



NIST Data Preservation and Migration Strategy: Virtualization

Wo Chang, wchang@nist.gov, NIST

February 02, 2006

Agenda:

- Disclaimer
- NIST IT Projects
- “Computing” Technology: Past, Present, Future
- GPO Legacy CD data and applications
- Data Preservation and Migration Path
- Conclusion
- Future Direction
- Demo

Disclaimer:

- NIST does not endorse any mentioned products
- NIST does not promote or demote any technologies

NIST IT Projects:

Information Technology Laboratory (ITL)

- **Security Division:**

Intelligent management of IT risks, vulnerabilities and protection needs

- **Mathematics and Computational Science Division:**

Mathematical modeling, numerical algorithms, mathematical software, high performance computing, and scientific visualization

- **Software Testing Division:**

Tools and methods to improve quality, standards conformance, correctness

NIST IT Projects (cont.):

Information Technology Laboratory (ITL)

- **Networking Research Division:**
Research and measurement in high-speed, wireless, multimedia, and internet networks
- **Statistical Engineering Division:**
Scientific and industrial research through application of statistical methods to experimentation and data analysis
- **Information Access Division:**
User interfaces, text retrieval, human language, biometrics, digital multimedia, distributed computing study and measurement

NIST IT Projects (cont.):

Information Access Division (IAD)

- **Human Language Technology:**

Measurement and evaluation of speech recognition, speaker recognition, spoken language understanding, information search, retrieval, and filtering, document understanding, summarization, interactive speech interfaces

- **Biometrics Technology: Identification and Verification Technology**

Measurement, evaluation and standards for fingerprint matching and interchange, criminal justice information systems, face recognition and multi-modal biometrics

NIST IT Projects (cont.):

Information Access Division (IAD)

- **Interactive Systems Technology:**

Measurement, standards, and evaluation for usability of interactive systems, web usability testing, usability reporting formats, accessibility testing

- **Digital Multimedia Technology:**

Measurement, evaluation and standards for multimedia, including searching, filtering, MPEG and JPEG standards, 3D on the web, video retrieval, motion image quality and optical media

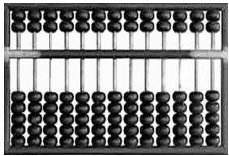
NIST IT Projects (cont.):

Campus operation (Chief Information Officer)

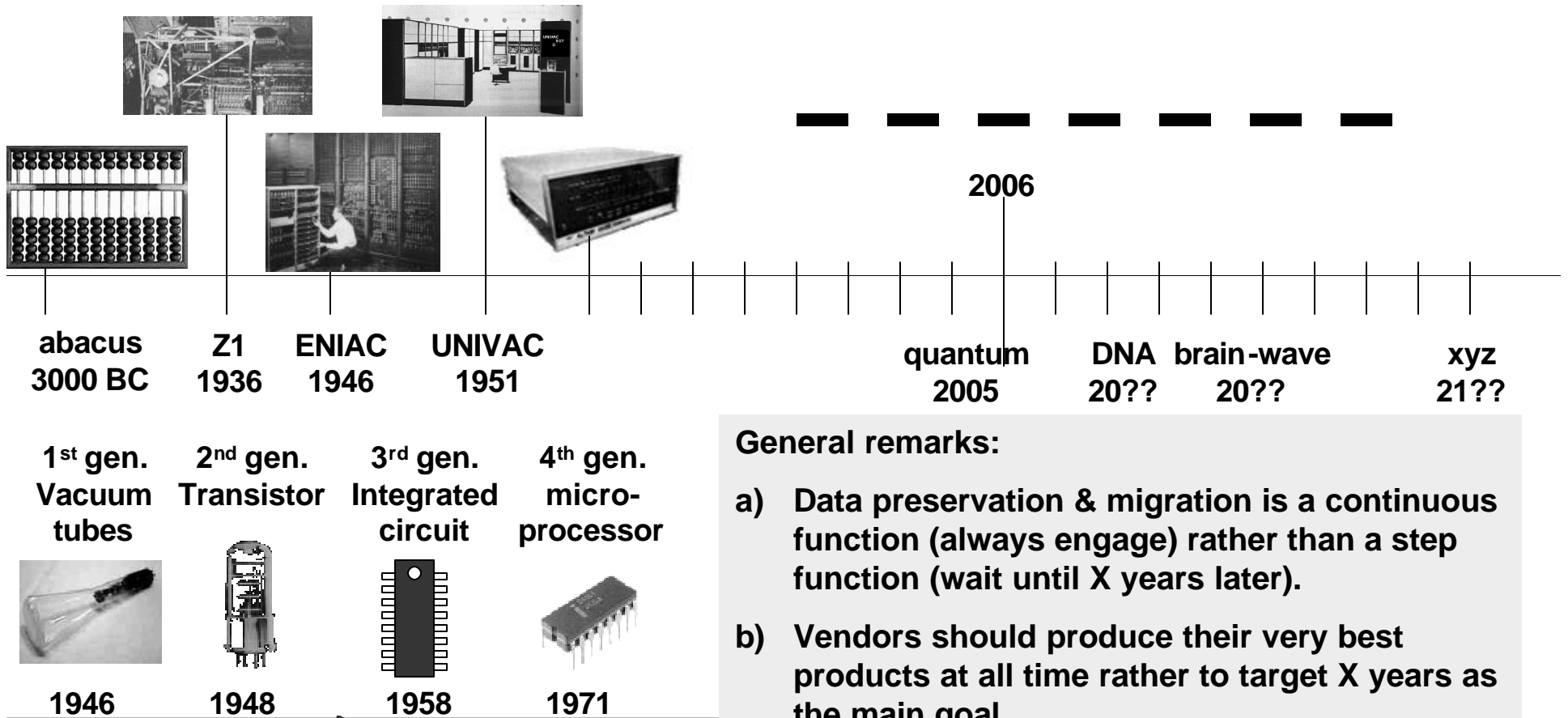
- Telecommunications
- IT Security and Networking
- Customer Access
- Applications Systems
- Enterprise Systems



“Computing” Technology: Past, Present, Future



“Computing” Technology: Past, Present, Future

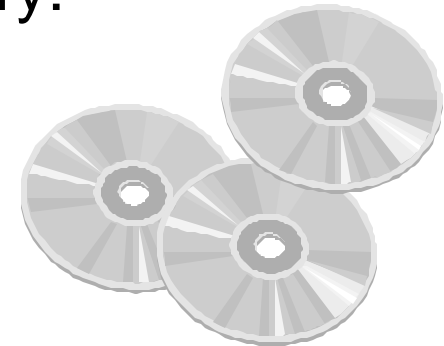


General remarks:

- a) Data preservation & migration is a continuous function (always engage) rather than a step function (wait until X years later).
- b) Vendors should produce their very best products at all time rather to target X years as the main goal.

GPO Legacy CD Data and Applications :

- Legacy CD data and applications are vital to GPO
- Due to physical medium degradation and lack of HW/SW support, these legacy data and applications may no longer be accessible.
- 20+ sample applications are in this category:
 - Date issued: 1988 – 2000
 - OS: DOS 3.11 – Windows 95



