



Promoting Access to R&D Results

**Walter L. Warnick, Ph.D.
Director**

**DoD Independent R&D
Technical Coordination Group Meeting
July 17, 2003**



Information Management Critical to DOE's R&D Mission

OSTI's Mission: To collect, preserve, disseminate, and leverage the scientific and technical information (STI) resources resulting from DOE's annual \$8.5 billion R&D investment

Purpose: To provide national and global STI for use by DOE, the scientific research community, academia, U.S. industry, and the science-attentive public to expand the knowledge base of science and technology, while safeguarding national security interests

COLLECT, PRESERVE, DISSEMINATE, LEVERAGE STI

*** SAFEGUARDING NATIONAL SECURITY INTERESTS ***



OSTI fulfills essential requirements on behalf of DOE:

Tracking R&D Investment: To ensure U.S taxpayers' investment in research is accounted for, OSTI tracks 10,000 current research projects funded by DOE per year and receives about 12,000 research deliverables per year in the form of R&D reports and other documents.

Collection and Preservation: OSTI has managed DOE's R&D results for over 55 years and maintains a central repository that includes:

- Over 5 million electronic citations of DOE and other energy-related research information,
- 1.5 million DOE-sponsored full-text reports, and
- Over 73,000 full-text reports publicly accessible electronically



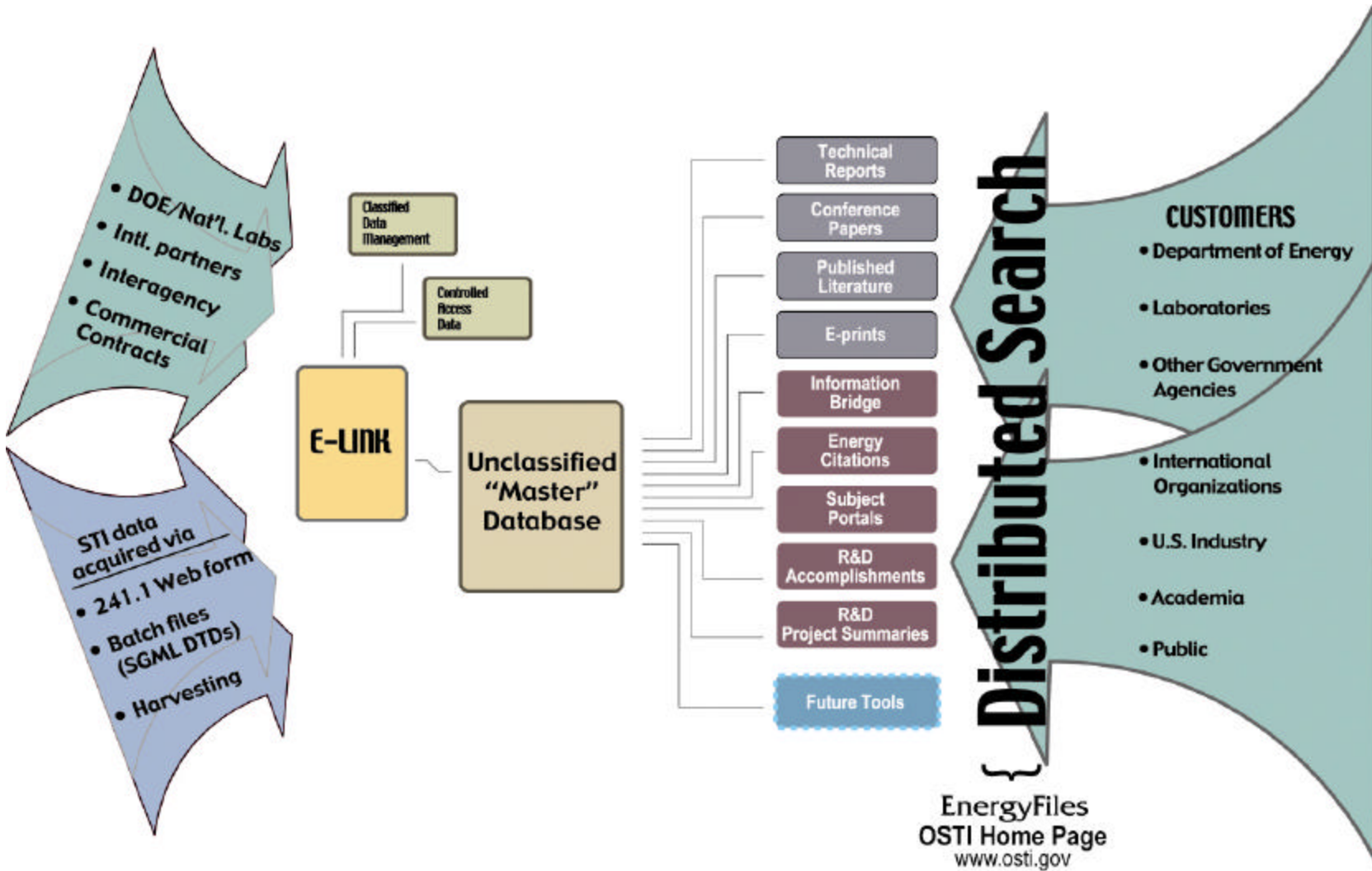
- OSTI coordinates the Department-wide STI Program, which involves Headquarters Program Offices, Field Offices, National Laboratories, other DOE facilities, and grantees.
- Collaboration enables us to link to the distributed STI full-text documents across the DOE complex via URLs in metadata.





