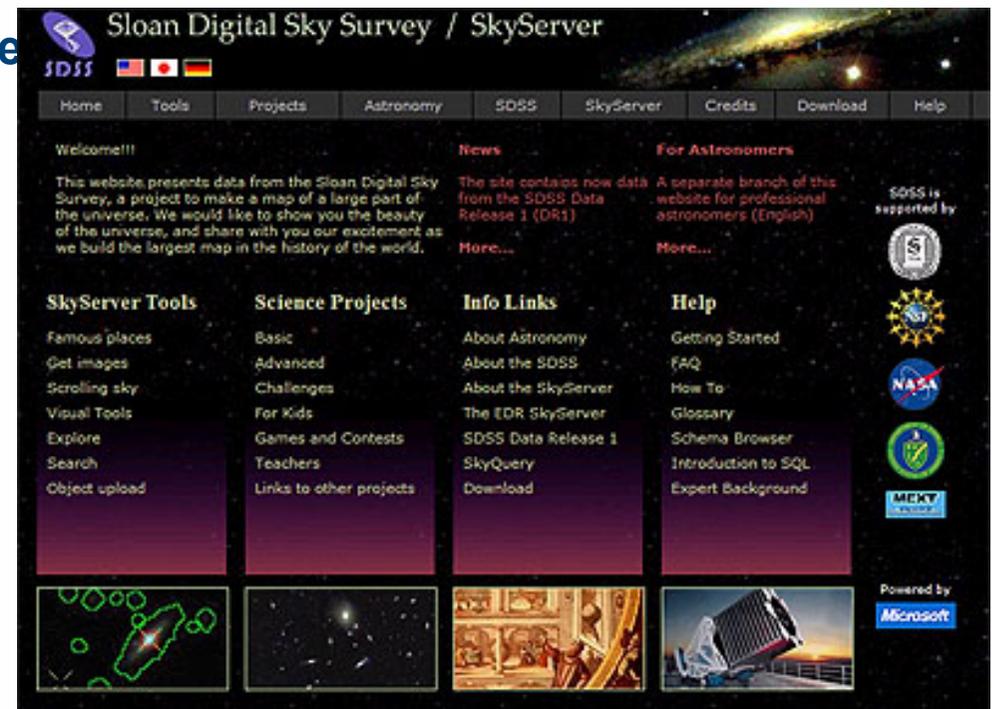
- **Ursprüngliche Durchmusterung wurde am 15.5.2008 beendet.**
- **Final data release (DR7) am 31.10.2008**
- **Final data archive**
  - **Publikationen: Bibliothek der U. Chicago**
  - **Digitales Archiv an der Johns Hopkins University**
  - **Spiegel bei FNAL + JHU + Potsdam**
- **Archiv enthält über 100TB an Daten**
  - **Alle Rohdaten**
  - **Alle reduzierten/kalibrierten Daten**
  - **Alle Versionen der Datenbank (>18TB)**
  - **Komplettes email-Archiv und technische Zeichnungen**
  - **Software code repository**
  - **Telescope sensor stream, IR fisheye camera, etc**





# Öffentliche Nutzung des SkyServer

- Prototyp für Veröffentlichung wiss. Daten
  - 650 Millionen web hits in 8 Jahren
  - 1.000.000 verschiedene Nutzer (vgl 15.000 professionelle Astronomen)
  - >100TB an Rohdaten
  - alles folgt einem Potenzgesetz
- Interaktive Workbench
  - Casjobs/MyDB
  - Eigene Datenbank für ~ 2.500 Power User, keine Zeitbeschränkung
  - Datenspeicherung beim Server, link zur Hauptdatenbank
  - Tools für einfache Datenanalyse (plots, etc)





# Galaxy Zoo

- Baut auf den SkyServer auf (C. Lintott et al)
- mehr als 40 Millionen visuelle Klassifikationen von Galaxien durch die Öffentlichkeit
- Enormes Presseecho (CNN, Times, Washington Post, BBC)
- 200,000 Teilnehmer, Blogs, Gedichte, ....
- Entdeckung eines ungewöhnlichen Himmelsobjekts durch eine Lehrerin in Holland
  - Folgebeobachtung mit Hubble, VLBA, GALEX
- Neue Art, wissenschaftliche Daten zu nutzen