NIST Data Preservation and Migration Strategy: Virtualization

NIST Data Preservation and Migration Testbed

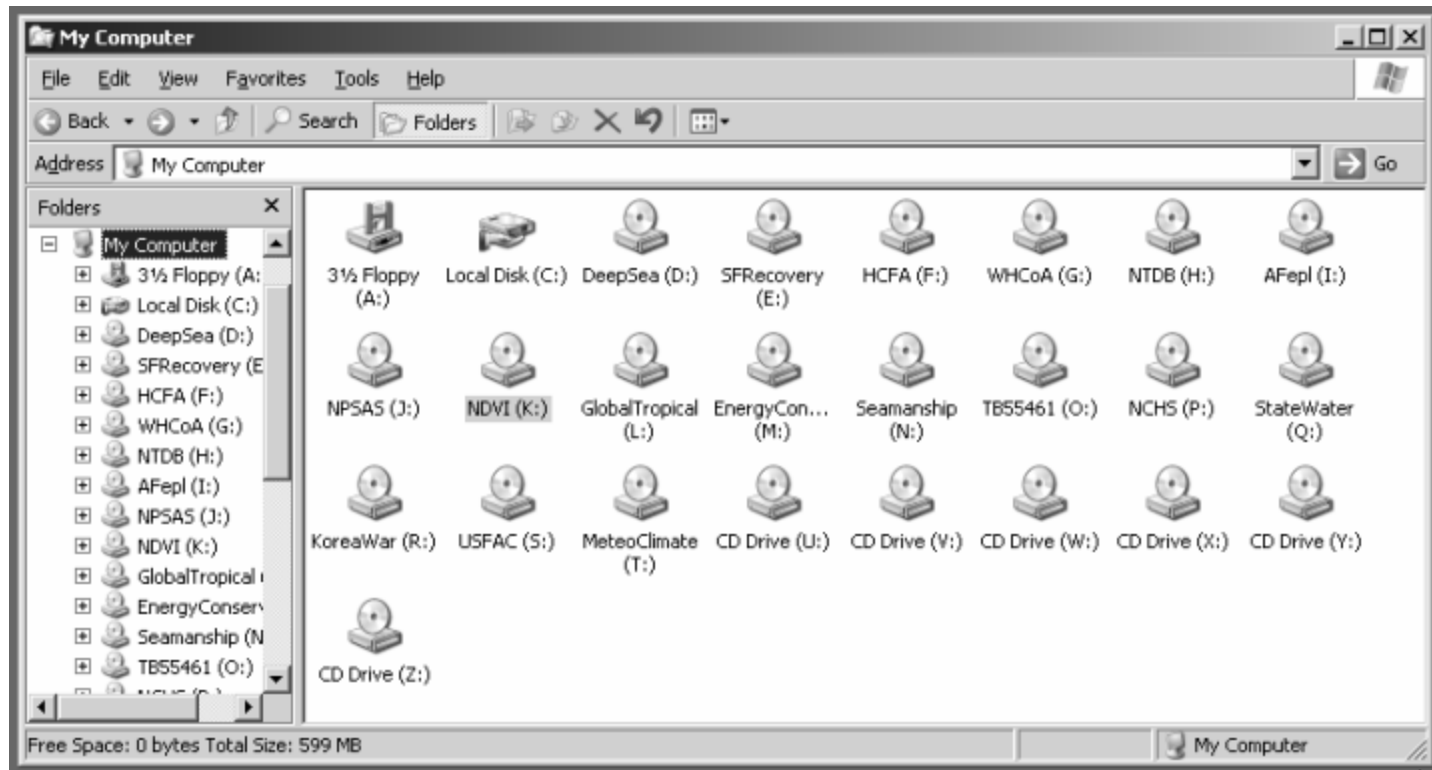
Apply virtual machines to host GPO legacy applications



NIST Data Preservation and Migration Strategy: Virtualization

NIST Data Preservation and Migration Testbed

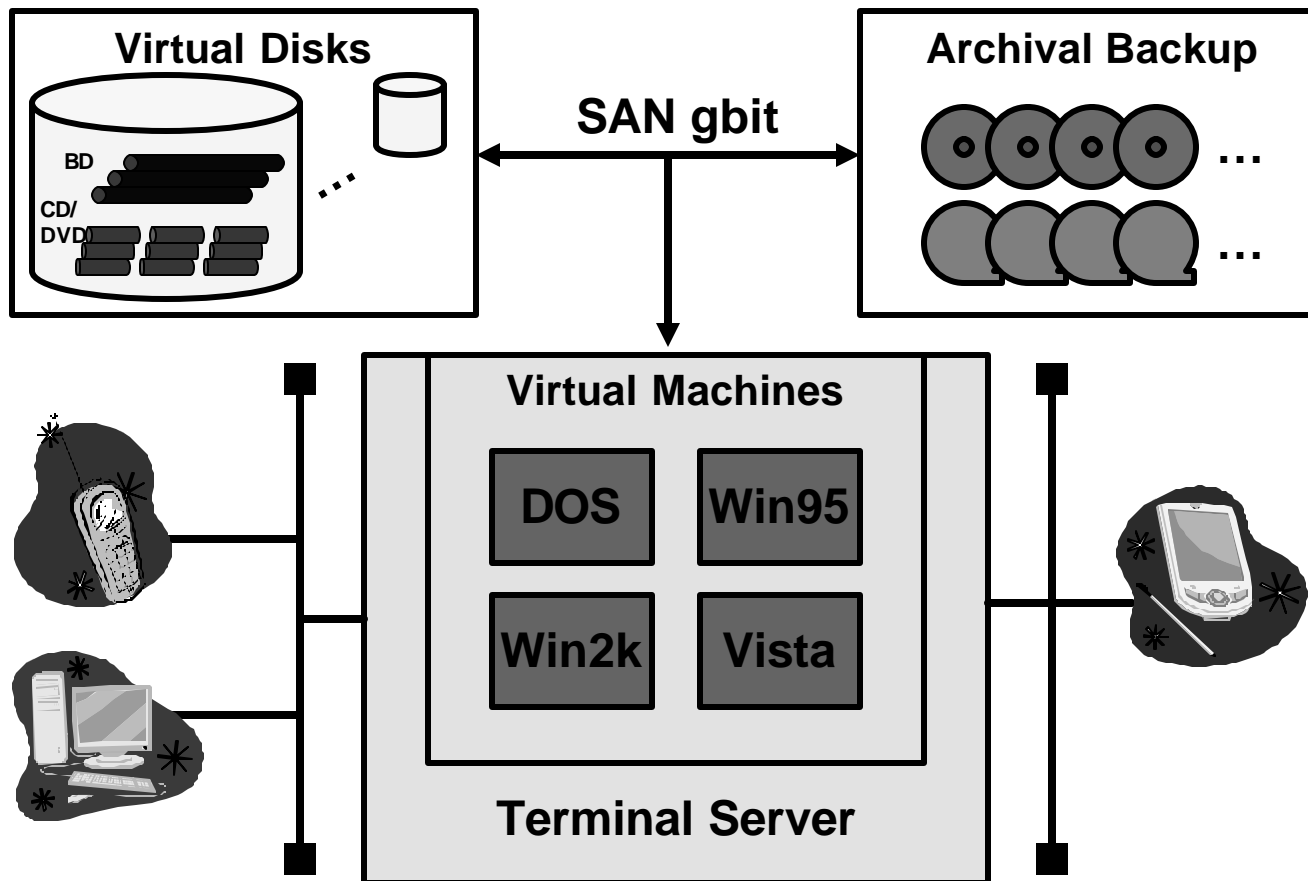
Apply virtual drives to emulate multiple CDs



NIST Data Preservation and Migration Strategy: Virtualization

NIST Data Preservation and Migration Testbed

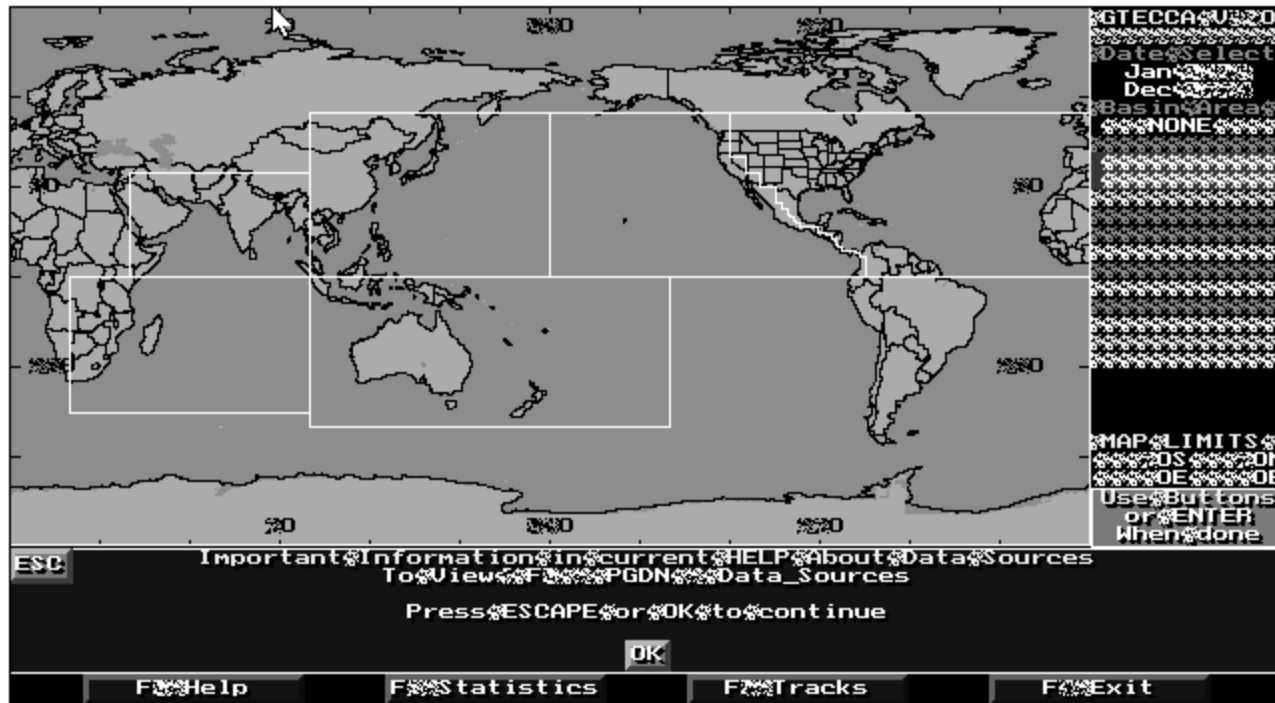
Long-term approach: Virtualization (need more work)



