1

Outline – Milestones in NASA Data "Openness"

- Beginning of the Space Age and the desire to open the data to others
- NASA establishes an Archive – National Space Science Data Center
- Enabling Technologies
  - NASA's Internet
  - Online data archives
- Active Archives defined and implemented
- World-Wide-Web
- Open Data Leads to successful new approaches (i.e. Citizen Science)

2

NASA Mission Data Foundation

- January 1958 Explorer-1 launched
- Newell (HQ science lead) - White Paper, **June 17, 1960** - Missions will be defined by NASA and the Science Community
- Centers were viewed as extensions of NASA Headquarters and are given assignments
  - No call for proposals, the early explorer series of spacecraft were assigned to GSFC
  - These were largely "radiation belt monitors" as a follow up of Explorer-1
- GSFC to HQ ~**1961**- Proposal for the IMP series (in support of Apollo) accepted by Headquarters which evolved into a series of about 10 explorer class missions
  - To GSFC, the IMPs were a natural extension of their early explorer missions series
  - IMPs designed to monitor the environment from the Earth to Moon
- NASA data was *only* held by the investigators

3

NASA Data Becomes "Open"

- Newell (HQ science lead) to Silverstein (HQ Spaceflight lead)

MEMORANDUM To Dr. Silverstein 15 Aug 60

SUBJECT: Policy on Release of Reduced Data Acquired from Experiments Carried on NASA Sounding Rockets, Satellites and Space Probes

1. The Space Sciences Steering Committee has considered the subject of the release of reduced data acquired in experiments flown on NASA sounding rockets, satellites, and space probes. The policy which it approved on 9 August and recommended for approval as a NASA policy is attached for your consideration.

2. The Space Sciences Steering Committee in its consideration of a policy kept two points in mind. The first was to give the individual scientific experimenter maximum freedom in the use of data that he obtained from his own experiments. The second was to avoid loss of data by an experimenter's inability to release them in the technical literature, and to specify the conditions under which NASA itself would release the data in advance of any actual case.

3. The policy recommended is practical, and it is simple administratively. It agrees with traditional scientific practice in giving the experimenter the right to his own data, but it does set up provisions for NASA to recover data that might otherwise be lost and for NASA to release the data for general use after the experimenter has finished with them himself.

Homer E. Newell
Deputy Director
Space Flight Programs

4

POLICY FOR RELEASE OF DATA ACQUIRED ON NASA SOUNDING ROCKETS, SATELLITES AND SPACE PROBES

Background

An important goal of the NASA is the encouragement of the widest possible participation in the space sciences program. … In assigning the use of its facilities for space research the NASA endeavors to choose the best experiments and experimenters available…. This paper defines the procedures which NASA will follow …regarding the publication of experimental data obtained in NASA sounding rockets, satellites, space probes or other spacecraft operations.

Basis of Policy

… In the usual case, an agreement that will assign the responsibility for the analysis and publication of the data to the scientist will exist between the NASA and the scientist. In the period before the publication of the results of the experiment it may happen that another scientist may wish to have access to the experimental data in its preliminary form. This is generally arranged informally by a request made directly to the experimenter. Occasions may arise when it will be necessary for NASA to release the results of experimental investigations made under its sponsorship, without waiting for publication by the cognizant scientist. It is expected that this will happen only when the responsible scientist is unable to meet his commitments within a reasonable time.

In many instances the data acquired from space vehicles, like the observations made in astronomy, will have cumulative value. Whenever this is the case, it is desirable that such data be made available for the general use of the scientific community. The original data received from a satellite or space probe is meaningless until the various channels are separated or demultiplexed, the calibration of the instrumentation is introduced, and the data put in chronological order. Provisions for these operations must be included in the arrangements made for supporting the applicable experiments.

5

Policy

The original data recordings received from a satellite or probe will be stored at the cognizant NASA center. Secondary records containing reduced data will be prepared under the cognizance of the responsible NASA center. On occasion this operation may take place at contractor's or experimenter's establishments. In the secondary records all extraneous material and all confusion of data will be fully removed possible. The records will be prepared by the NASA activity or contractor with the cooperation of the experimenter to satisfy the experimenter's and NASA's needs for reduced data.

The experimenter will be given whatever original or reduced data from his experiment that his agreement with NASA calls for as soon as possible after they are received. The cognizant NASA center will retain copies of the reduced data furnished the experimenter. The experimenter will have sole use of the reduced data for study for a period previously decided upon between him and the cognizant research center, with the approval of Headquarters.

No fixed time is set for the period of exclusive use of reduced data by the experimenter. Each case will be considered individually, and no general rule is established.

If, during the period in which the experimenter has not reported the results of his experiment, information on the progress of the experiment or portions of the reduced data are requested by another scientist for designing space science experiments, or for the preparation of a technical paper, informal arrangements for the release of the needed data should be made by the experimenter concerned. Both the cognizant NASA center and the NASA Headquarters should be informed of these arrangements.

At the end of the agreed upon period, during which the experimenter will have the opportunity to analyze and report on the reduced data, the reduced data may be released for general use if such a course seems best for developing the space sciences. An extension of the period of sole use requested by the experimenter will require the approval of Headquarters.

All releases of the reduced data by the NASA centers to other than the original experimenter must be approved by Headquarters.

6

Definition of Terms
Used in the Policy Statement

Original data records

Those records made by the various telemetering and/or tracking stations as part of the basic field operations. Those recordings will generally require editing for the removal of useless segments, playback to introduce calibration factors, correlation with related recordings and the application of specialized processing techniques to recover weak signals from noisy recordings.

Reduced data records

Those records prepared from the original data records by editing, introduction of calibration factors and inter-record correlations. These will contain a minimum of extraneous information and/or noise and will generally present the value of the physical quantity measured as a function of time and position. It is from these records, or the tabulations or graphs prepared directly therefrom that the responsible scientist will form his analysis and conclusions.

Analyzed data

Those data which have been reviewed and correlated by either the original experimenter or another scientist to form the experimental basis for a technical paper submitted for publication to either NASA or a scientific journal.

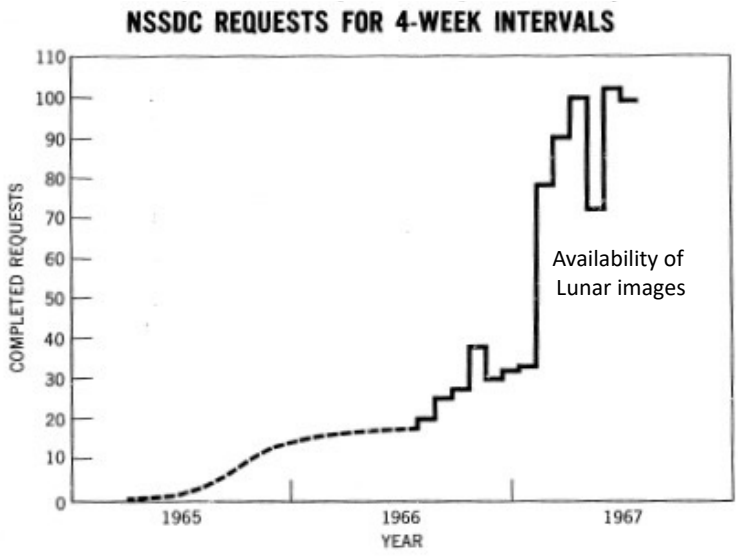
7

Establishing the NSSDC

- In 1966 Newell establishes the National Space Science Data Center at the Goddard Space Flight Center
- NSSDC to enable US investigators to archive their "reduced" data at a central location providing a mechanism to help the investigators distribute their data to only US scientists upon request
  - PIs would provide NSSDC with microfilm, photographic images, and data on digital tapes (256, 800, and "high density" 1600 bits per inch)
- In Sept 1968 Truve (Chair of NAS Geophysical Research Board) requests that NAS transfer its designation as a World Data Center to NASA
- Newell assigned that role to the NSSDC
  - Officially making NSSDC held data available to foreign scientists upon request

8

Operation of the National Space Science Data Center - 1967



9

Volume of Data at National Space Science Data Center - 1967

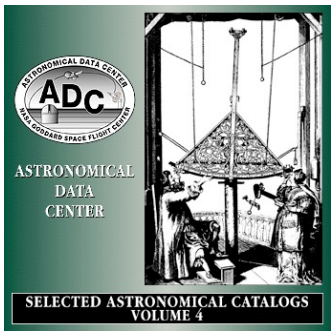
Present (August 1967) and Projected, Annually (January 1970)