www.osti.gov

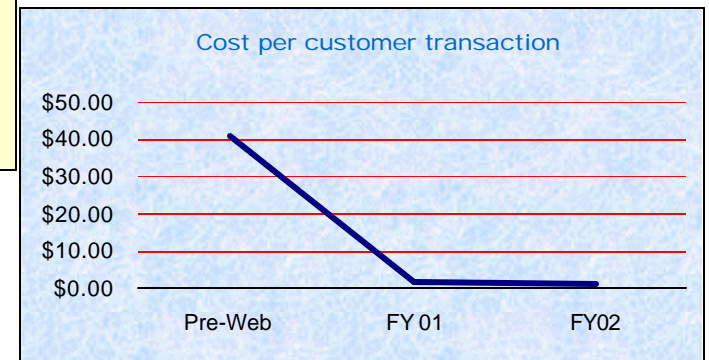
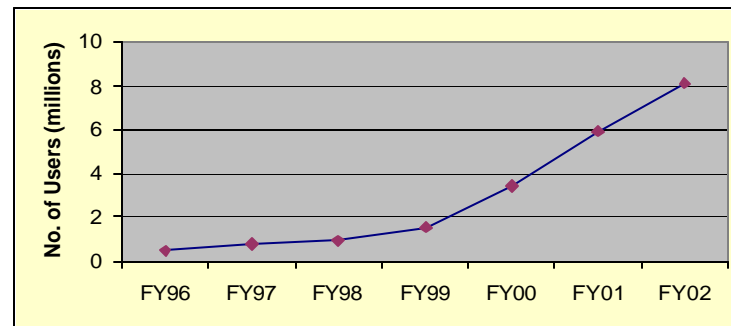
Overview of OSTI's Information Process





Information Access and Retrievability: OSTI has developed information tools covering the ways that scientists disseminate their research results: Journal literature, E-prints, and DOE's technical reports

With These Tools,
8.1 Million User
Transactions in FY
2002: A 35% Increase
from Previous Year



Efficiency and Effectiveness: Streamlined data requirements and use of new technologies in the R&D tracking and deliverables reporting system have reduced costs and time.