NIST Data Preservation and Migration Strategy: Virtualization

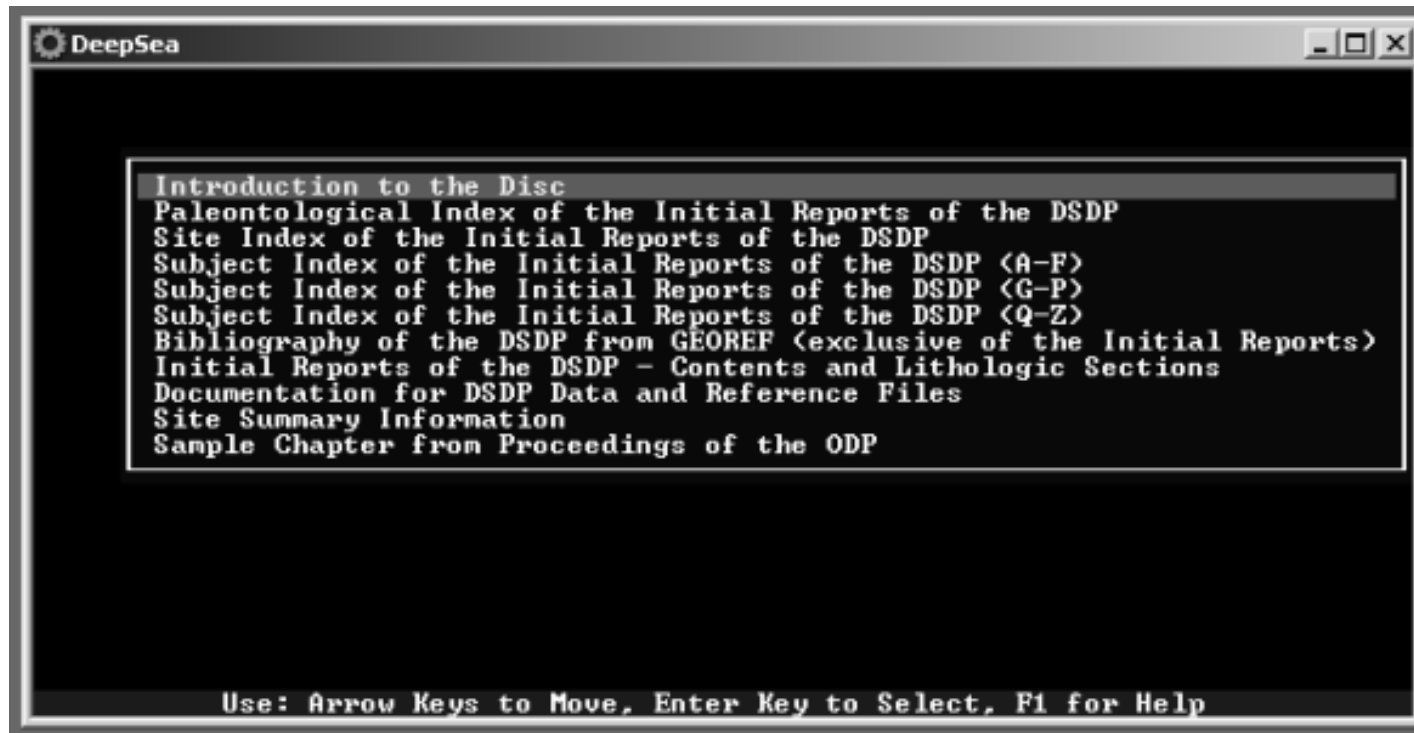
NIST Data Preservation and Migration Testbed

GPO Application: Global Tropical Cyclone Climatic (DOS 3.21, 1996)



NIST Data Preservation and Migration Testbed

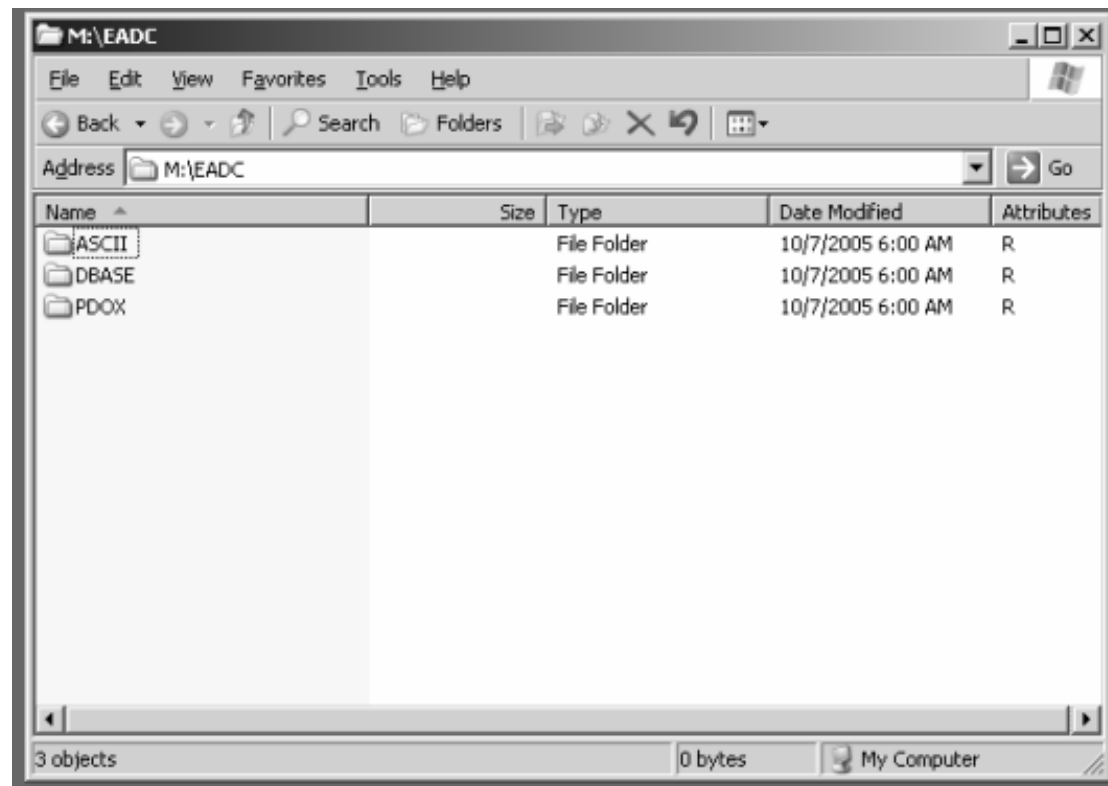
GPO Application: Deep Sea Drilling Project (DOS, 1992)