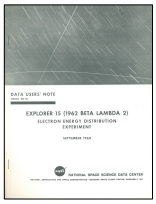
MEDIUM	ON HAND	INCOMING
SHEETS AND BOUND VOLUMES, SHEETS	175,000	100,000
ROLL CHARTS, LINEAR FEET	360,000	150,000
DIGITAL MAGNETIC TAPES, 1/2"x2400'	291	10,000
ANALOG MAGNETIC TAPES, 1/2"x2400'	1,035	0
MICROFILM, 100-FT ROLLS	7,800	2,000
PHOTOGRAPHIC FILMS:		
9-1/2" WIDTH, LINEAR FEET	14,000	4,000
70-mm WIDTH, LINEAR FEET	33,200	12,000
4- x 5-INCH, EACH	2,100	1,000
16- x 20-INCH, EACH	93	--
20- x 24-INCH, EACH	2,200	800
PHOTOGRAPHIC PRINTS:		
8- x 10-INCH	600	500
11- x 14-INCH	200	300
16- x 20-INCH	93	--
20- x 24-INCH	2,200	800

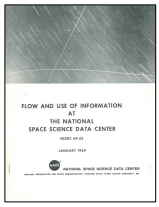
10

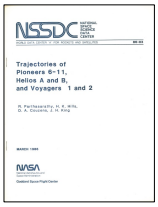
Early Responsibilities of NSSDC

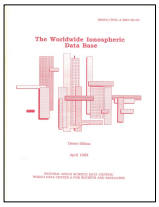
- Archiving satellite data
- Maintaining sounding rocket database
- SPACEWARN bulletin – announcement of all rocket launches
- Assigning international spacecraft IDs
- Publication of data catalogs and supporting documentation
- Safeguarding magnetic tape, microform, and hardcopy media
- Developing and promoting data standards
- Make data available to scientists and the general public





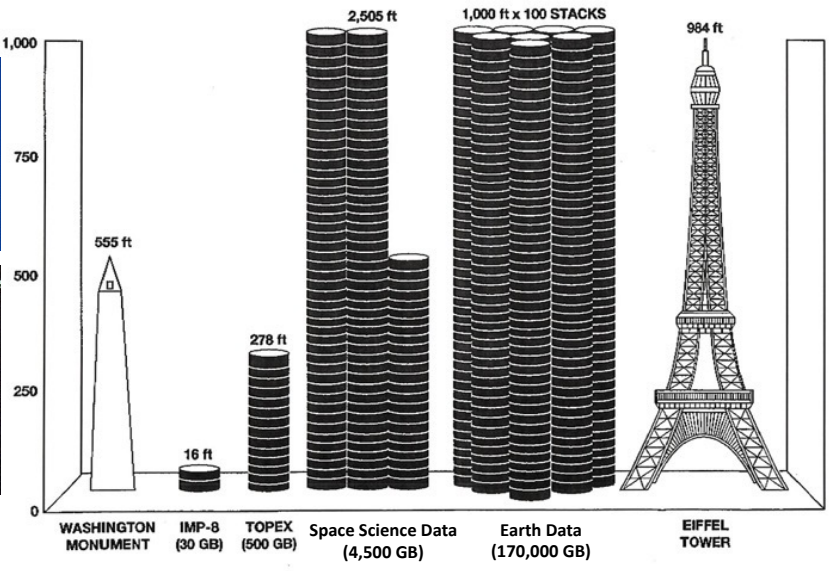




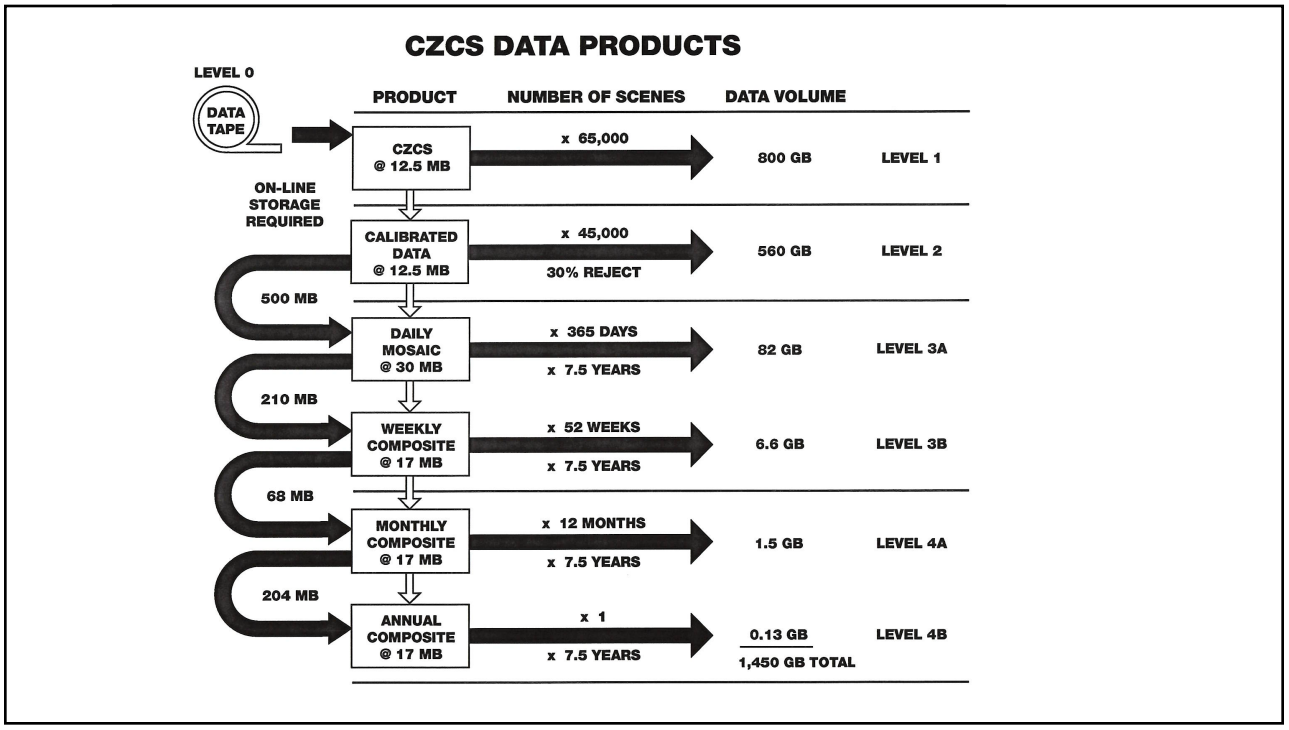


11

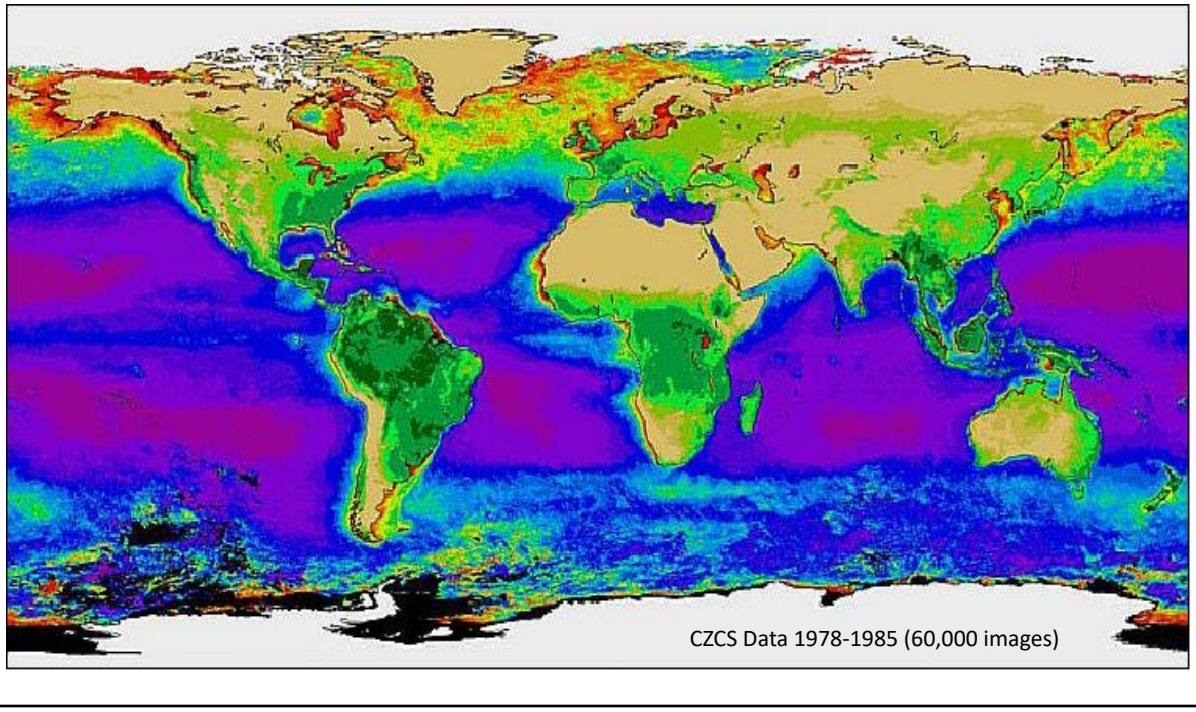
DATA VOLUMES IN MAGNETIC TAPES/YEAR



12

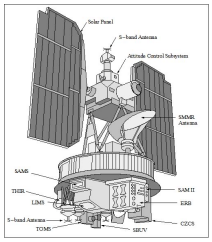

CZCS DATA PRODUCTS
LEVEL 0
DATA TAPE
PRODUCT
NUMBER OF SCENES
DATA VOLUME
CZCS @ 12.5 MB
x 65,000
800 GB
LEVEL 1
ON-LINE STORAGE REQUIRED
CALIBRATED DATA @ 12.5 MB
x 45,000
30% REJECT
560 GB
LEVEL 2
500 MB
DAILY MOSAIC @ 30 MB
x 365 DAYS
x 7.5 YEARS
82 GB
LEVEL 3A
210 MB
WEEKLY COMPOSITE @ 17 MB
x 52 WEEKS
x 7.5 YEARS
6.6 GB
LEVEL 3B
68 MB
MONTHLY COMPOSITE @ 17 MB
x 12 MONTHS
x 7.5 YEARS
1.5 GB
LEVEL 4A
204 MB
ANNUAL COMPOSITE @ 17 MB
x 1
x 7.5 YEARS
0.13 GB
1,450 GB TOTAL
LEVEL 4B