- Collaboration with DOE STIP partners to implement STI policy and best business practices
- Partnerships with other government agencies and institutions
- Development, delivery, and maintenance of information tools and services for a variety of audiences
- Expertise and tools to help solve information management challenges

DOE STIP OGA'S PUBLIC ETDE INIS EXPERTISE



Innovative information technology to offer comprehensive access at the desktop

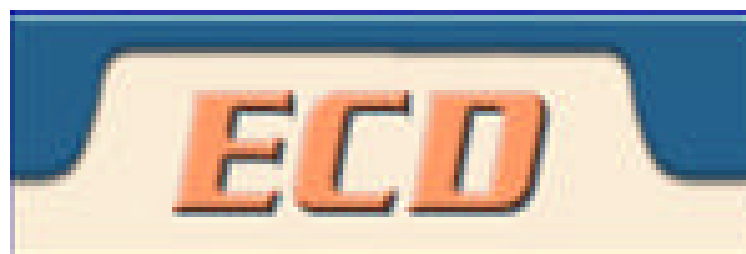
Access to the primary types of literature used by scientists, academia, and science-attentive citizens: journal literature, technical reports or gray literature, and e-prints

Advanced tools to assist in mining, assimilating and manipulating information





- Bibliographic citations from AEC, ERDA, and DOE
- Released October 2001
- DOE Version – 4 million citations
- Public Version – 2 million citations
- One-stop source for 1948 forward
- Links to electronic full text if PURLs are provided



www.osti.gov/energycitations (Public)

www.osti.gov/doeeed (DOE)

1940's

1950's

1960's

1970's

1980's

1990's

2000's

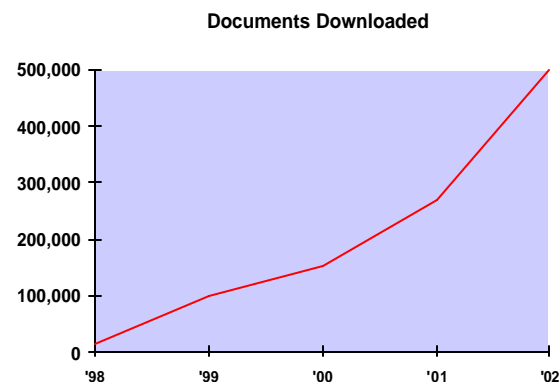


- Over 73,000 publicly available full-text DOE research reports and their bibliographic records and abstracts from 1995 forward
- Provides over 5.7 million full-text pages in an organized, searchable format
- Released April 1998
- Enables users to access, locate, search, retrieve, and download full-text and/or bibliographic information from their desktops
- Public access via partnership with GPO
- 500,000 full-text documents downloaded in 2002



www.osti.gov/bridge (Public)

www.osti.gov/doebridge (DOE)



FULL-TEXT FROM DOE NATIONAL LABS AND GRANTEEES



**E-PRINT
Network**

Research Communications for Scientists and Engineers

www.osti.gov/eprints

- A vast, integrated network of electronic scientific and technical information created by scientists and research engineers
- A gateway to over 10,000 Web sites and databases worldwide, containing e-prints in disciplines of interest to DOE.
- Specialized tools and services to facilitate the exchange and use of information
 - a unique Deep Web search capability combining full-text searching of over 150,000 PDF documents residing on e-print Web sites with a distributed search across 28 large e-print databases
 - access to e-prints and related scientific Web sites organized by subject category
 - an Alert Service that provides automatic updates on new e-prints in areas of interest
 - links to over 1,000 relevant scientific societies
- Released July 17, 2003

PROVIDES EARLY ACCESS TO RESEARCH FINDINGS



- Long-standing agency-to-agency partnerships with GPO, NTIS, and DTIC
- DOE representative of CENDI and Science.gov Alliance

- Science.gov enables you to search two kinds of interagency information – selected Web sites and databases of technical reports, journal articles and other published materials. These can be searched simultaneously or separately.