# A peek into the future: the Large Synoptic Survey Telescope

**SDSS:**

a digital color  
snapshot of the  
night sky

**LSST:**

a digital color  
movie of the sky



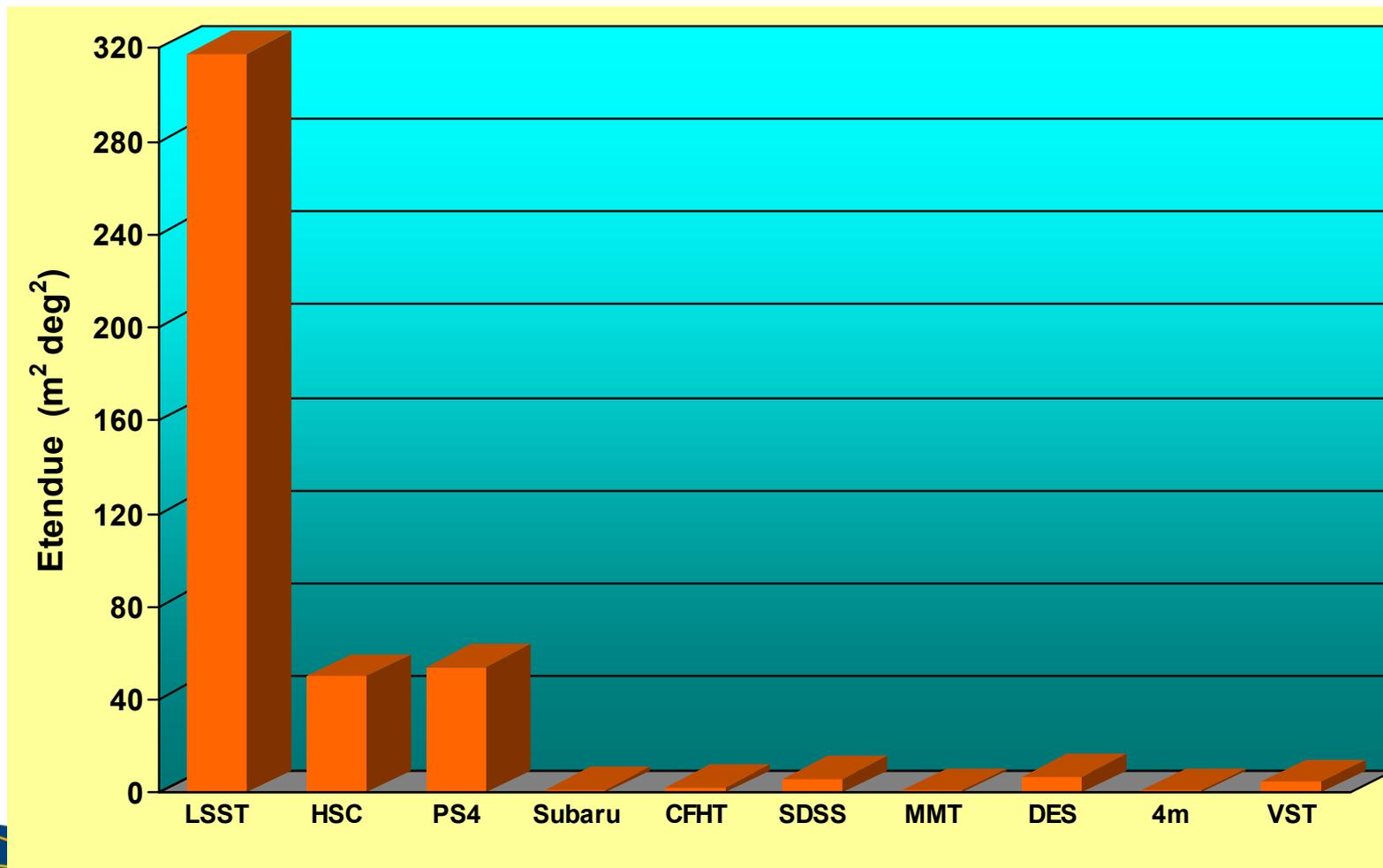
# LSST survey of 20,000 sq deg

- 4 billion galaxies with photo-z
- Time domain:
  - several million asteroids
  - several million supernovae
  - several million lenses
  - new phenomena



# Etendue: Ein Maß für die Mächtigkeit einer Durchmusterung

*Information/time ~ rate of sky coverage ~ Etendue*





# LOFAR: das modernste Radioteleskop der Welt



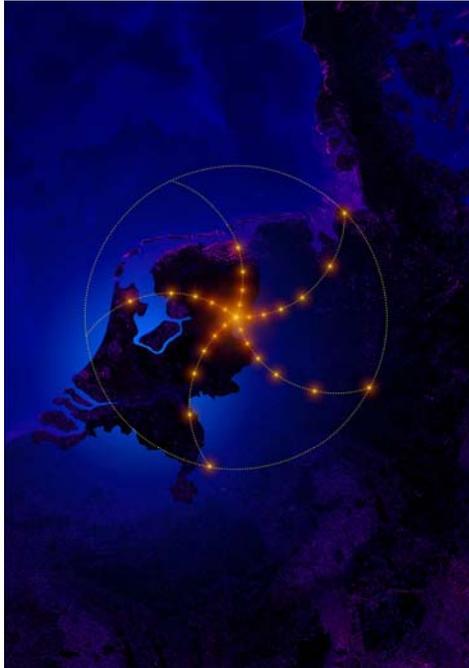


# LOFAR: das modernste Radioteleskop der Welt



© AEROPHOTO EELDE

# Low Frequency Array (LOFAR)



- Teleskop so groß wie die Niederlande + Teile Europas
- Frequenzbereich 30-80, 120-240 MHz
- Basislinie: 2.5-100 km to 1000 km
- ~ 45 Stationen europaweit
- Jede Station ~5Gbit/s sustained
- Keine beweglichen Teile: elektronischer Strahl
- Supercomputer setzt Signal zusammen
- Im Prinzip wird der ganze Himmel erfasst.  
Computerleistung limitiert Zahl der Sichtlinien





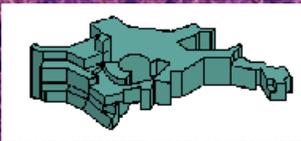
# Massiv parallele Astrophysik

- **Galaktische Struktur einschließlich der lokalen Gruppe**
- **Seltene, sich bewegende Objekte**
- **Gamma Ray Bursts & Supernova bis zu hohen Rotverschiebungen**
- **Gravitationslinsen (stark, schwach, mikro)**
- **Physik der dunklen Materie**
- **Dunkle Energie**
- **Variable Sterne/Galaxien: Akkretionsprozesse**
- **Optische Ausbrüche bis zur 25 mag: unbekanntes Gebiet!**
- **5-Farben photometrischer Survey bis 27m : repräsentatives Volumen des Kosmos**
- **Sonnensystem: NEO, Kometen, ...**