13


CZCS Data 1978-1985 (60,000 images)

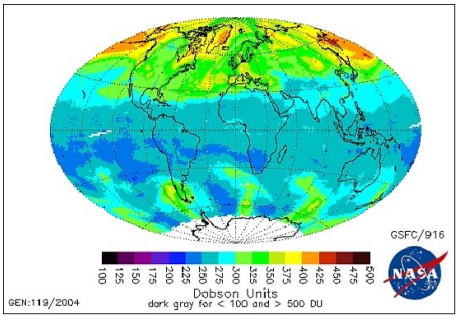
14

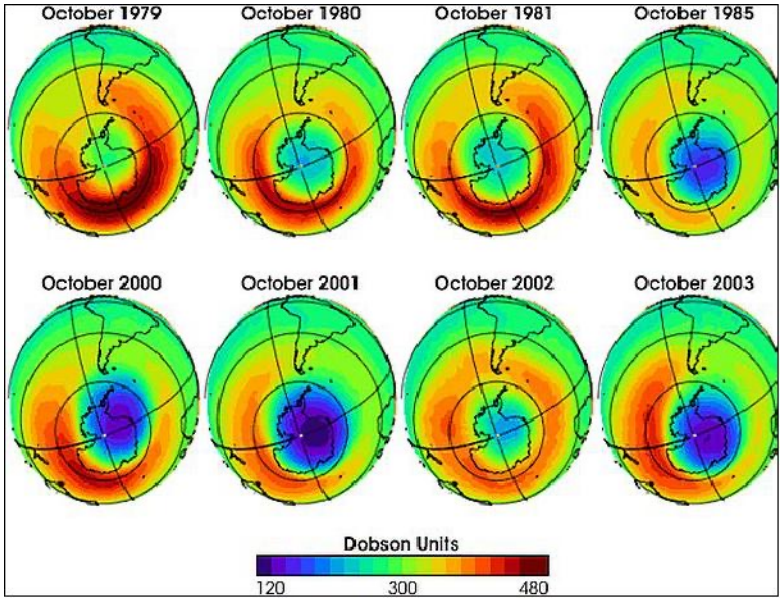
Ozone Concentrations (1978-1985)



Nimbus-7

Total Ozone May 1993





15

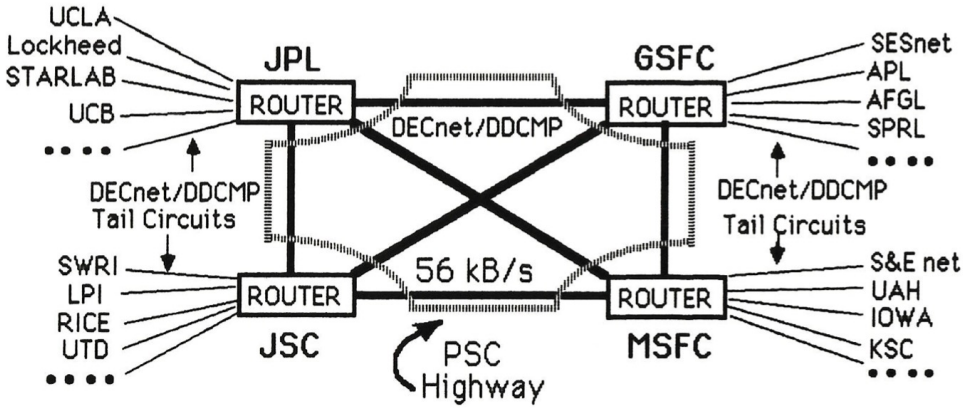
Computer Networking

Without good “open” science networks, rapid access to space-derived data and information would be impossible

16

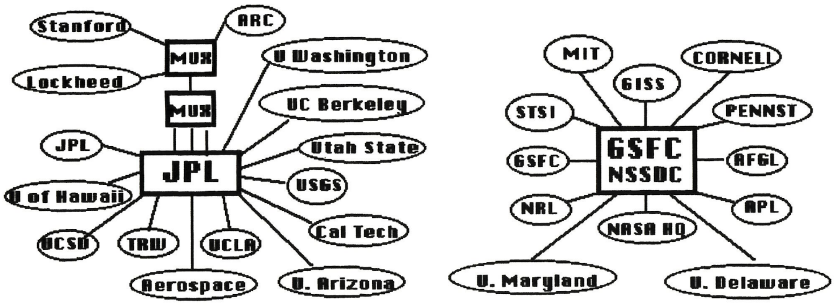
Space Physics Analysis Network (SPAN)

SPAN started in 1980 at MSFC

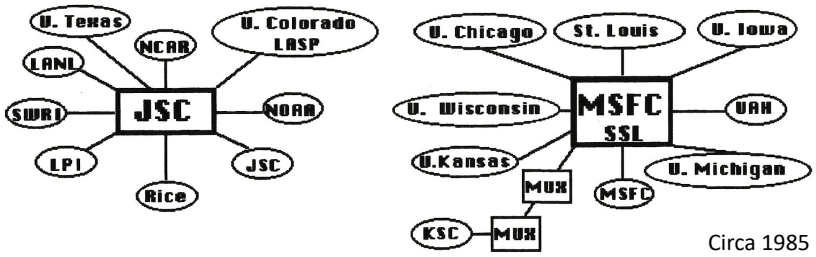


Circa 1985

17



SPAN CONFIGURATION



Circa 1985

18

ESA SCIENCE NETWORK

Established to facilitate the transfer of near-real time ICE data to European Investigators