- Science.gov Alliance of 15 information offices from 11 major science agencies.
- OSTI is DOE's representative, and we are the "home" of science.gov web site.
- Web portal indexes over 1700 resources.
- Deep Web searching of 30 databases is a key feature, developed by OSTI to offer greater access to R&D results.

Science.gov aims to bring the substantial resources of the federal science and technology enterprise together, in one place. Working together, federal agencies have assembled countless pages of government research, data, and reports. The site is a great example of e-government in action.

-Dr. John H. Marburger, Director, Office of
Science and Technology Policy

Alliance members are: Department of Agriculture, Department of Commerce, Department of Defense, Department of Education, Department of Energy, Department of Health & Human Services, Department of Interior, Environmental Protection Agency, Government Printing Office, National Aeronautics & Space Administration, National Science Foundation

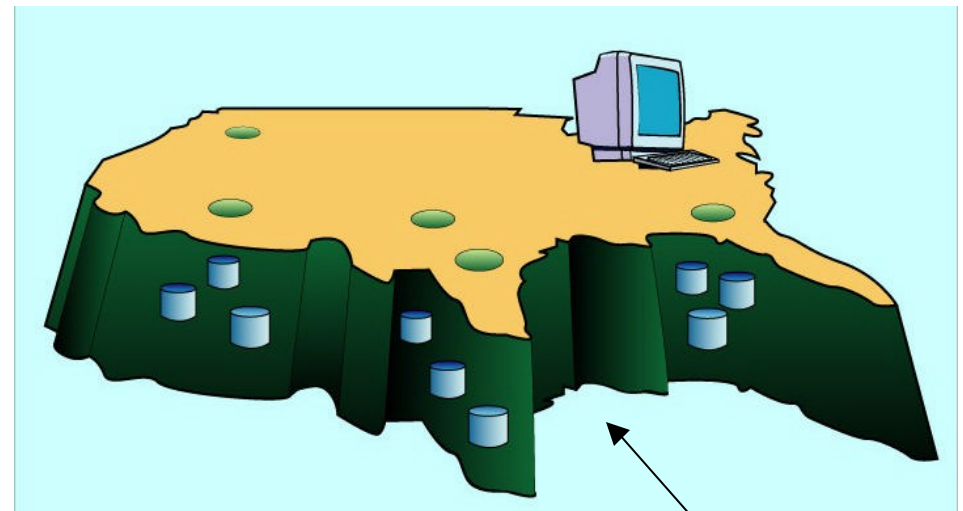


www.osti.gov

Vast Government R&D in Databases

R&D Results – *The Government’s Most Important R&D Offerings* - Reside in Databases

- In vast databases of gray literature
- In scientific peer-reviewed journal literature
- Such databases are generally not searchable via Google