1 Gpc/h

# Millennium Run

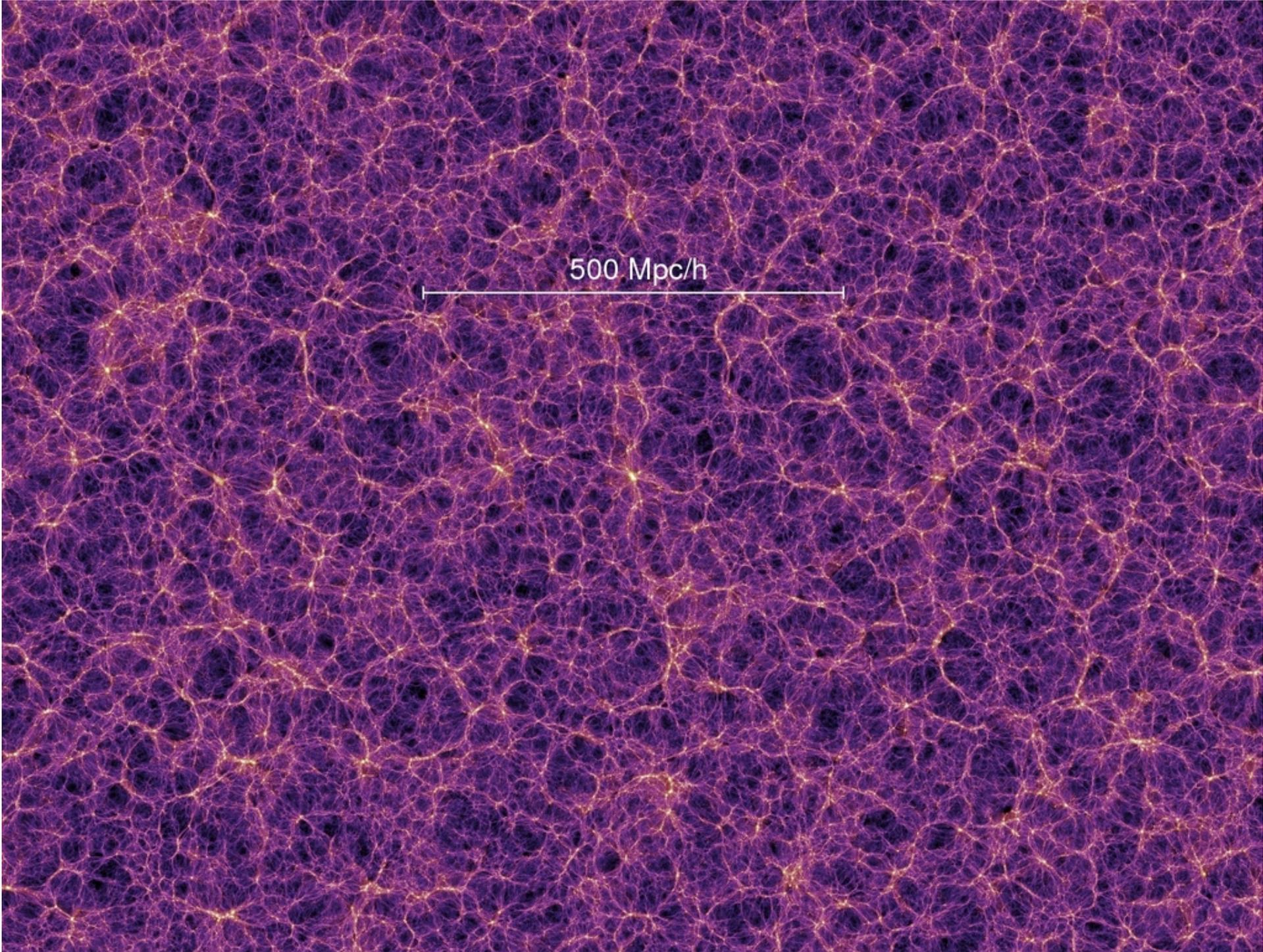
10.077.960.000 particles