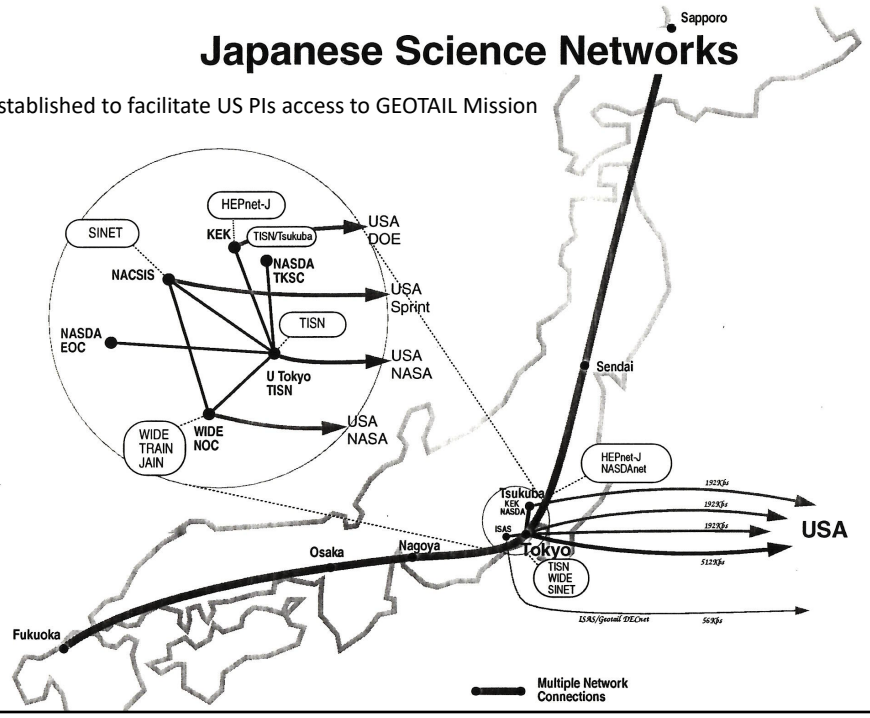
Established in 1992

Circa 1985

19

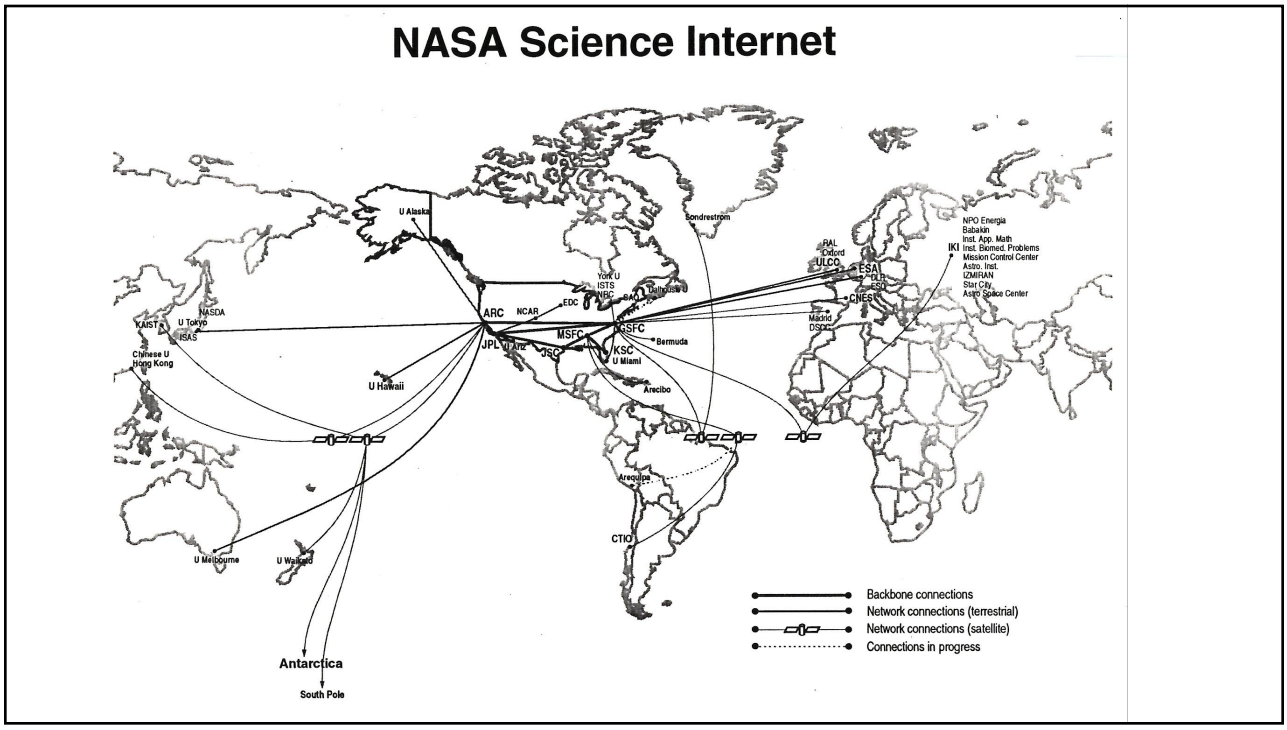
Japanese Science Networks

Established to facilitate US PIs access to GEOTAIL Mission



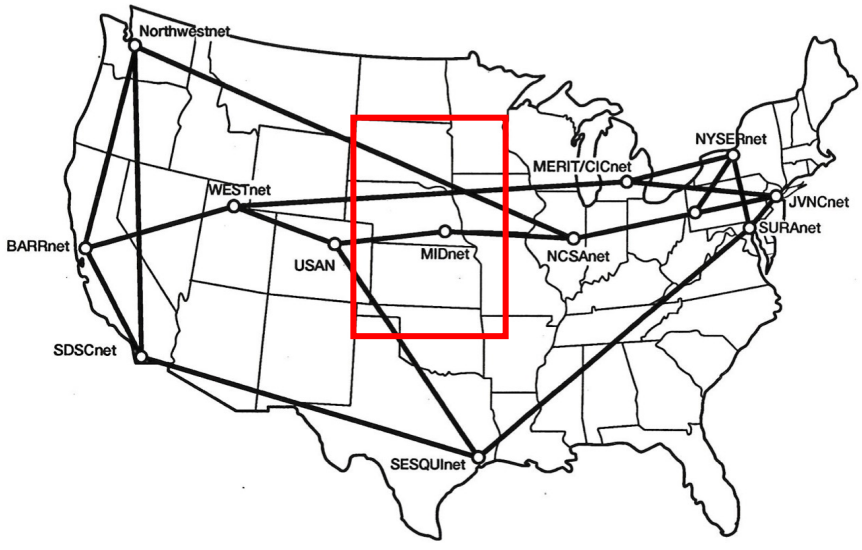
Circa 1987

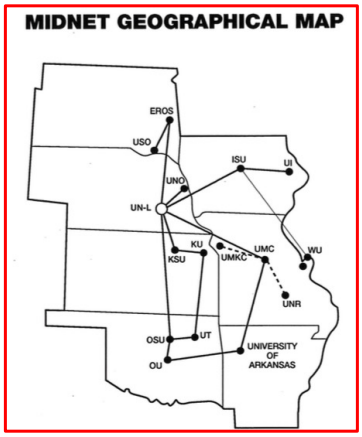
20



21

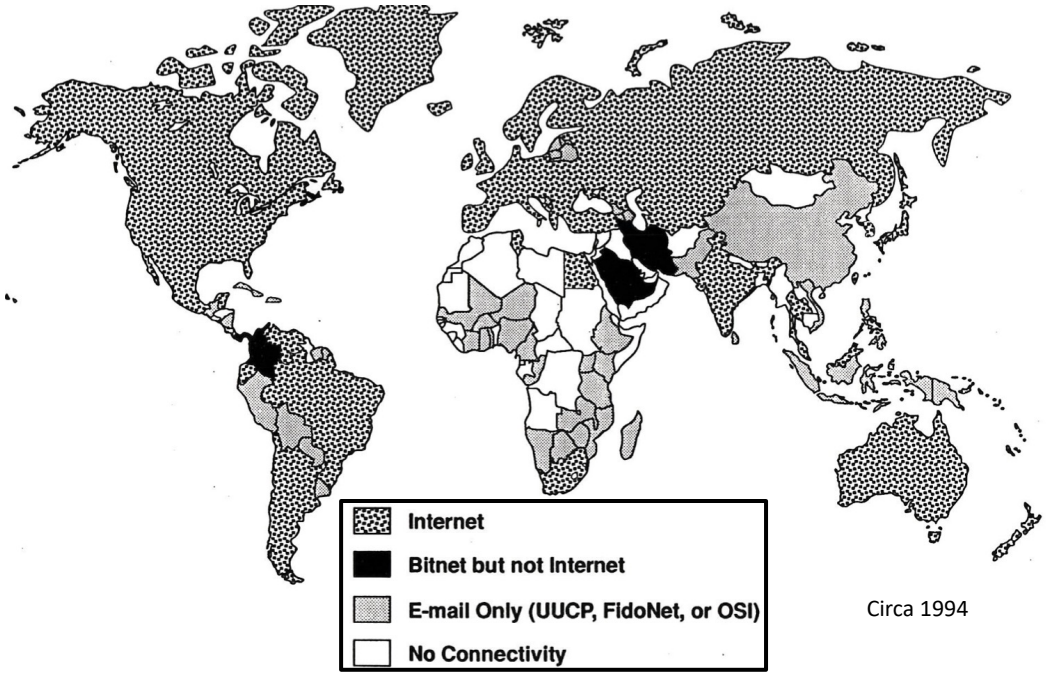
NSF Backbone Network





Circa 1993

22

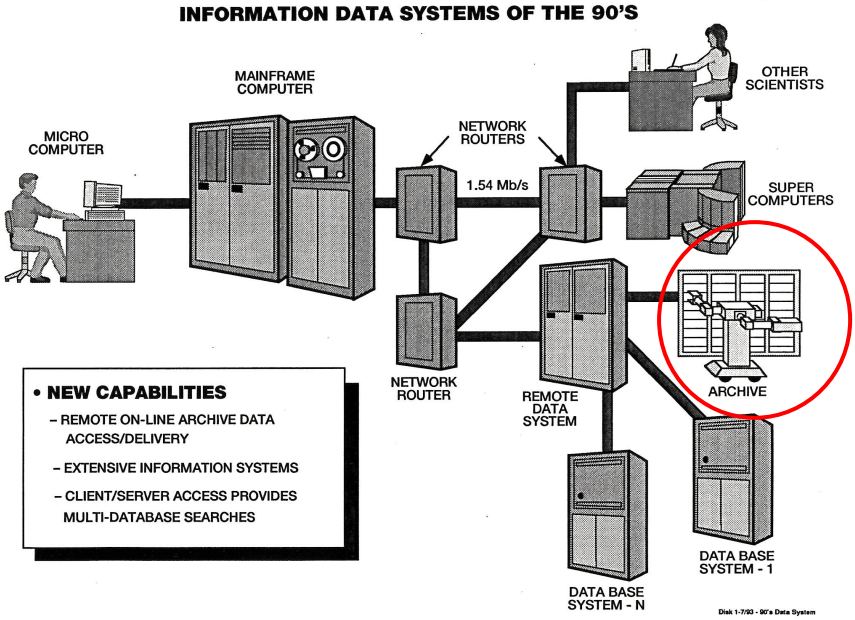


Circa 1994

23

Online Data Systems to Distributed Active Archives

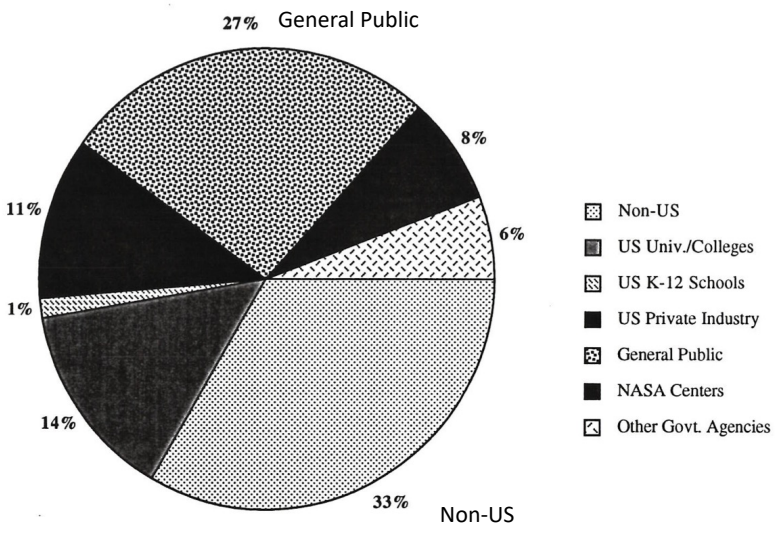
24



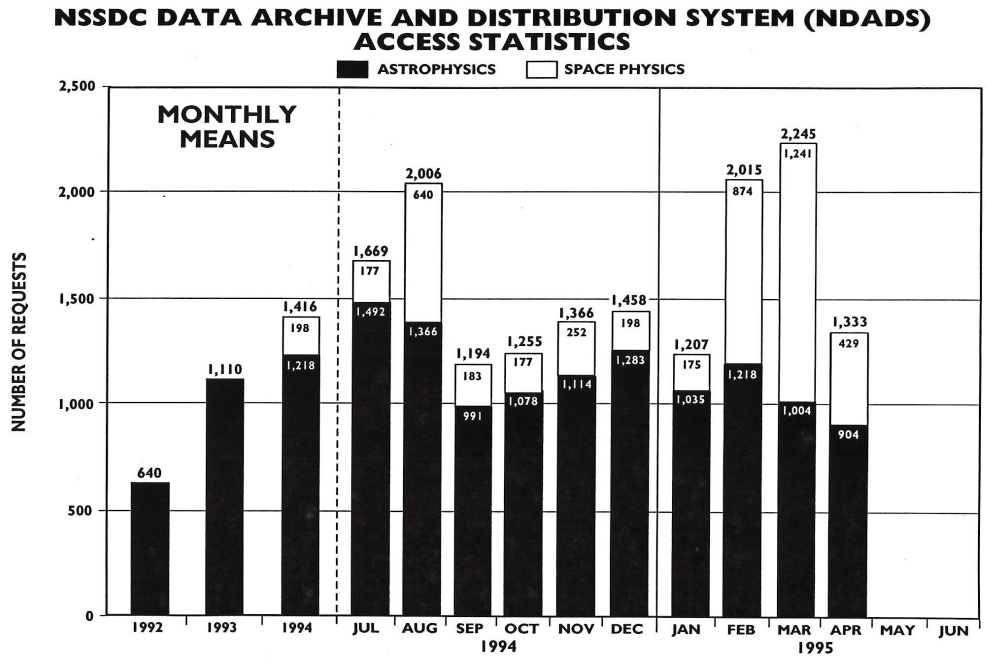
25

National Space Science Data Center (circa 1994)

- **NSSDC established in 1966**
  - Maintains over 8 TB of digital data
  - Several 10's of millions of feet of film
  - Supports:
    - ❖ Astrophysics
    - ❖ Space Plasma Physics
    - ❖ Planetary Physics