Databases in Deep Web



www.osti.gov

The Web Has Two Parts

Surface Web

Accessible by
Traditional Search
Engines; e.g., Google

Deep Web

Not Accessible by
Traditional Search
Engines

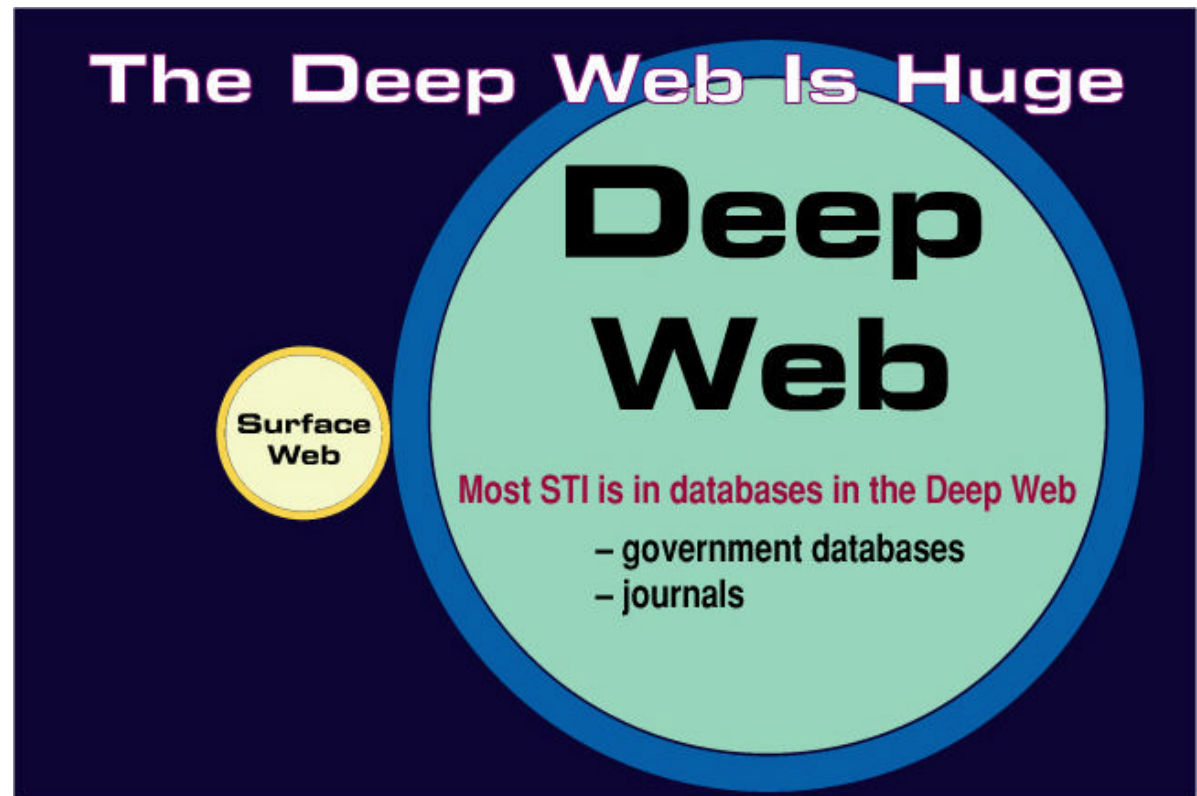
Surface Web and Deep Web
Intersect at Pages Where Patrons
Launch Searches of Databases