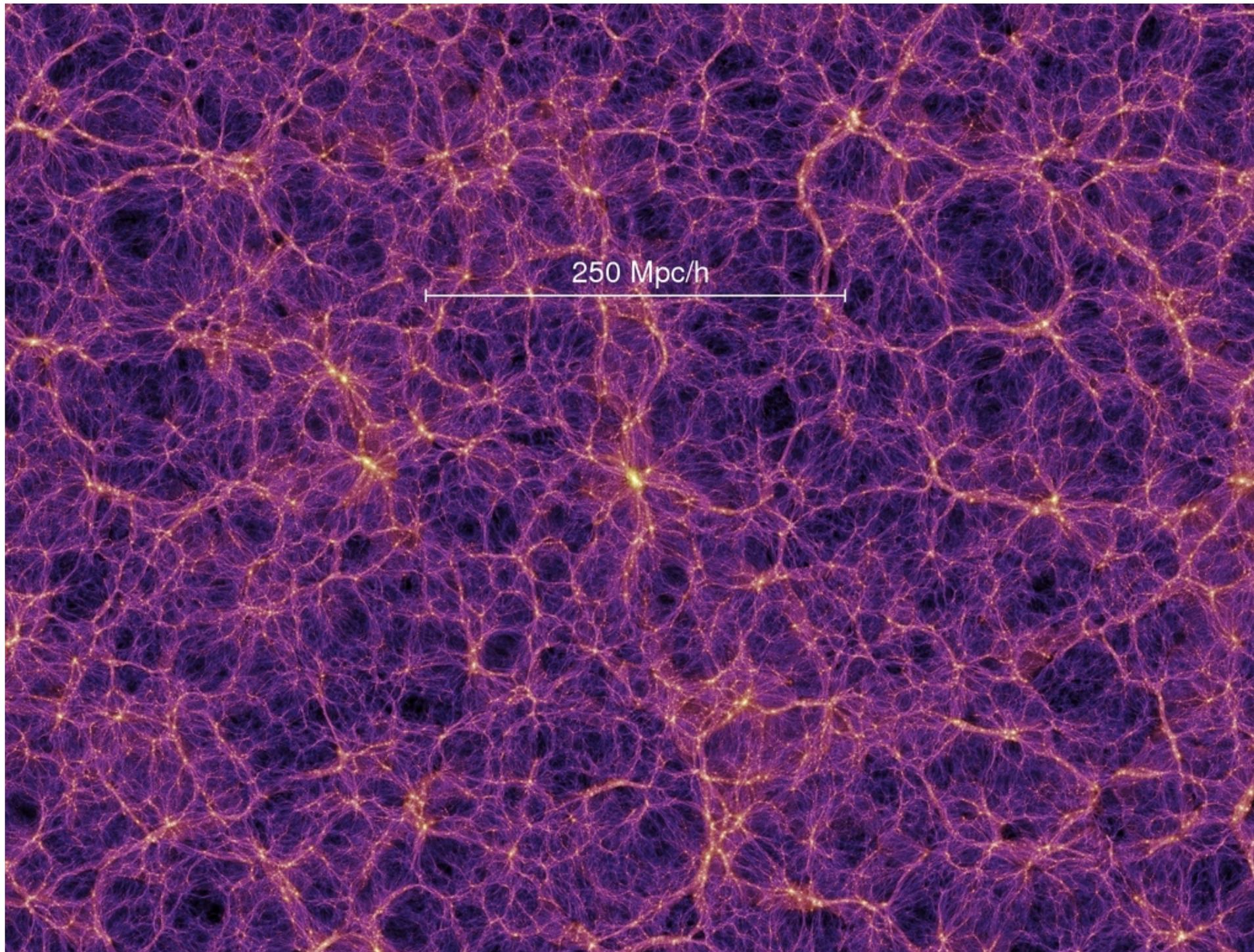


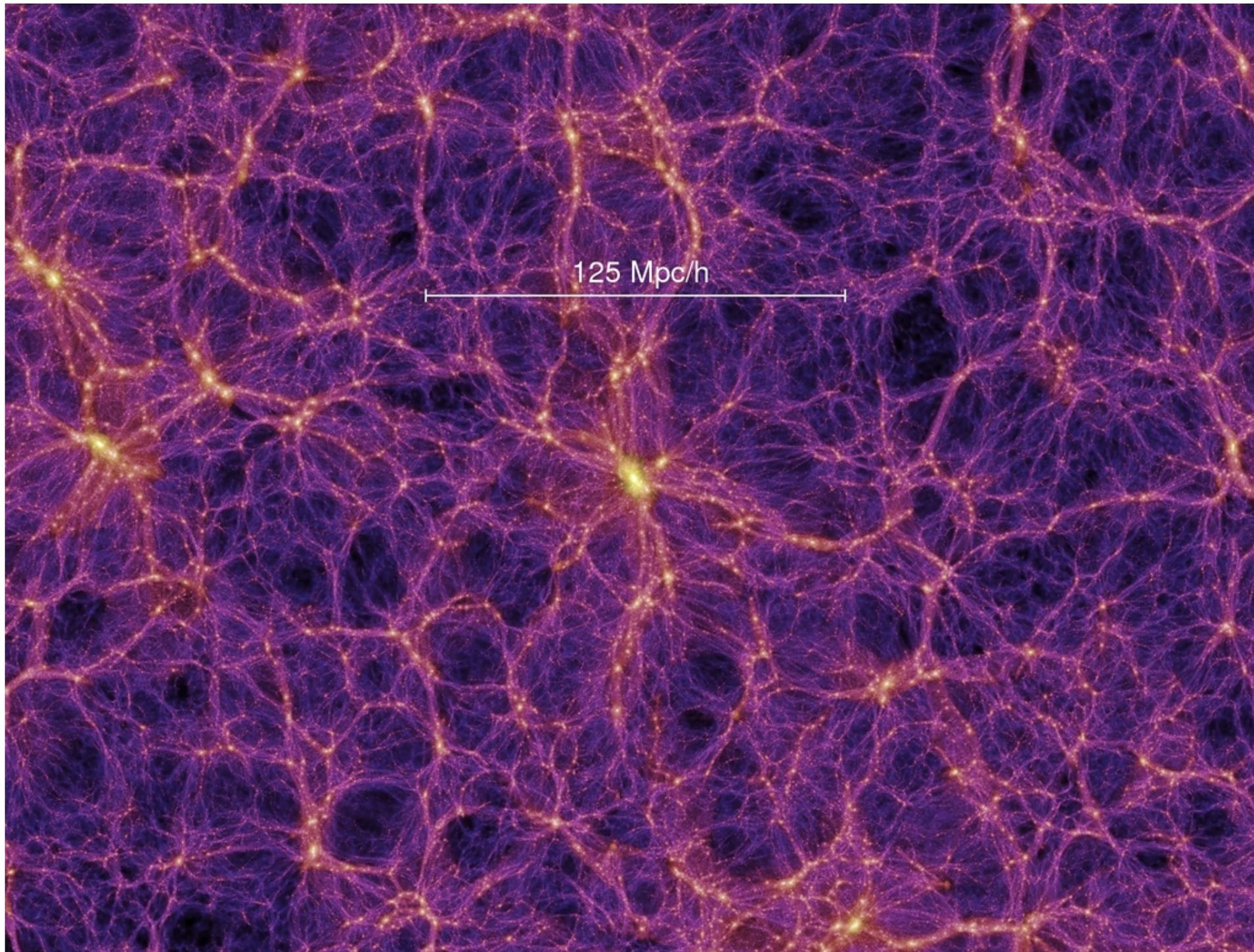
Springel et al. (2004)