NIST Data Preservation and Migration Testbed

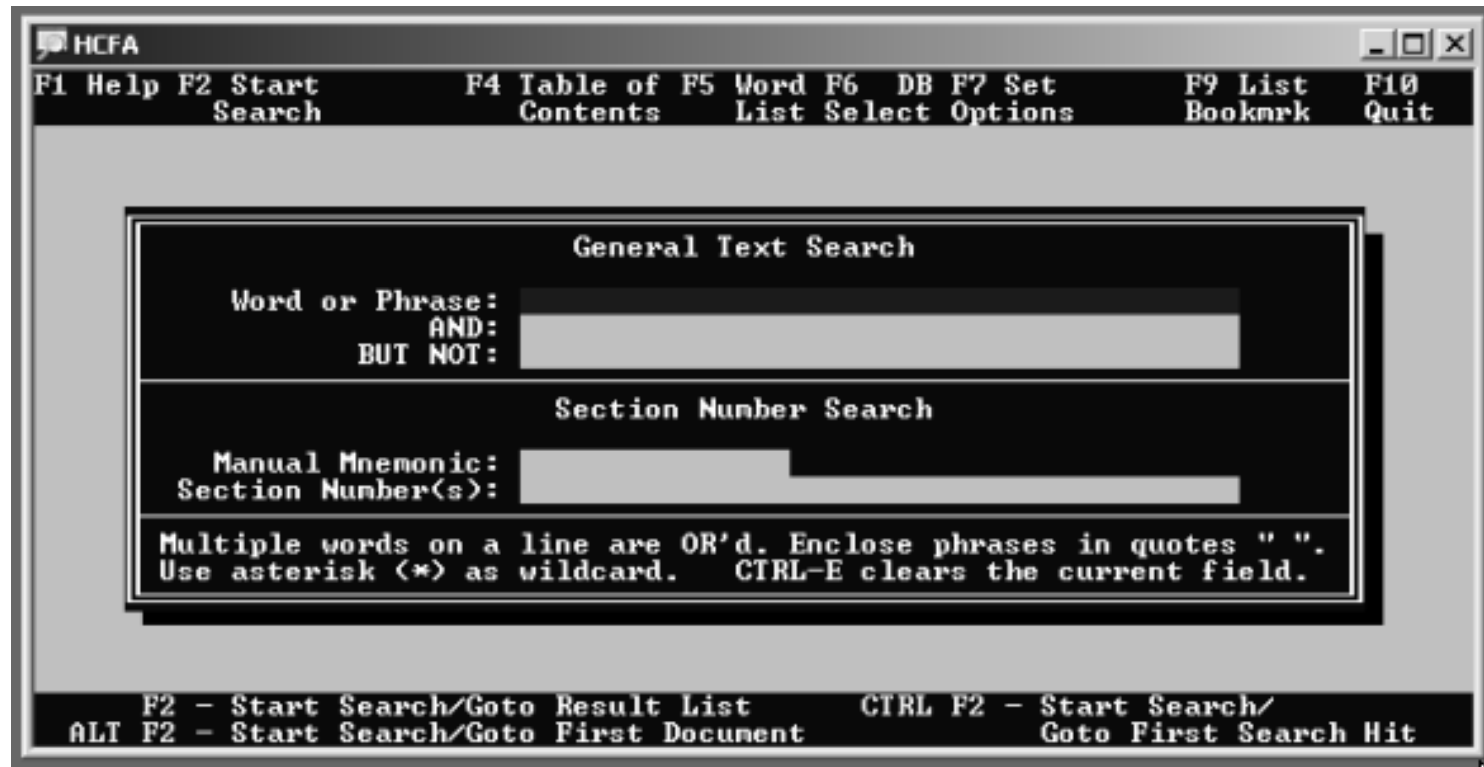
GPO Application: Deep Sea Drilling Project (DOS, 1992)

[needs dBase or Paradox to run]



NIST Data Preservation and Migration Testbed

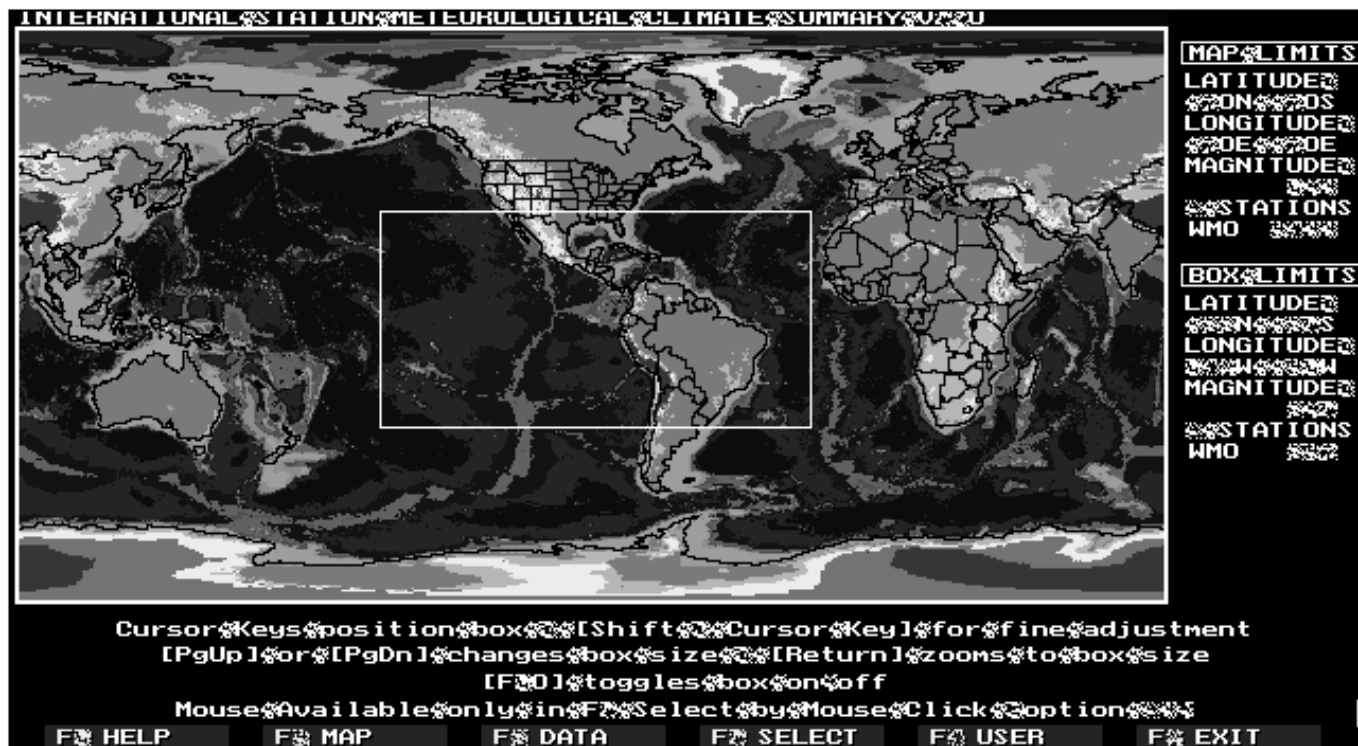
GPO Application: HCFA's Laws Regulations Manuals (Win, 1994)