www.osti.gov

Getting to the R&D Results

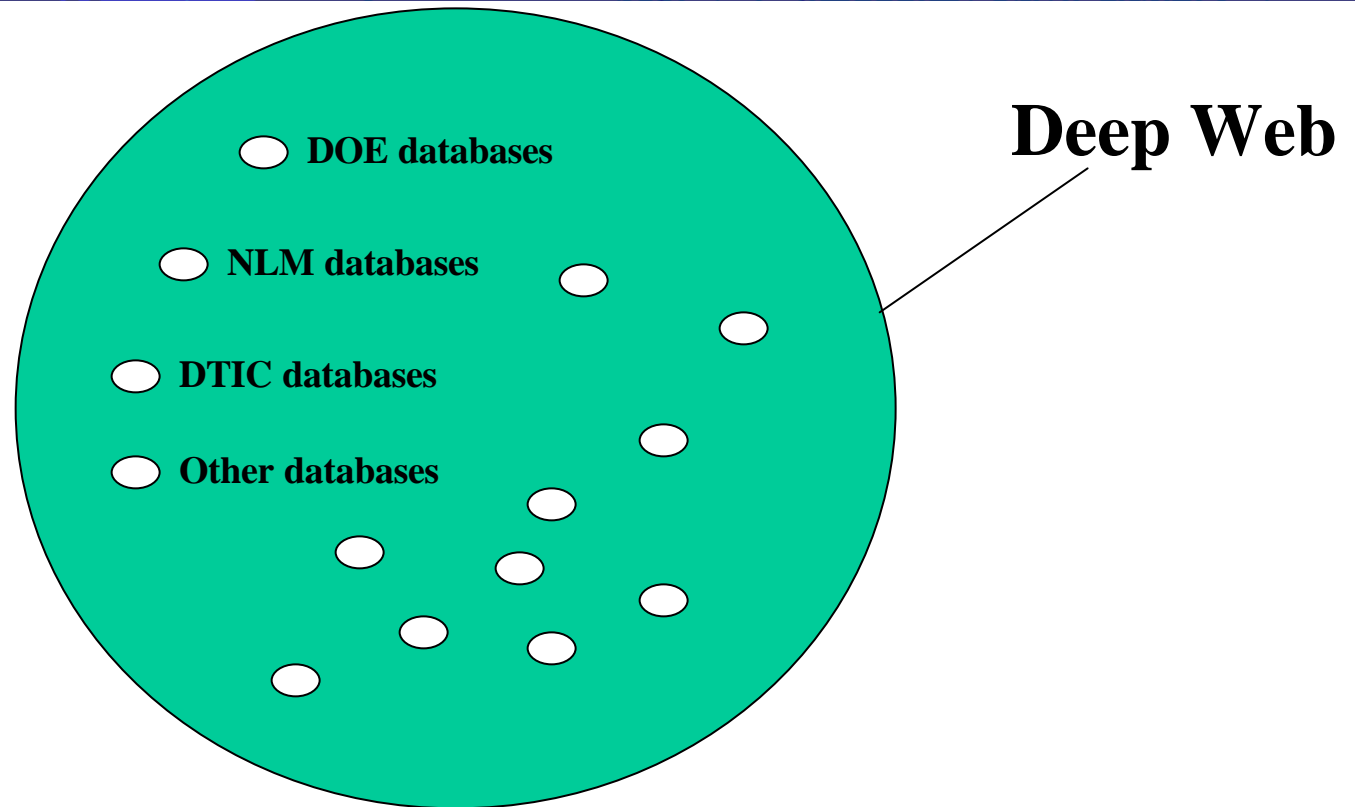
- **The Surface Web is accessible to popular search engines such as Google.**
- **But less than 1% of government R&D results are currently accessible to crawlers.**