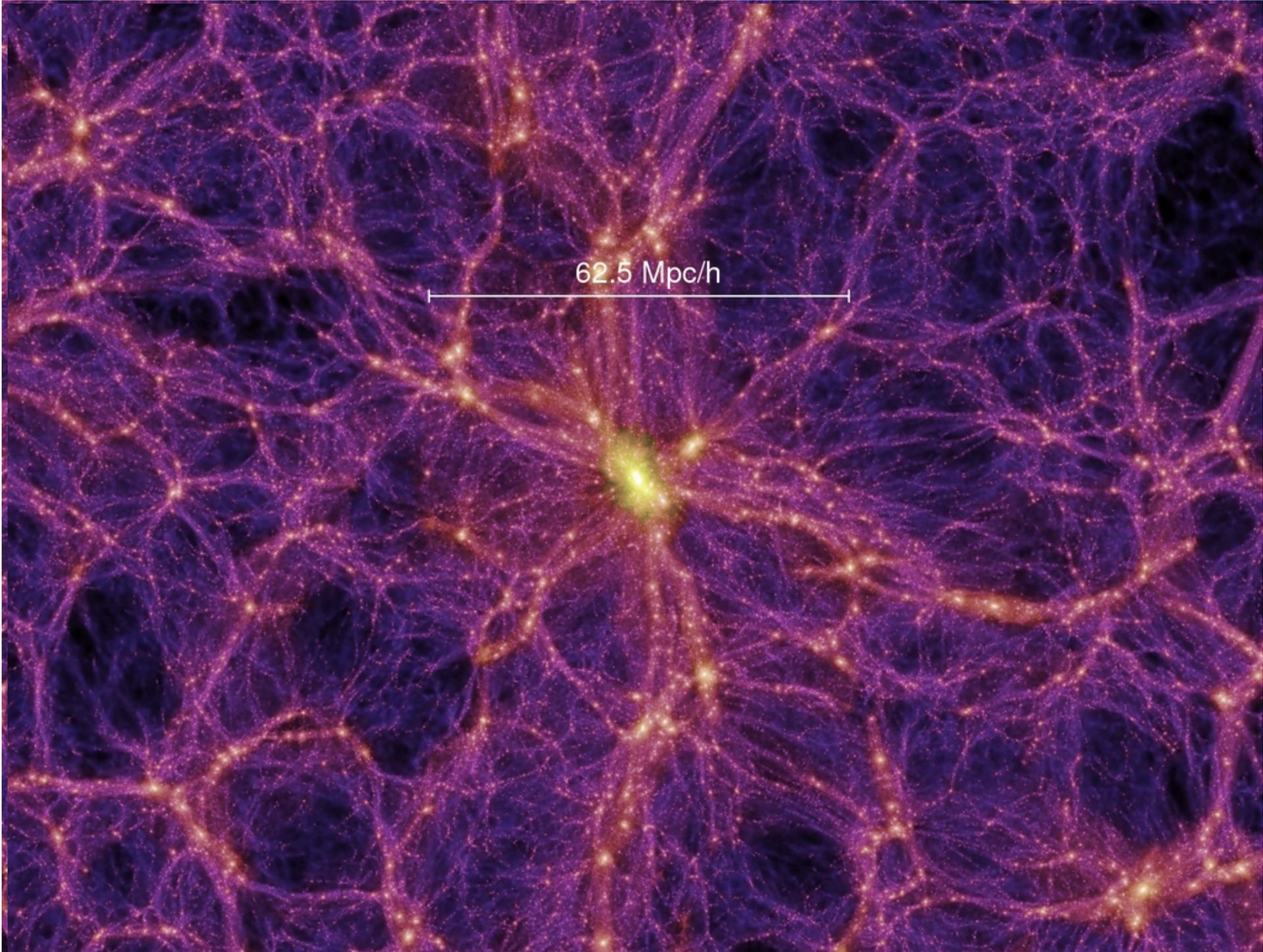


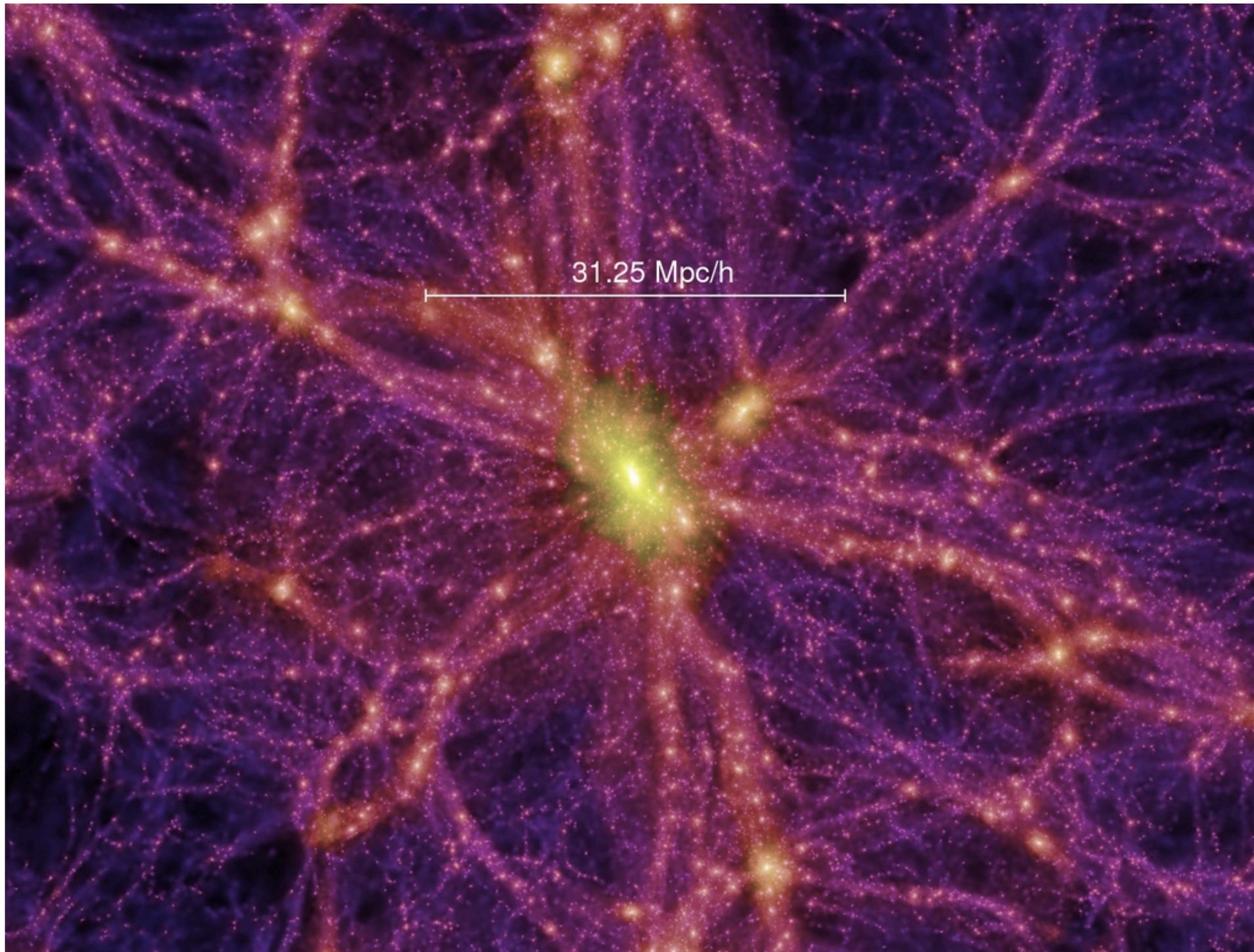
Max-Planck-Institut für  
Astrophysik

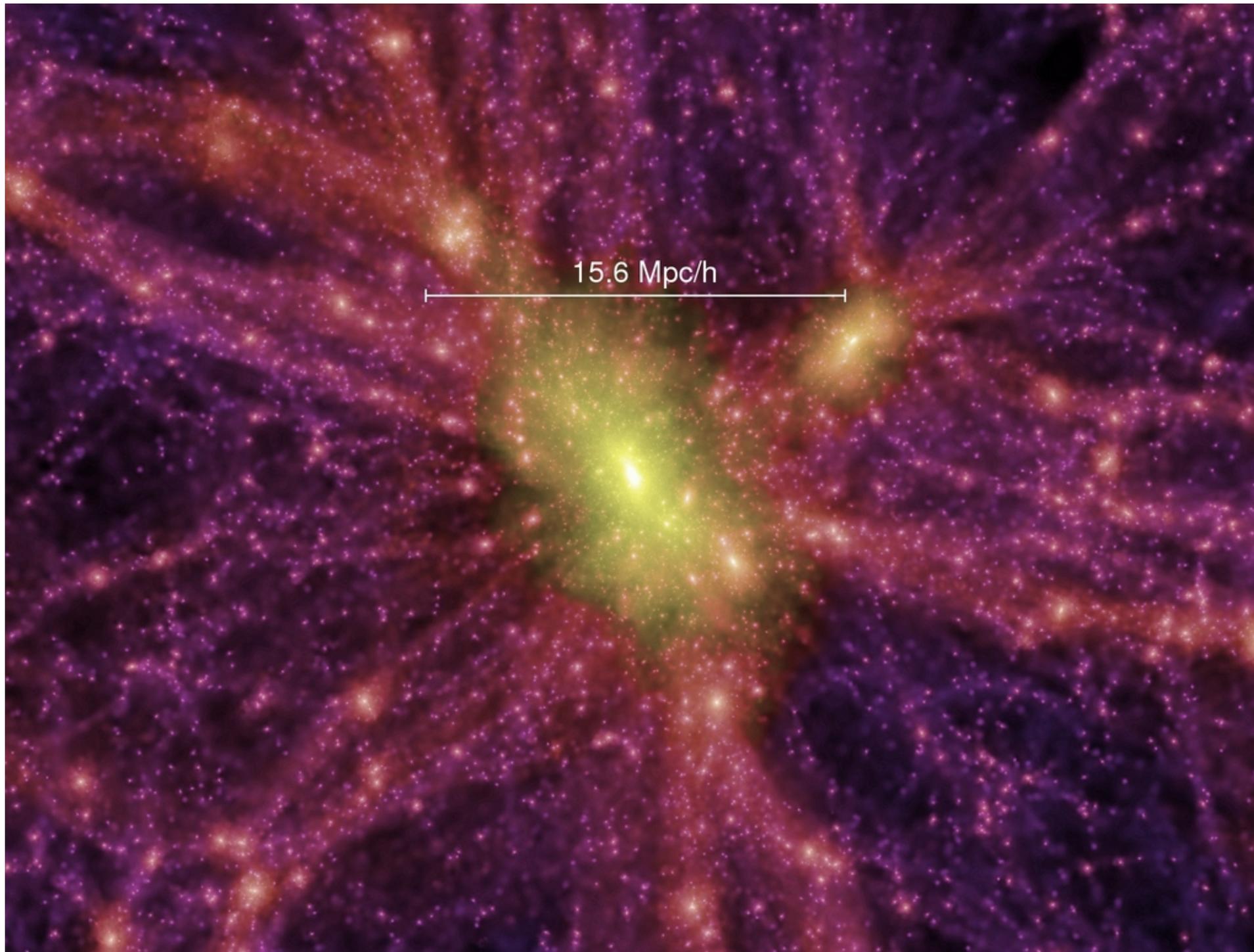


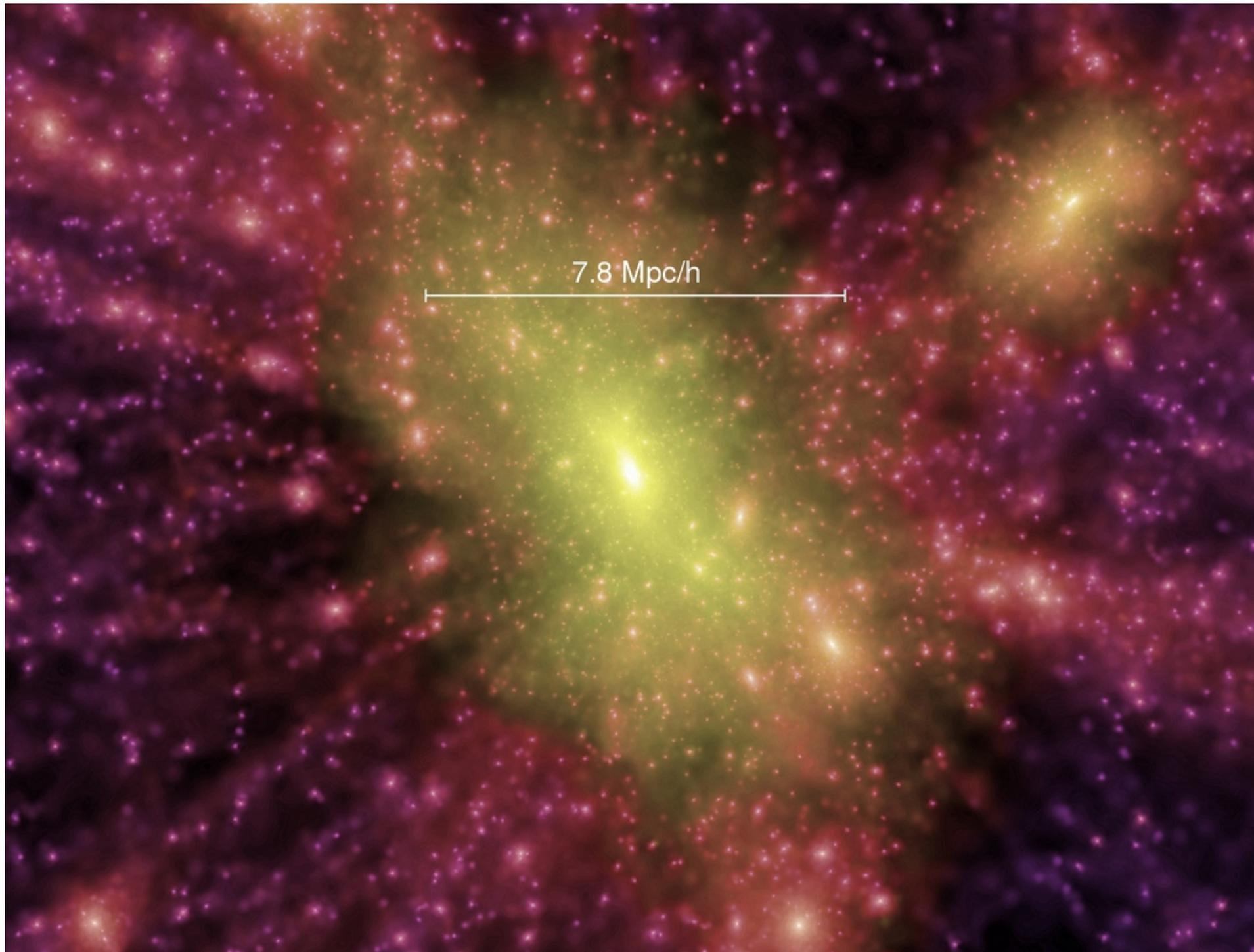


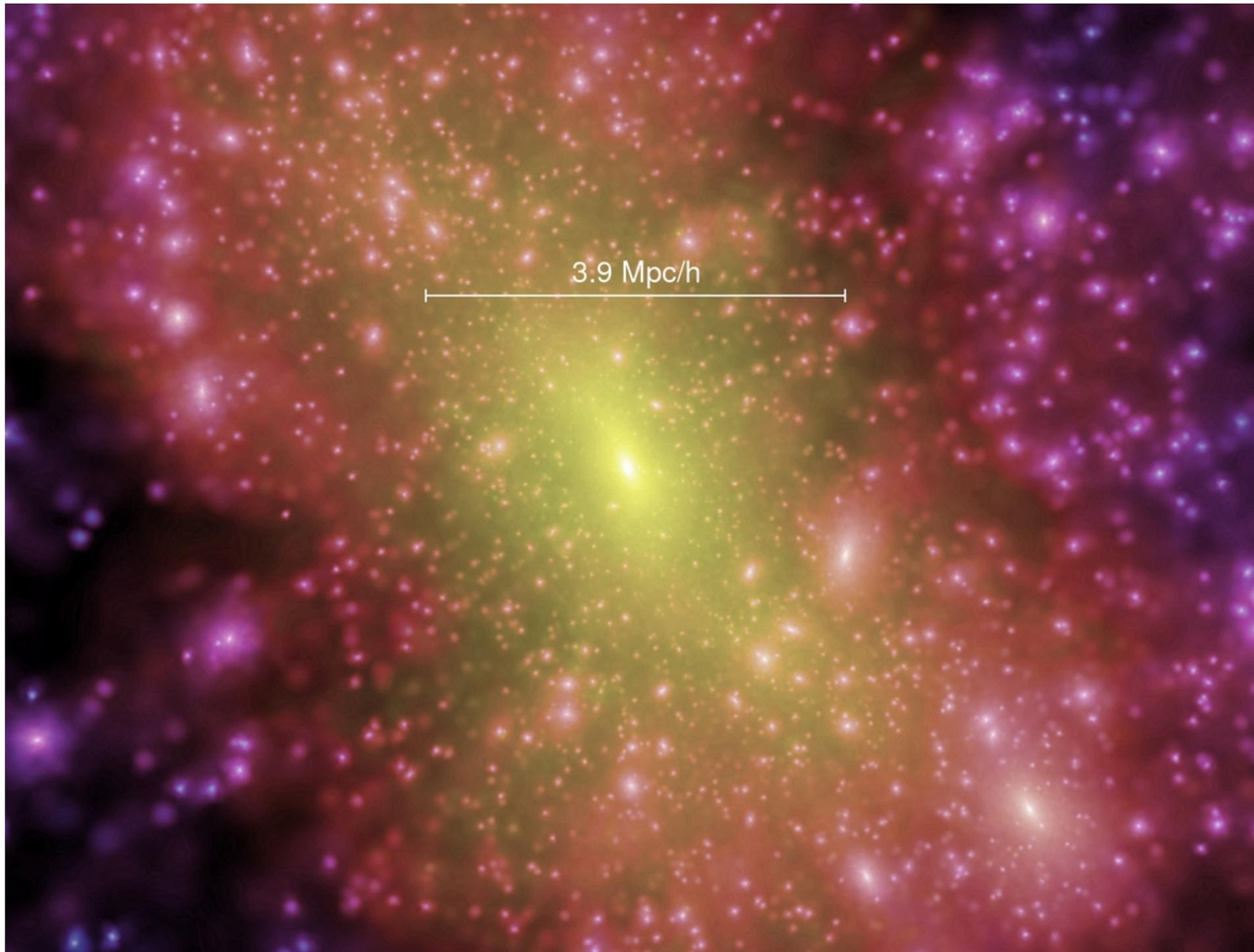














# Die Daten-Herausforderung

- **Petabytes von Daten pro Jahr pro Projekt**
- **Viele Epochen**
- **Viele Frequenzen (optisch, IR, UV, Röntgen, Radio ...)**
- **All Sky**
- **All Sky = 40000 sqdeg, bei einer Auflösung an der atmosphärischen Grenze (0.2'')**:
  - Eine Bitmap des Himmels sind 13 Tpix (pro Farbe und Epoche)



# Herausforderungen und Trends für das Virtuelle Observatorium

- **Änderung in der Astronomie-Demographie**
  - “arm chair astronomer” wird zur Regel  
→ “data miner”
  - “survey builder”
- **Kosten für SDSS: 30% Software, für LSST: 50% Software**
  - Kosten der durchschnittlichen Publikation: \$100.000
  - SDSS: \$100M, 2000 referierte Paper
- **Wir nähern uns der „power wall“**
- **Das Virtuelle Observatorium ist unverzichtbar und unvermeidbar !!!**