    - ❖ Some Earth science
- **Facilities include:**
  - Offline data storage (~ 7 TB)
  - Online data storage (~ 1 TB)
  - CD-ROM Browse Facility
- **Data Access**
  - Phone, fax, visit
  - Electronic mail, ftp, WWW



26



27

Early Review of NASA Data Archiving And Access

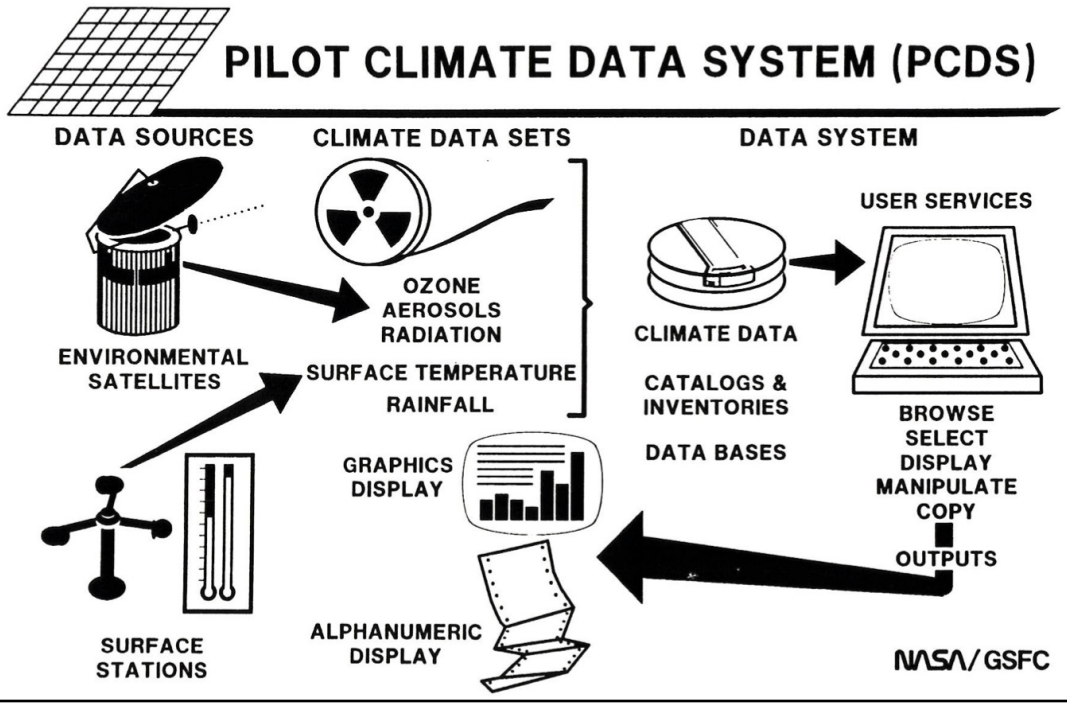
- Archiving in the 1960s and 1970s focused on depositing data sets and documentation into the NSSDC
- National Research Council's Space Science Board initiated a Committee on Computation and Data Management (CODMAC) in 1982
  - Determine the state of archiving and make recommendations for improvement
- Key CODMAC recommendations:
  - The most useful archives are those managed by scientists who use the data
  - Distributed archives situated at data scientist's institutions should maximize the ability to make new discoveries