Deep Web Before Science.gov

www.osti.gov



Databases for CENDI agencies were scattered around the Deep Web



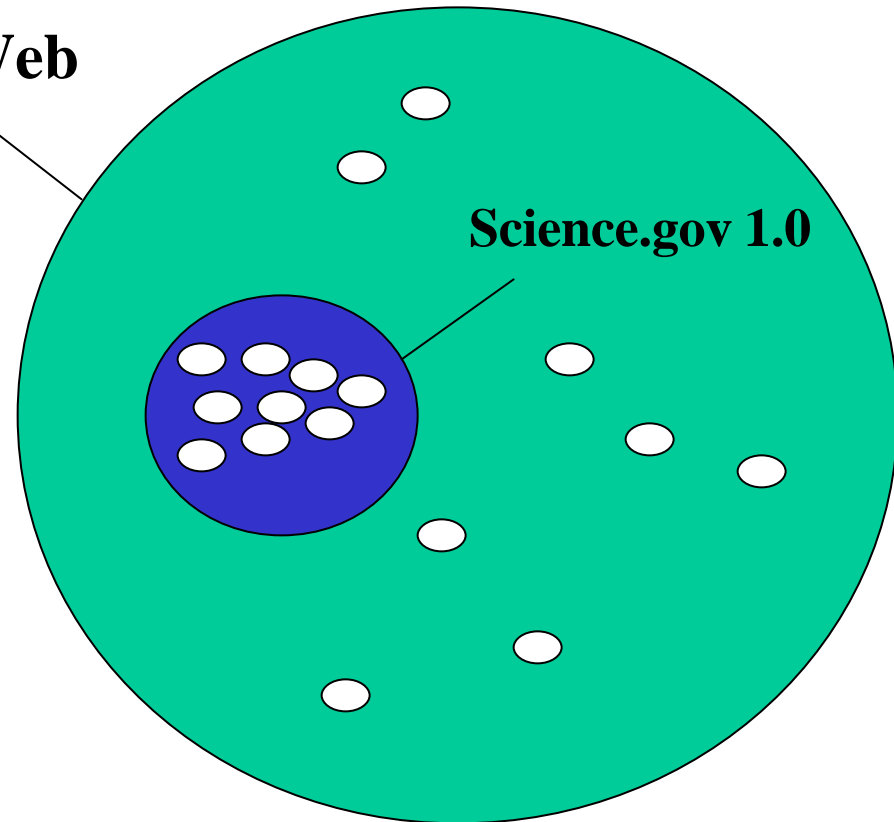
www.osti.gov

Deep Web After Science.gov 1.0

(December 2002)

- **Science.gov 1.0 enables cross-database searching**
- **Searches still lack relevancy ranking**

Deep Web





www.osti.gov

Current Limitations of Science.gov

- **Selecting databases: Has a key database been overlooked?**
- **Sorting through hits: Which results are the most relevant to the individual patron?**

Information that is hard to locate is information that is little used.

“Articles freely available online are more highly cited. For greater impact and faster scientific progress, authors and publishers should aim to make research easy to access.”

Nature, 411:6837, p. 521 (2001)



www.osti.gov

Science.gov 2.0 with Relevancy Ranking

Solution in principle:

- Retrieve everything: not only bibliographic data, but also full-text content within the Deep Web databases.
- Then, relevancy rank the results **BEFORE** delivery to the patron.
- Eliminate limitations of selecting specific databases and reviewing unranked hits.