NIST Data Preservation and Migration Strategy: Virtualization

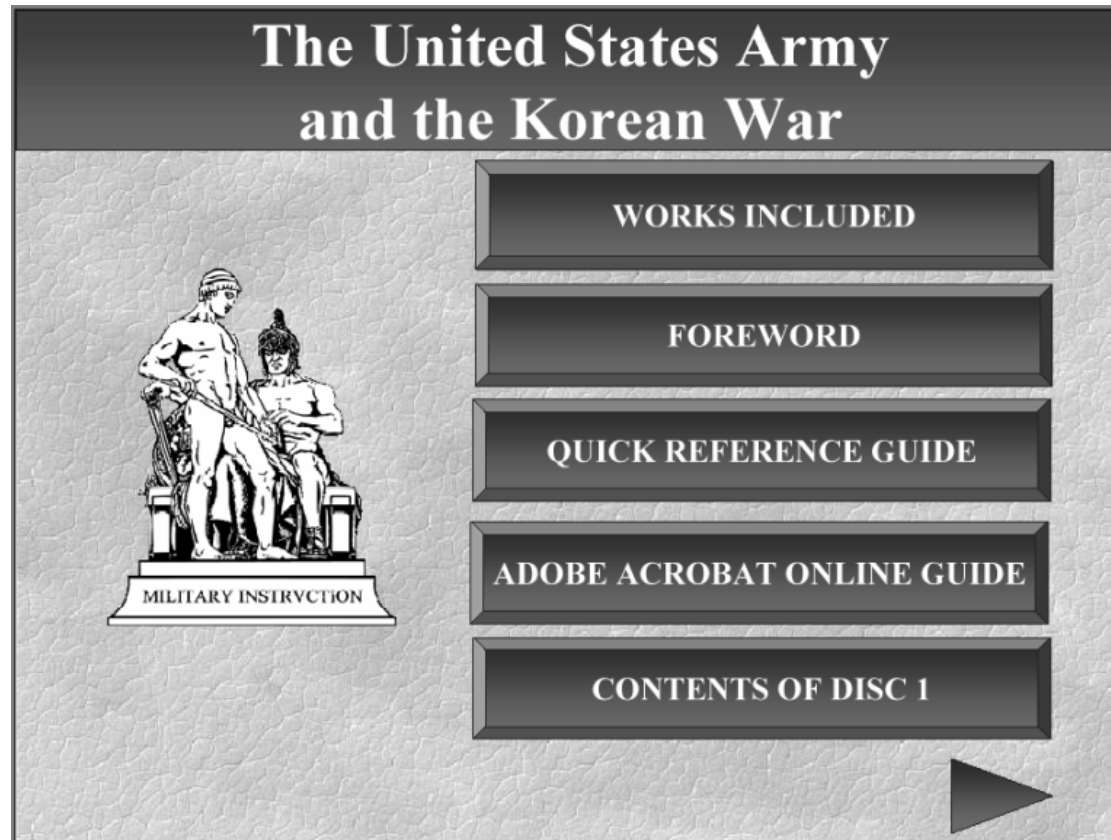
NIST Data Preservation and Migration Testbed

*GPO Application: International Station Meteorological Climate Summary
(DOS 4.01, 1996)*



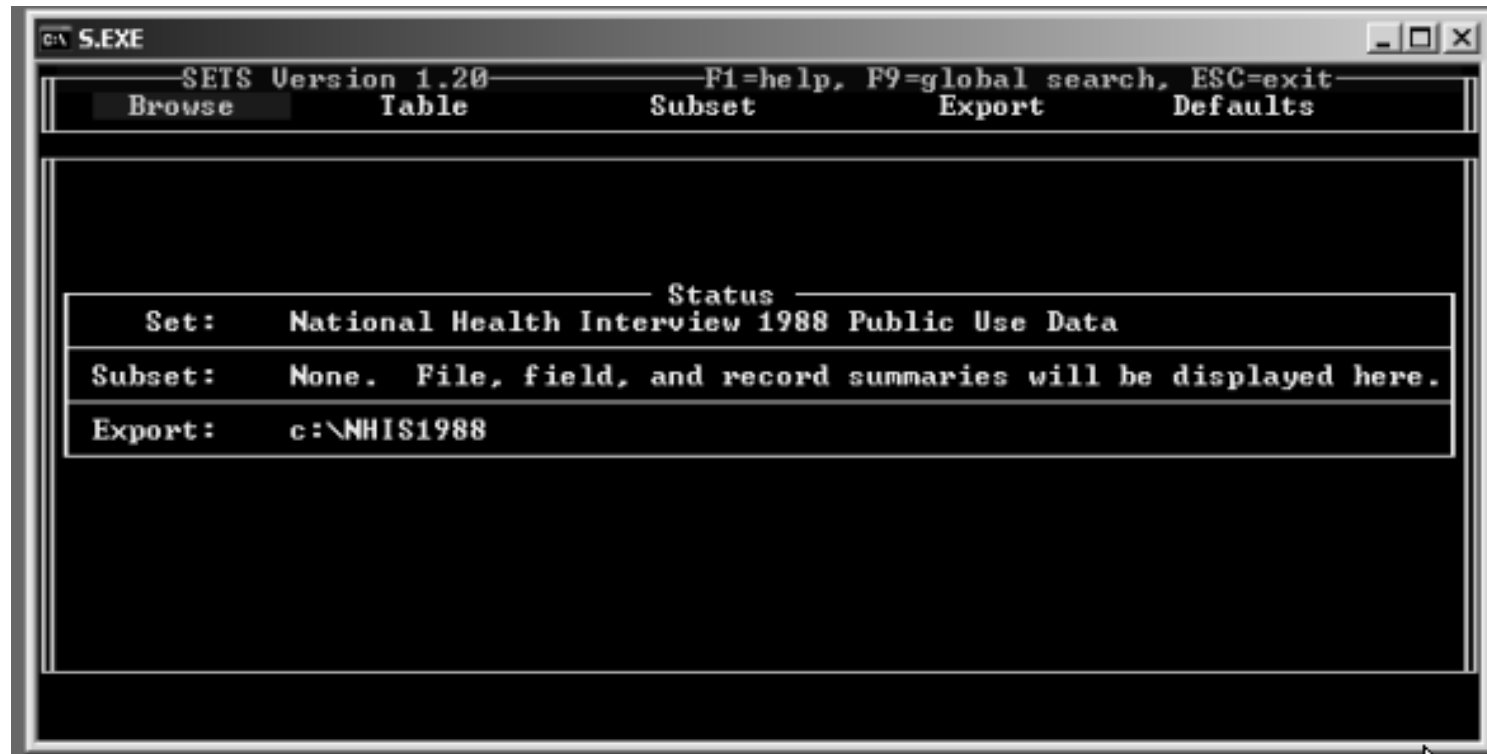
NIST Data Preservation and Migration Testbed

GPO Application: The United States Army and the Korean War (Win, 2000)



NIST Data Preservation and Migration Testbed

GPO Application: National Health Interview Survey 1990 (DOS, 1992)



NIST Data Preservation and Migration Testbed

GPO Application: NDVI and derived products (DOS, 1997)



NIST Data Preservation and Migration Testbed

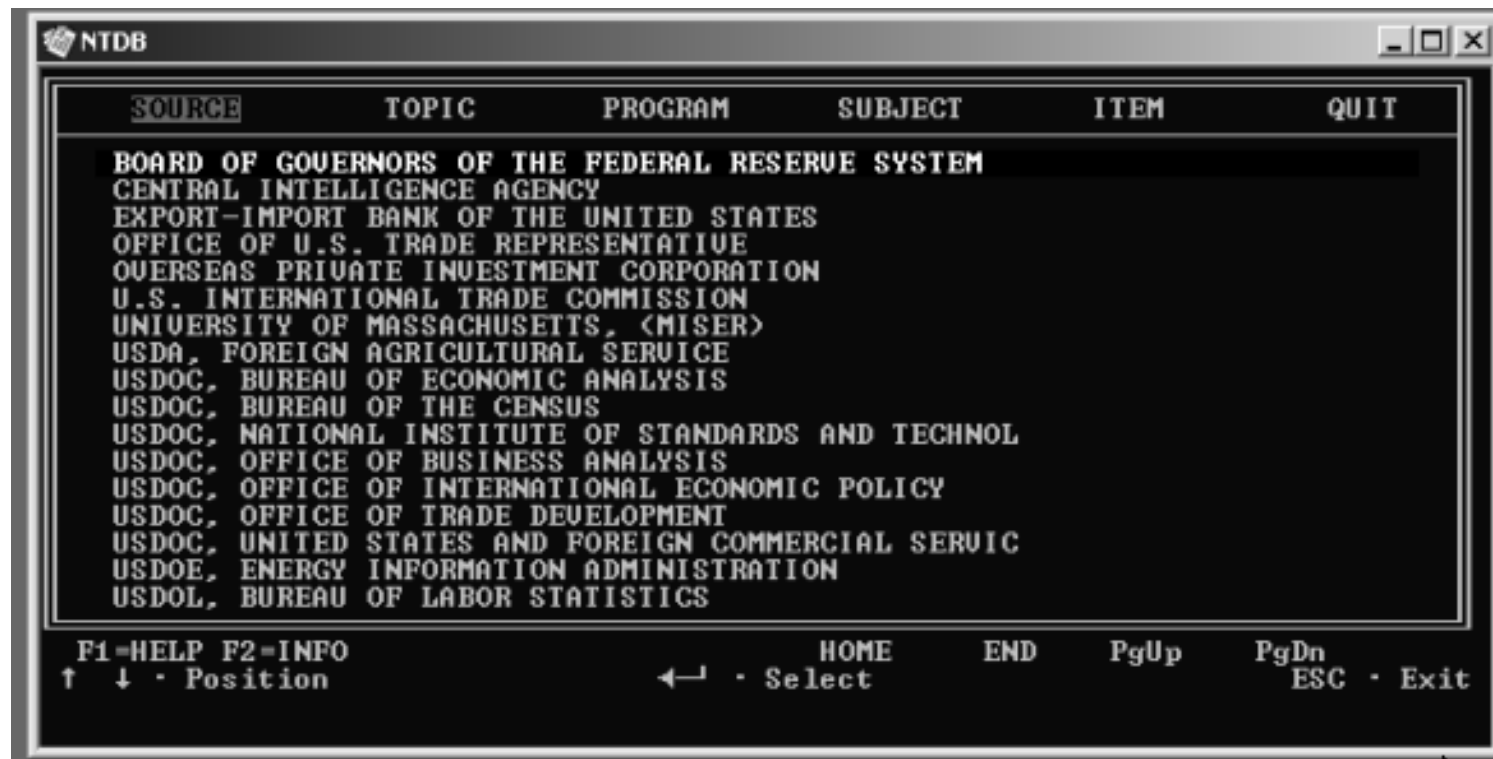
GPO Application: National Post-Secondary Student Aid Study