28

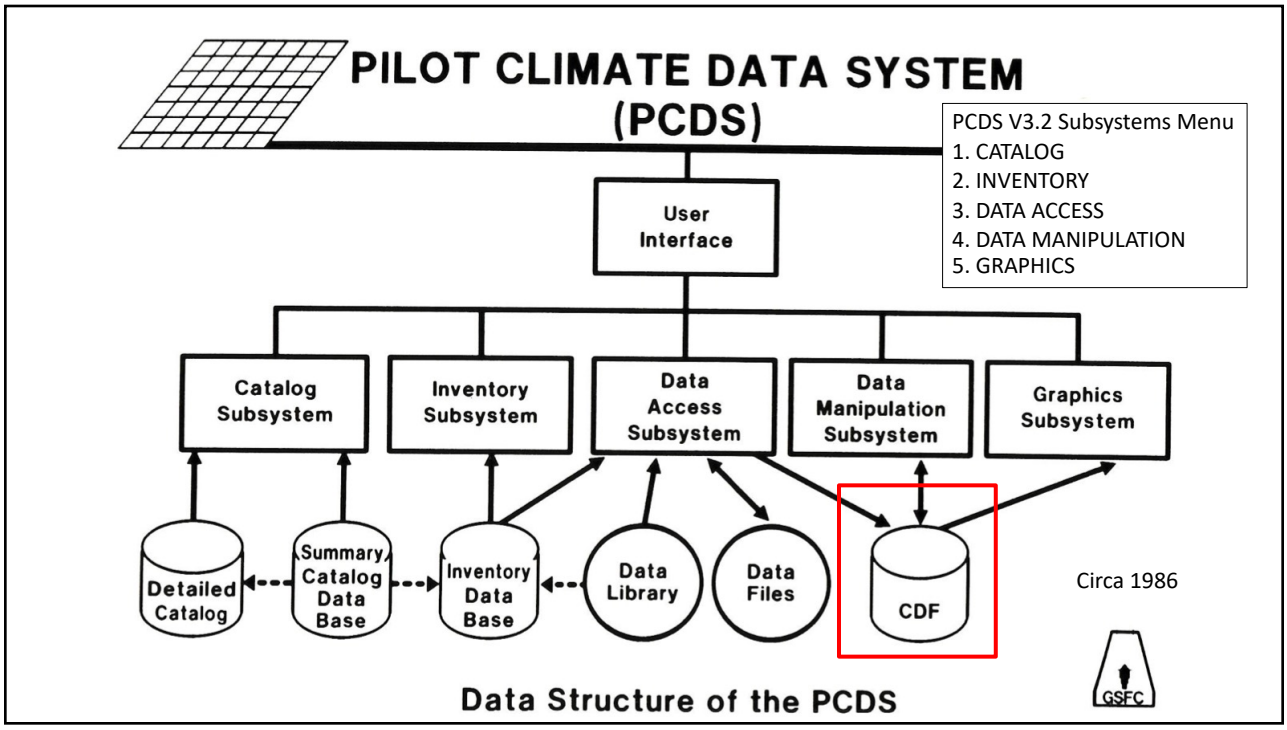
Evolution of NASA Data Systems

- The Astronomical Data Center (1977)
  - NSSDC, GSFC Lab for Astronomy & Solar Physics, Centre de Donnees Stellairs
- Crustal Dynamics Data Information System (1981 - 1994) – transferred to EOSDIS
- Pilot Climate Data System (1983 - 1994) – transferred to EOSDIS
- Pilot Land Data System (1984- 1994) – transferred to EOSDIS
- NASA Ocean Data System (1987-1994) – transferred to EOSDIS
- Planetary Data System (1988)
- Astrophysics Data System (1989)
- Space Physics Data Facility (1990)
- Earth Observing System Data and Information System (1990)

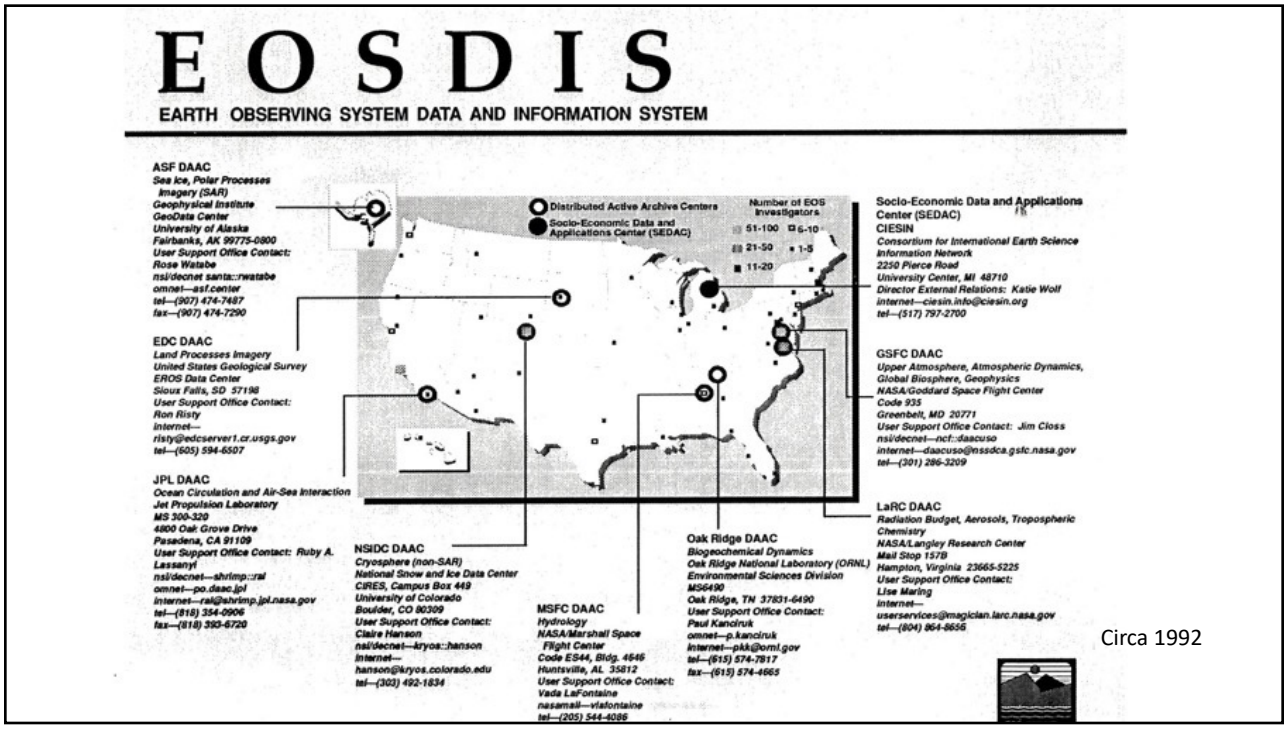
29



30


PILOT CLIMATE DATA SYSTEM
(PCDS)
PCDS V3.2 Subsystems Menu
1. CATALOG
2. INVENTORY
3. DATA ACCESS
4. DATA MANIPULATION
5. GRAPHICS
User Interface
Catalog Subsystem
Inventory Subsystem
Data Access Subsystem
Data Manipulation Subsystem
Graphics Subsystem
Detailed Catalog
Summary Catalog Data Base
Inventory Data Base
Data Library
Data Files
CDF
Circa 1986
GSFC
Data Structure of the PCDS

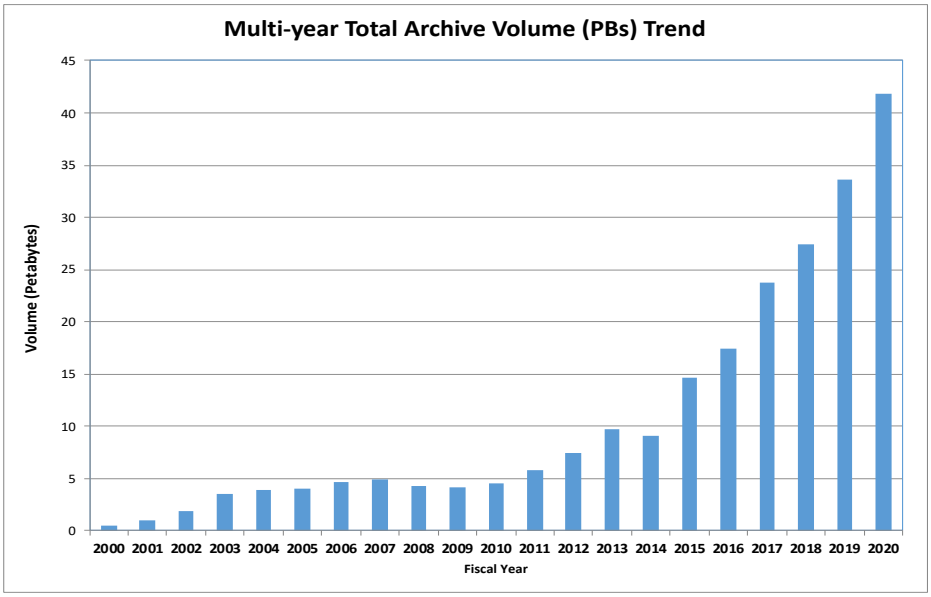
31


EOSDIS
EARTH OBSERVING SYSTEM DATA AND INFORMATION SYSTEM
ASF DAAC
Sea Ice, Polar Processes
Imagery (SAR)
Geophysical Institute
GeoData Center
University of Alaska
Fairbanks, AK 99775-0800
User Support Office Contact:
Rose Watabe
nsi/decnet santa::rwatabe
omnet—asf.center
tel—(907) 474-7487
fax—(907) 474-7290
EDC DAAC
Land Processes Imagery
United States Geological Survey
EROS Data Center
Sioux Falls, SD 57198
User Support Office Contact:
Ron Risty
internet—
risty@edcserver1.cr.usgs.gov
tel—(605) 594-6507
JPL DAAC
Ocean Circulation and Air-Sea Interaction
Jet Propulsion Laboratory
MS 300-320
4800 Oak Grove Drive
Pasadena, CA 91109
User Support Office Contact: Ruby A. Lassanyi
nsi/decnet—shrimp::rsl
omnet—po.daac.jpl
internet—rsl@shrimp.jpl.nasa.gov
tel—(818) 354-0906
fax—(818) 393-6720
NSIDC DAAC
Cryosphere (non-SAR)
National Snow and Ice Data Center
CIRES, Campus Box 449
University of Colorado
Boulder, CO 80309
User Support Office Contact:
Claire Hanson
nsi/decnet—kryos::hanson
internet—
hanson@kryos.colorado.edu
tel—(303) 492-1834
MSFC DAAC
Hydrology
NASA/Marshall Space Flight Center
Code ES44, Bldg. 4646
Huntsville, AL 35812
User Support Office Contact:
Vada LaFontaine
nasamail—vlafontaine
tel—(205) 544-4086
Oak Ridge DAAC
Biogeochemical Dynamics
Oak Ridge National Laboratory (ORNL)
Environmental Sciences Division
MS6490
Oak Ridge, TN 37831-6490
User Support Office Contact:
Paul Kanciruk
omnet—p.kanciruk
internet—pkk@ornl.gov
tel—(615) 574-7817
fax—(615) 574-4665
LaRC DAAC
Radiation Budget, Aerosols, Tropospheric Chemistry
NASA/Langley Research Center
Mail Stop 157B
Hampton, Virginia 23665-5225
User Support Office Contact:
Lise Maring
internet—
userservices@magician.larc.nasa.gov
tel—(804) 864-8656
GSFC DAAC
Upper Atmosphere, Atmospheric Dynamics, Global Biosphere, Geophysics
NASA/Goddard Space Flight Center
Code 935
Greenbelt, MD 20771
User Support Office Contact: Jim Closs
nsi/decnet—ncf::daacuso
internet—daacuso@nssdca.gsfc.nasa.gov
tel—(301) 286-3209
Socio-Economic Data and Applications Center (SEDAC)
CIESIN
Consortium for International Earth Science Information Network
2250 Pierce Road
University Center, MI 48710
Director External Relations: Katie Wolf
internet—ciesin.info@ciesin.org
tel—(517) 797-2700
Distributed Active Archive Centers
Socio-Economic Data and Applications Center (SEDAC)
Number of EOS Investigators
51-100
21-50
11-20
6-10
1-5
Circa 1992