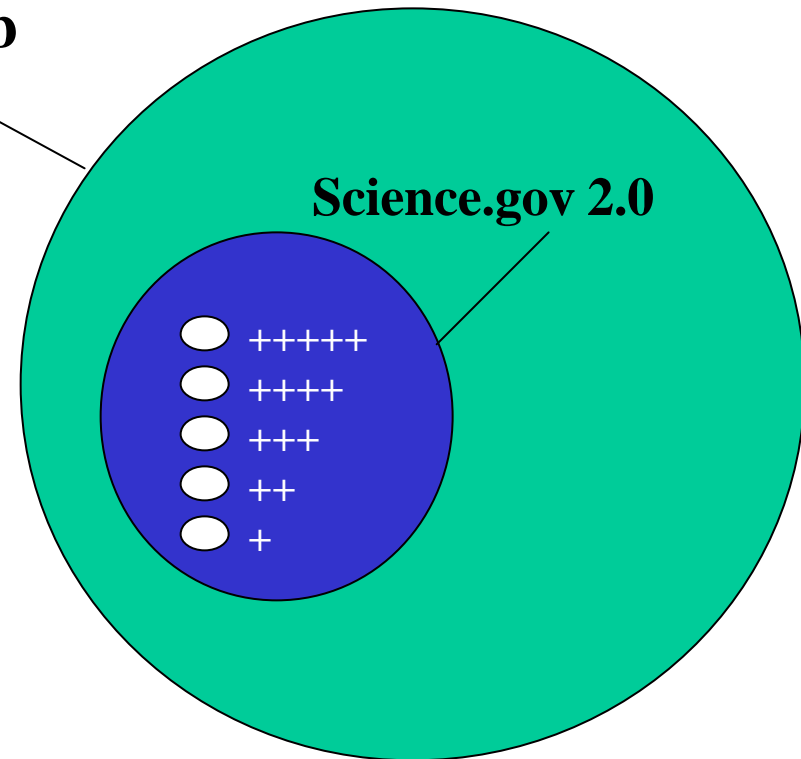
www.osti.gov

Deep Web After Science.gov 2.0

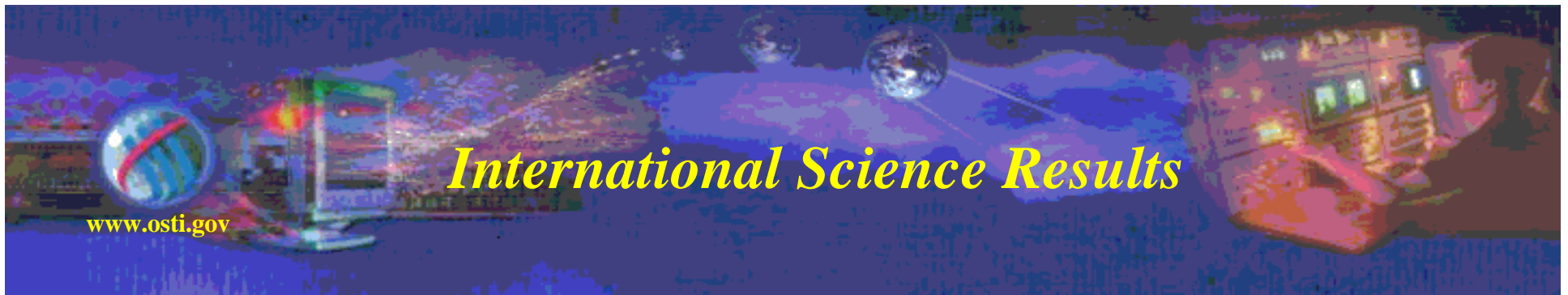
(May 2004???)

Deep Web

- Science.gov 2.0 will add relevancy ranking to cross-database searching



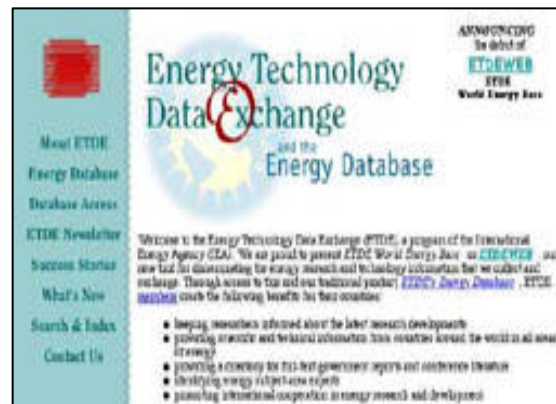
Cross-database searching with relevancy ranking



- OSTI acquires 80,000 summaries of foreign research results per year from the Energy Technology Data Exchange (ETDE) and the International Nuclear Information System (INIS) representing more than 100 countries; now available to the U.S. partners via ETDEWEB
- OSTI is working to bring full-text capability to these sources of information
- OSTI is the U.S. Representative/Liaison for ETDE and INIS, serving on the ETDE Executive Committee and as the U.S. INIS Liaison Officer



www.iaea.or.at/programmes/inis



www.etde.org



www.etde.org/etdeweb



DOE's Laboratory Directed Research and Development (LDRD) Program

- DOE major laboratories do not have authority for independent R&D comparable to DoD.
- The LDRD Program at DOE multi-program national laboratories has existed since the mid-70's, but significant funding changes occurred in FY 2000.
- The allowable funding level for LDRD was reduced from 6% of a laboratory's total operating budget to 4%. FY02 projects at 11 DOE laboratories totaled about \$335 million.
- LDRD projects are defined as research and development work of a creative and innovative nature selected by the director of a laboratory for the purpose of maintaining the vitality of the laboratory.
- LDRD provides DOE laboratory directors with flexibility to invest in long-term, high-risk, and potentially high-payoff research activities that enhance the science and technology capabilities of the laboratories.
- The Department's requirements for LDRD are in DOE Order 413.2A. The reporting requirements include providing OSTI with reports for completed projects.