(DOS, 1992)



NIST Data Preservation and Migration Testbed

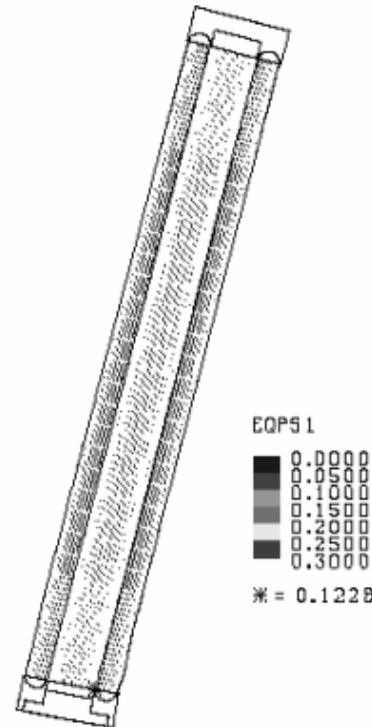
GPO Application: NTDB (Win, 1991)



NIST Data Preservation and Migration Testbed

GPO Application: Reexamination of Spent Fuel Shipment Risk Estimates (PC, 2000)

Figure A-1 Steel-lead-steel truck cask following a 30 mph corner impact



NIST Data Preservation and Migration Testbed

GPO Application: South Florida Multi-Species Recovery Plan (Win, 1999)



NIST Data Preservation and Migration Testbed

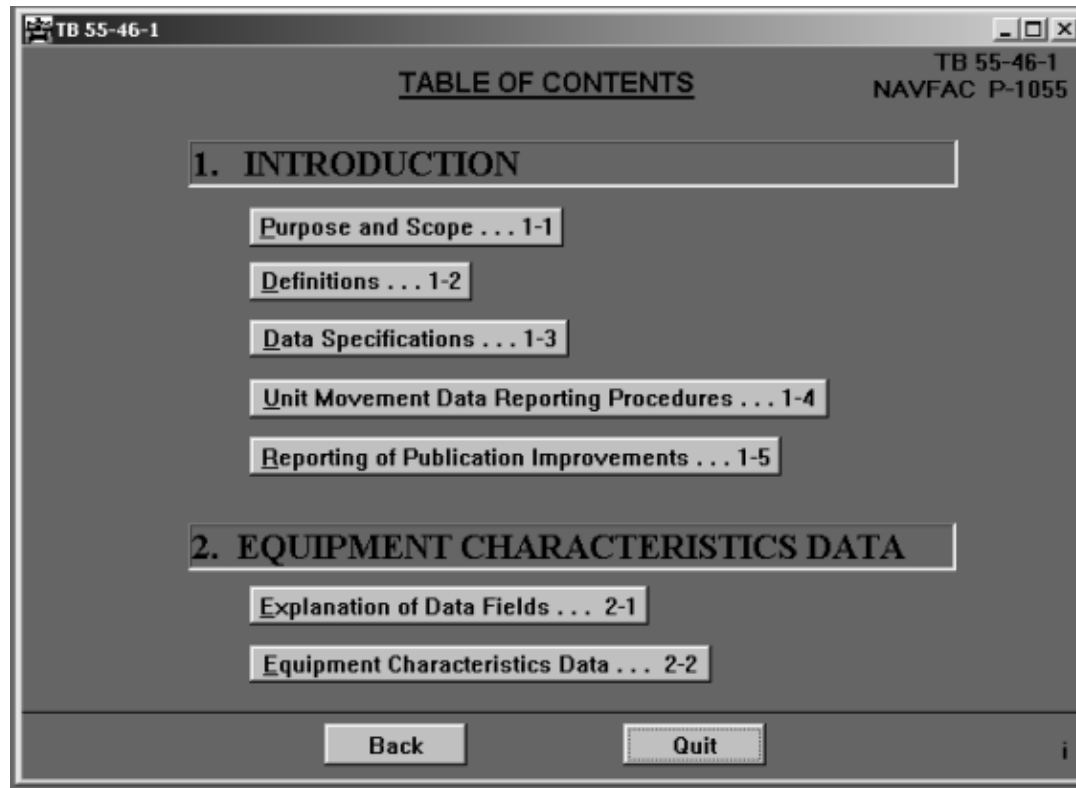
GPO Application: State Water Data Reports Open File Report 91463z

(DOS, 1990)



NIST Data Preservation and Migration Testbed

*GPO Application: Standard Characteristics for Transportability
(Win, 1997)*



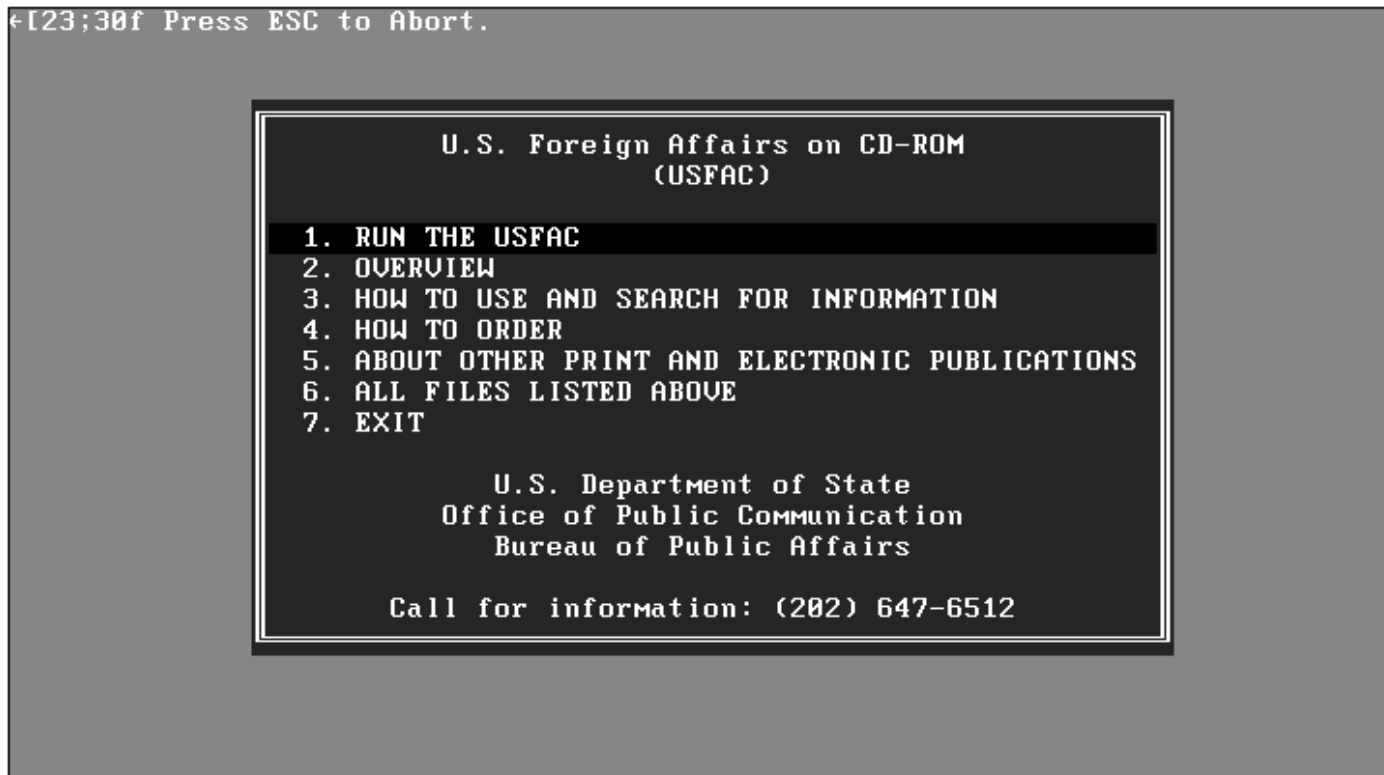
NIST Data Preservation and Migration Testbed