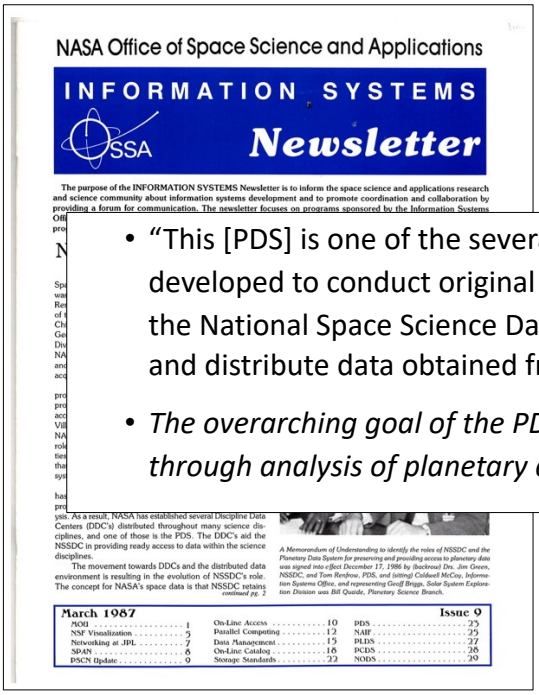
32

EOSDIS Total Archive Trend

Volume By FY	Total Volume (PBs)
2000	0.51
2001	0.99
2002	1.92
2003	3.55
2004	3.94
2005	4.06
2006	4.62
2007	4.91
2008	4.29
2009	4.20
2010	4.50
2011	5.76
2012	7.40
2013	9.76
2014	9.12
2015	14.63
2016	17.50
2017	23.81
2018	27.37
2019	33.60
2020	41.86



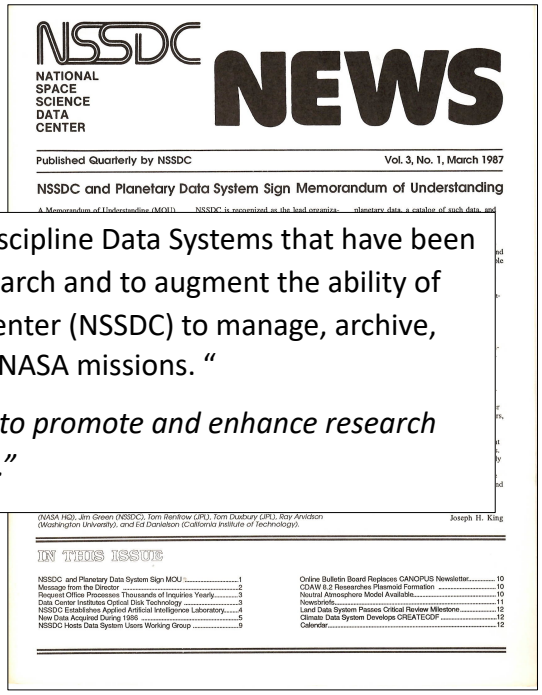
33


NASA Office of Space Science and Applications

INFORMATION SYSTEMS Newsletter

The purpose of the INFORMATION SYSTEMS Newsletter is to inform the space science and applications research and science community about information systems development and to promote coordination and collaboration by providing a forum for communication. The newsletter focuses on programs sponsored by the Information Systems

March 1987 Issue 9


NSSDC NEWS

NATIONAL SPACE SCIENCE DATA CENTER

Published Quarterly by NSSDC Vol. 3, No. 1, March 1987

NSSDC and Planetary Data System Sign Memorandum of Understanding

IN THIS ISSUE

- "This [PDS] is one of the several Discipline Data Systems that have been developed to conduct original research and to augment the ability of the National Space Science Data Center (NSSDC) to manage, archive, and distribute data obtained from NASA missions. "
- *The overarching goal of the PDS is to promote and enhance research through analysis of planetary data."*

34

The Planetary Data System

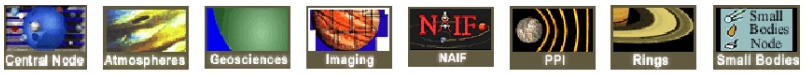
The PDS acquires, preserves, and distributes the large volume of unique and valuable data returned by Solar System Exploration missions

Key PDS Products and Services:

- High quality peer-reviewed data archives
- Data distribution to planetary community
- Archiving expertise to planetary missions
- Scientific expertise and support for users
- Create value-added aggregated data products
- Education and outreach data products and services

http://pds.jpl.nasa.gov

NSSDC would maintain a backup role to PDS

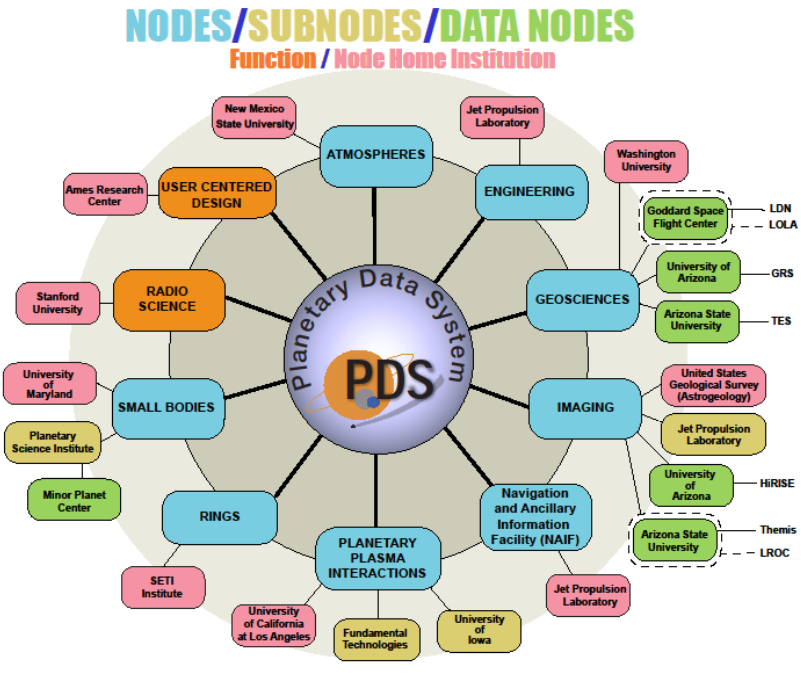


JPL | New Mexico State Univ. | Wash Univ. St. Louis | JPL/USGS Flagstaff | JPL | UCLA | NASA Ames | U. Maryland