# Warum ist Astronomie “special”?

- **Partner: google, Microsoft, Oracle ...**
- **Hohe Attraktivität für die Öffentlichkeit**
- **kein kommerzieller Wert**
  - keine Rechte, freier Zugang für alle
  - ideal um mit Algorithmen zu experimentieren
- **Real und wohl dokumentiert**
  - Multidimensional
  - räumlich, zeitlich
- **Divers und verteilt**
  - Viele verschiedene Instrumente an vielen verschiedenen Orten zu vielen verschiedenen Zeiten
- **There is a lot of it (petabytes)**



# These: Wissenschaftliche Daten öffentlich zugänglich machen

- **Kosten und Risiken**

- Verlangen nach einem höherem Standard, einschließlich Dokumentation, Metadaten etc.
- Kosten der Verfügbarmachung (Server, help desk etc.)
- Gefahr des „being scooped“

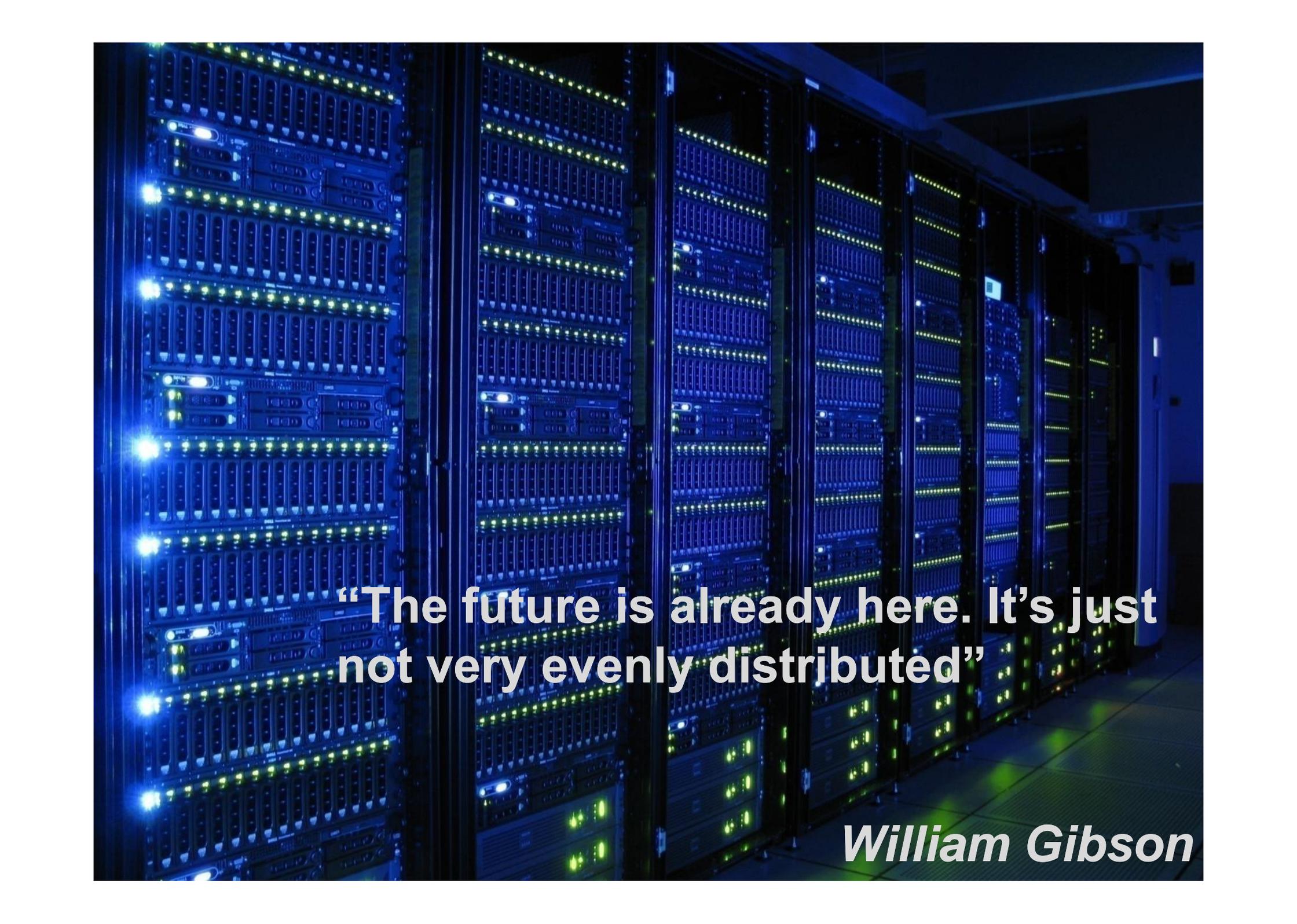
- **Weitere Erwägungen**

- Wer nutzt die Daten, gibt es überhaupt Nutzer?
- Sind sie in der Lage die Daten auch zu nutzen?
- Sicherheitsaspekte (Militärsatelliten in astron. Daten)
- Eigentumsaspekte (J. Gray: Astronomische Daten



# 10 Vorteile offen zugänglicher Daten (nach Ivezic et al)

1. Eine frühe Veröffentlichung von Daten verbessert das Endprodukt
2. Frühe Veröffentlichung von Daten ermöglicht rechtzeitige Analyse
3. Mehr Wissenschaft vom gleichen Datensatz
4. Reproduzierbarkeit wissenschaftlicher Ergebnisse
5. Synergien aus verschiedenen Datensätzen
6. Intradisziplinäre und interdisziplinäre Forschung
7. Manchmal der einzige Weg, knappe Ressourcen zu sichern
8. Mehr Zitierungen und Prestige für das Team
9. Ausbildung und Öffentlichkeitsarbeit
10. Ethische Aspekte und weiterer Impact

A photograph of a server room with multiple rows of server racks. The racks are filled with server units, and the room is dimly lit with blue and green light emanating from the equipment. The perspective is from the end of a long aisle, looking down the center.

**“The future is already here. It’s just not very evenly distributed”**

***William Gibson***