GPO Application: U.S. Coast Guard Seamanship (Win, 1999)

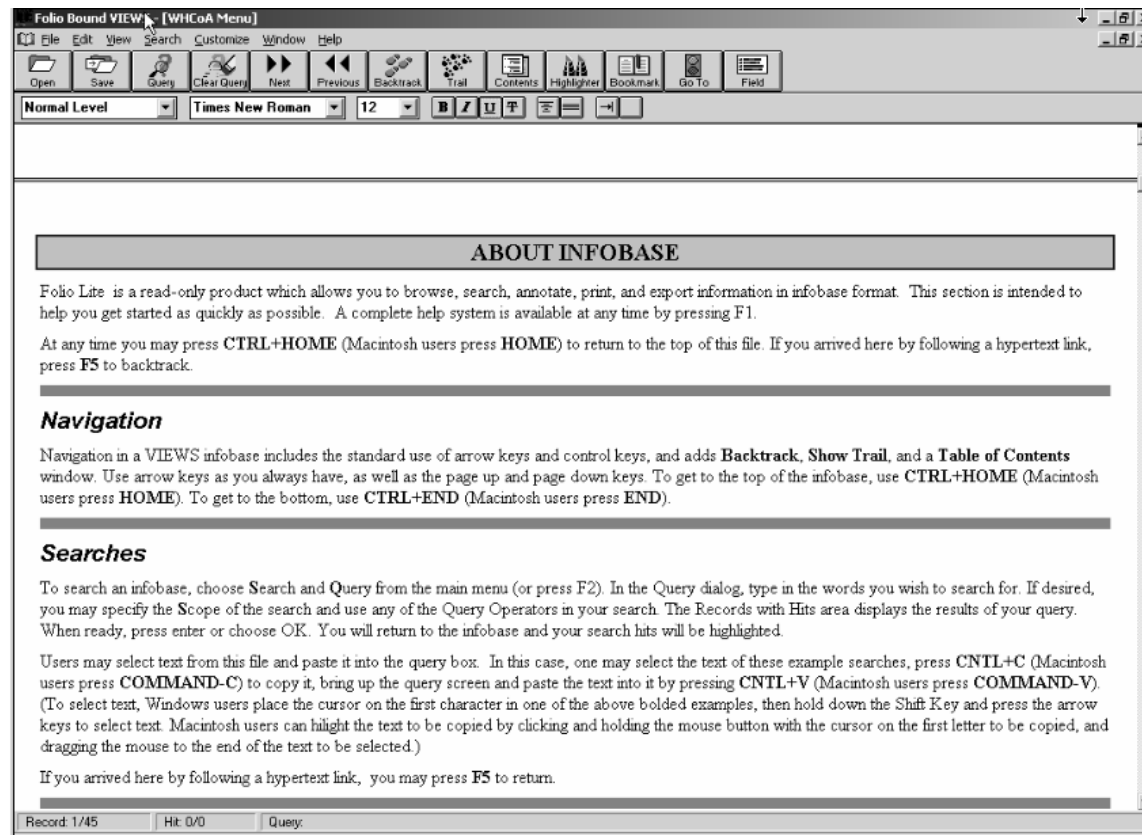


NIST Data Preservation and Migration Testbed

GPO Application: USFAC (DOS/Win, 1997)

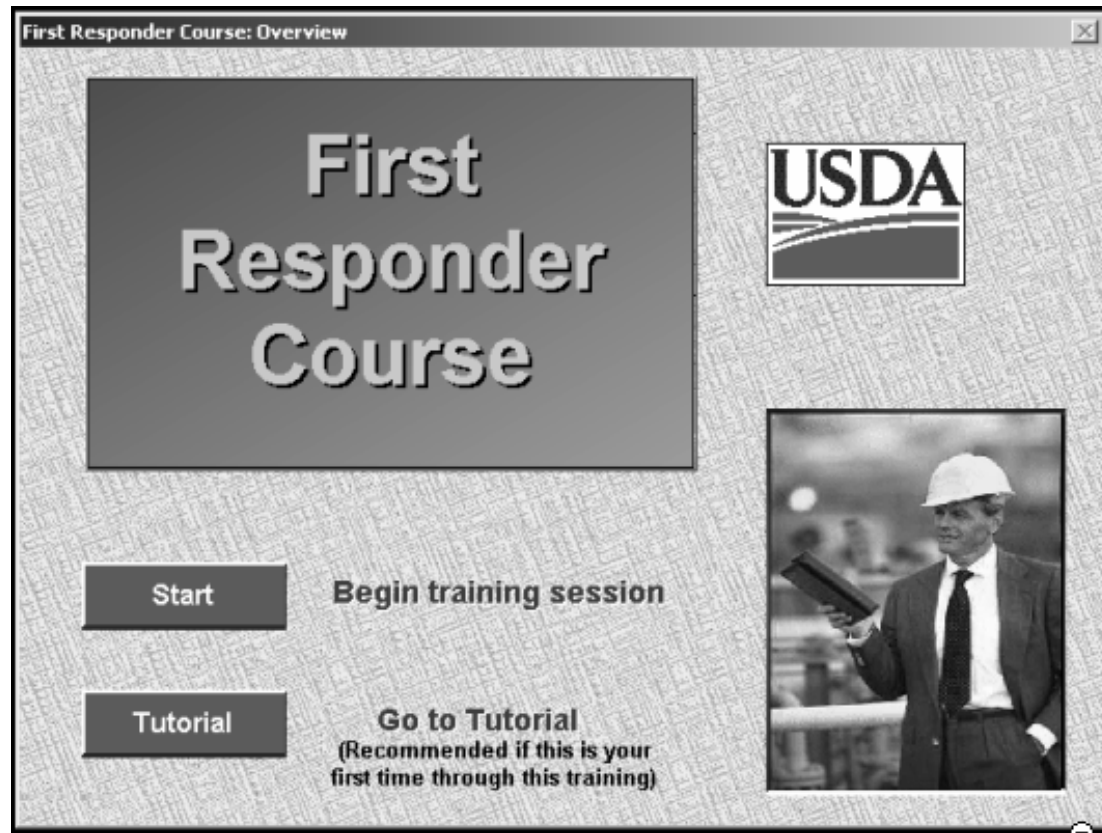


NIST Data Preservation and Migration Testbed *GPO Application: White House Conference on Aging (Win, 1995)*