Node structure provides focus on key disciplines

Circa 1991

35

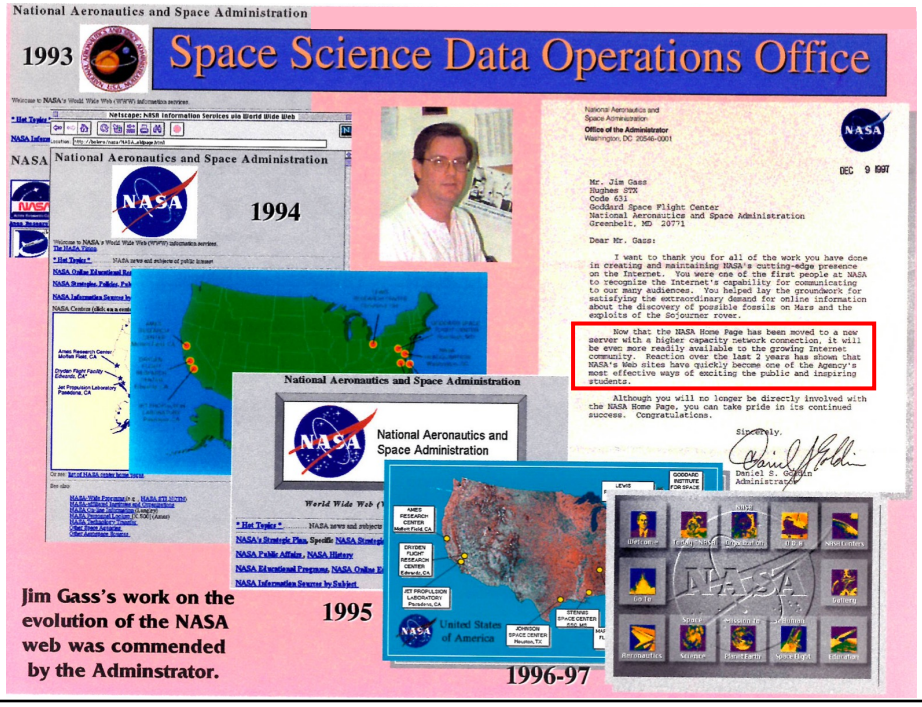


Circa 2014

36

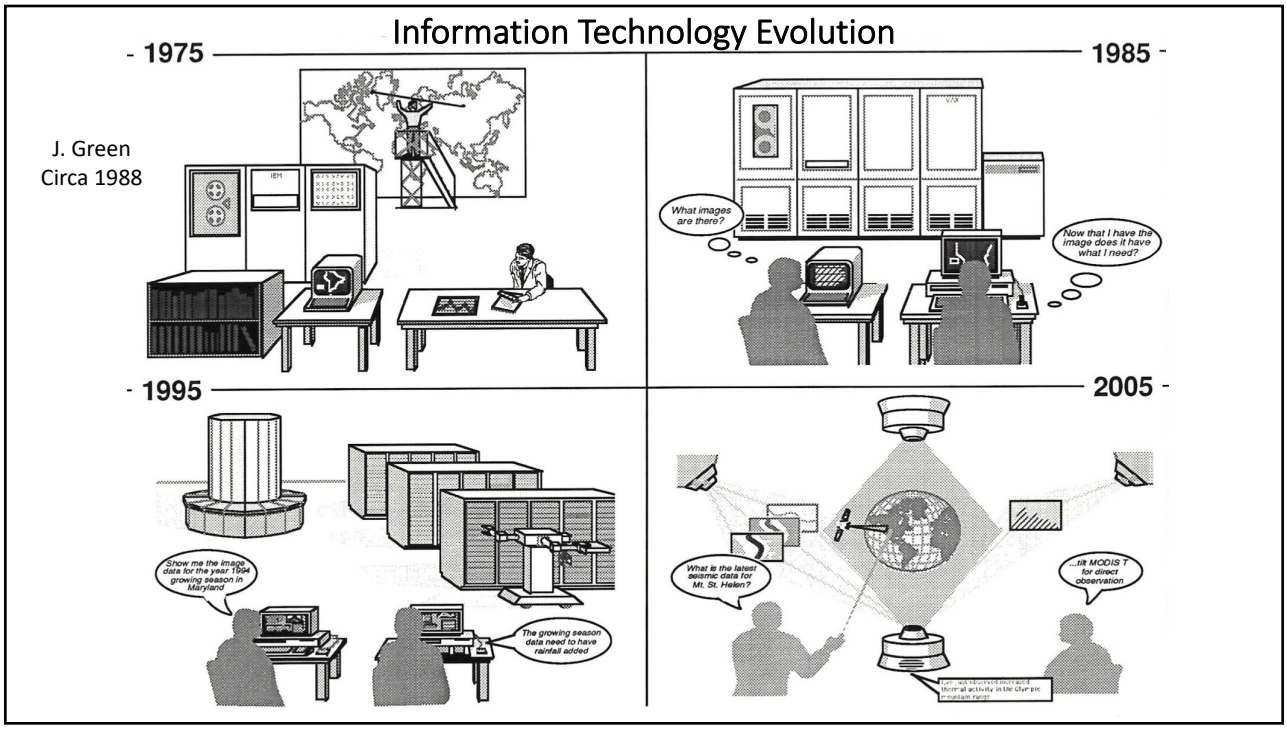
World Wide Web

37



NASA Public Affairs given
The responsibility to manage
WWW access to the public

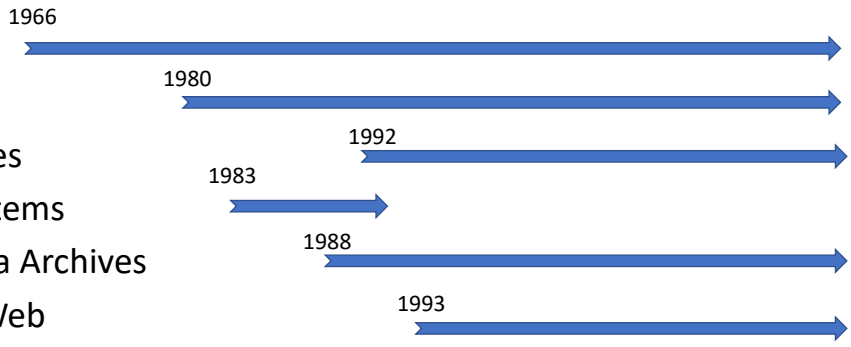
38



39

Summary NASA Data Access Evolution

- NSSDC Offline
- Networks
- Online Archives
- Pilot Data Systems
- Discipline Data Archives
- World Wide Web

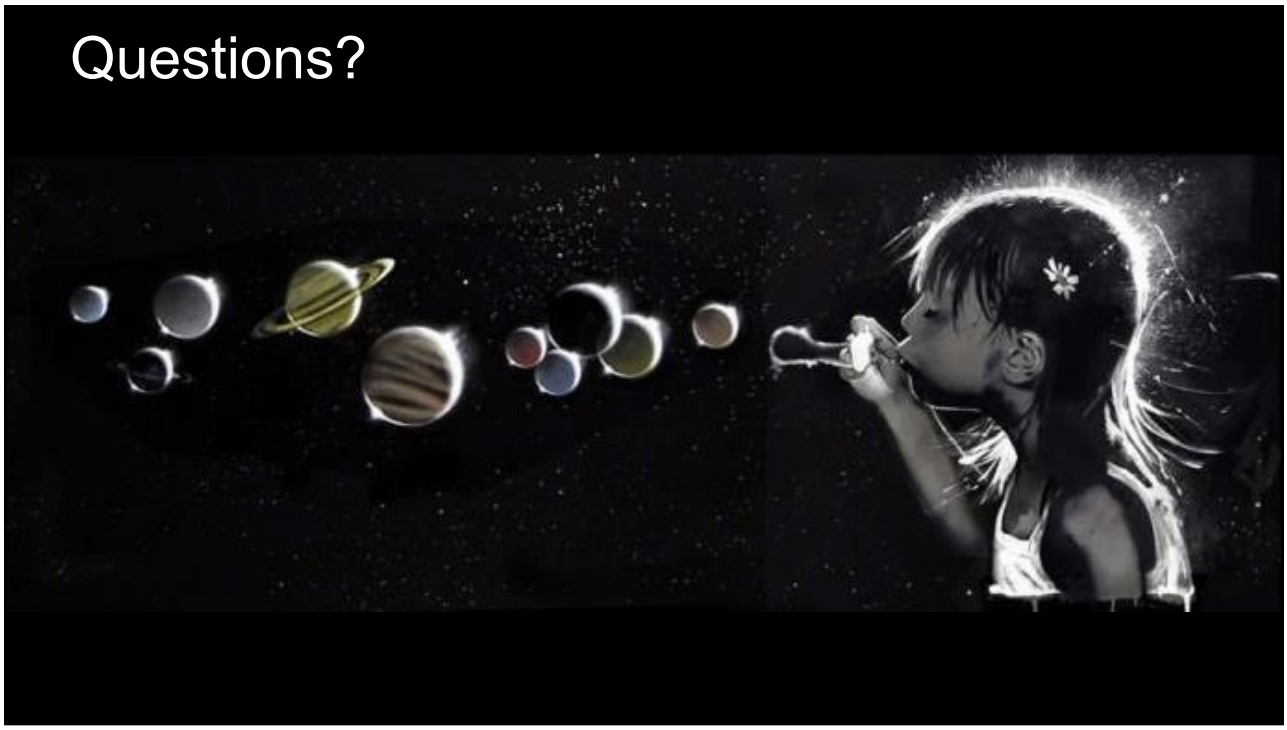


Along the way common data formats were developed (FITS, CDF, …)

40


Citizen Science
• 26 active projects and engage over 1.5 million volunteers. From studies of our planet's biodiversity to exploring the sun, comets, and planets outside of our solar system.
• 22 other project under development.
Exoplanets
Northern lights
Planet Forming Disks
Algae Blooms

41


Questions?

42