NIST Data Preservation and Migration Testbed

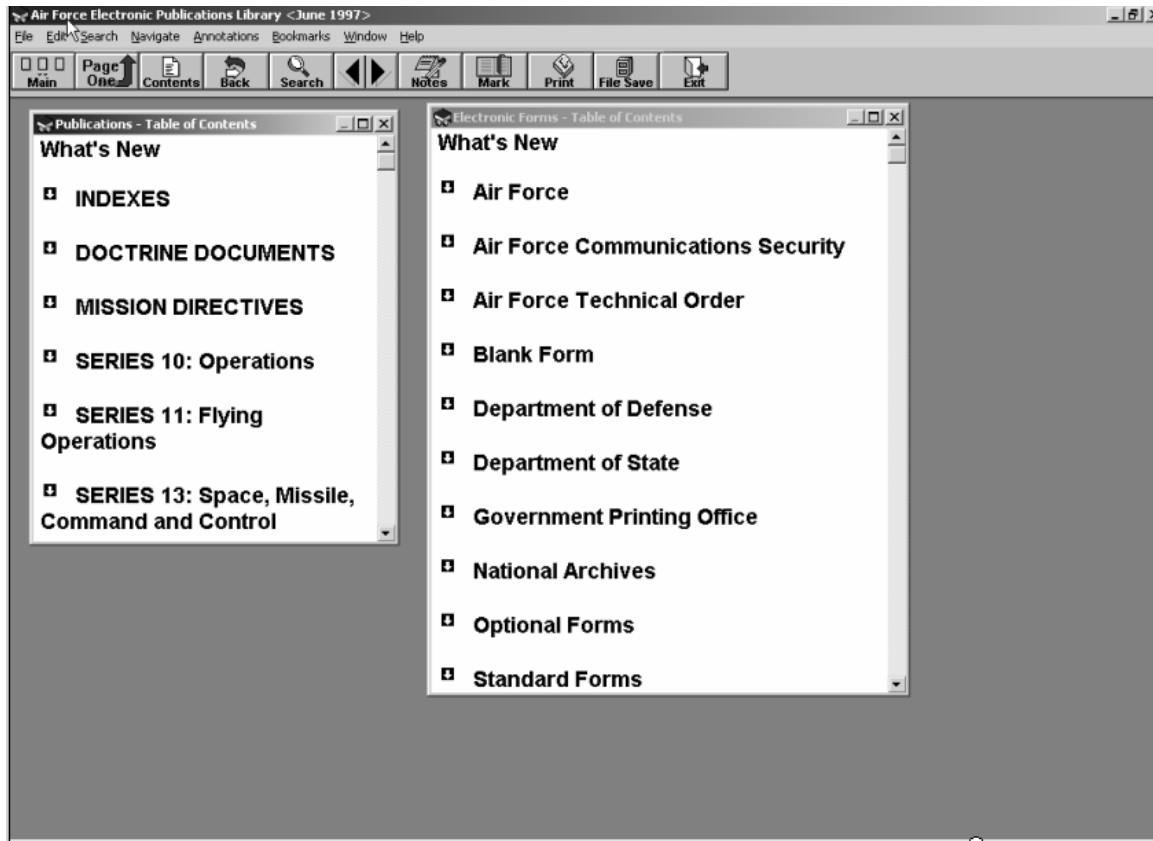
GPO Application: First Responder Course (USDA) (Win, 1999)



NIST Data Preservation and Migration Strategy: Virtualization

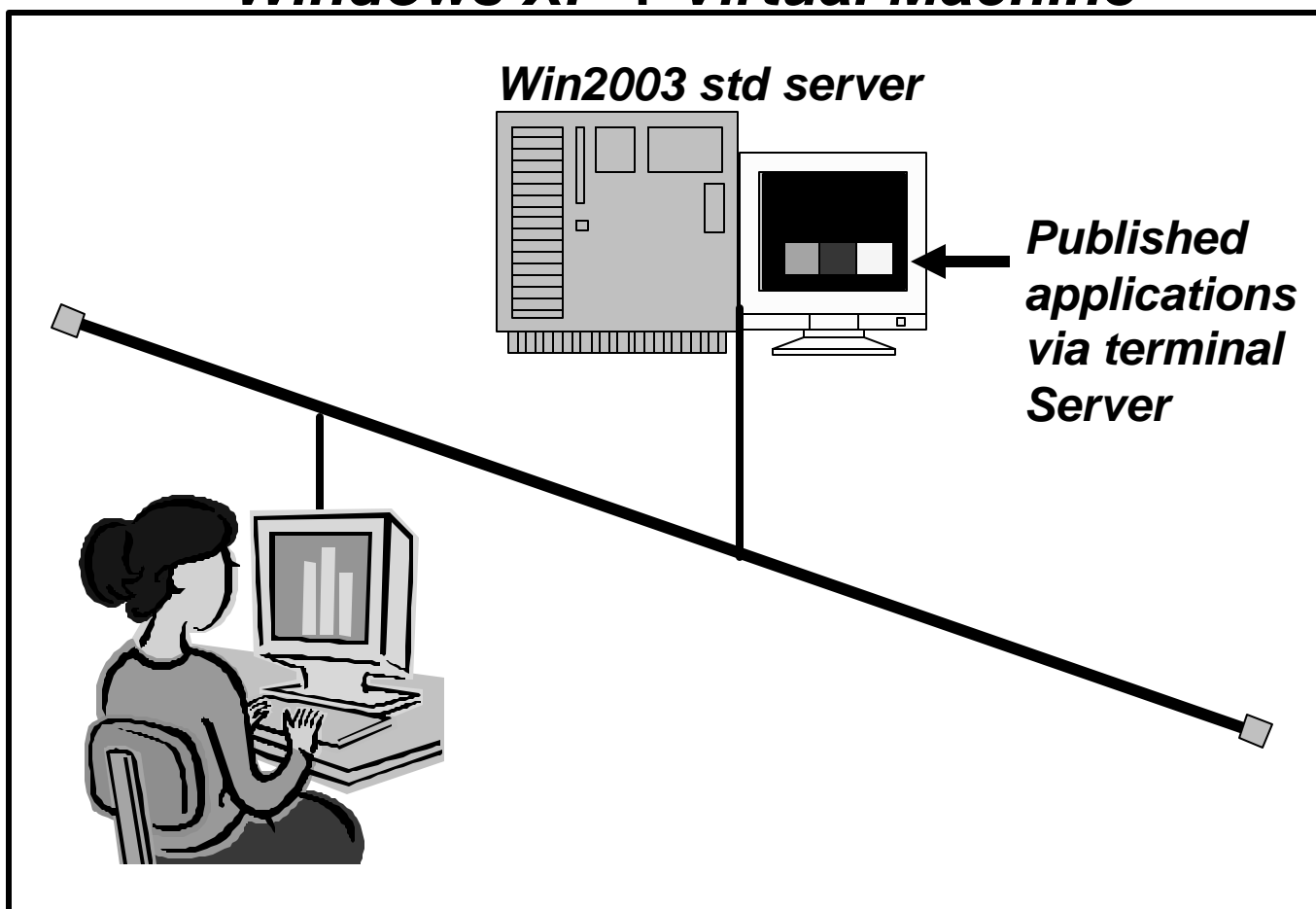
NIST Data Preservation and Migration Testbed

GPO Application: Air Force Electronic Pub Library (Win, 1997)



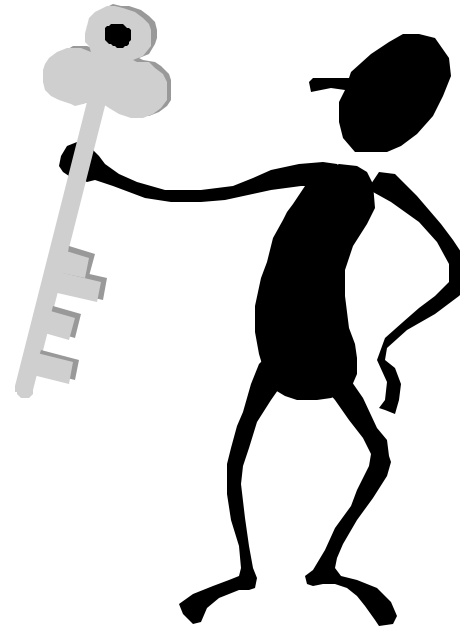
Demo:

Windows XP + Virtual Machine



Conclusion:

- Data preservation and migration are very hard problems
- No single solution fits all
- Approach: requires constant
Engagement
Analysis
Transformation



Future Direction:

- Areas of focus:
 - Legacy system, multimedia-based preservation and migration.
 - Apply OAIS (Open Archival Information System) ref. model.
 - Apply ISO/IEC MPEG-21 standard implementation.
- Collaborate with other agencies:
 - to compile and understand general data preservation and migration requirements via GIPWoG's DPM meetings.
 - NIST is a technology agency and happy to implement the DPM Test Bed.
- Deliverables:
 - Integrated best practices into a "testbed" environment.
 - Validate and benchmark the "testbed".
 - Generate standard best practice guidelines.
- Others...

