Office of Scientific and Technical Information

U.S. Department of Energy
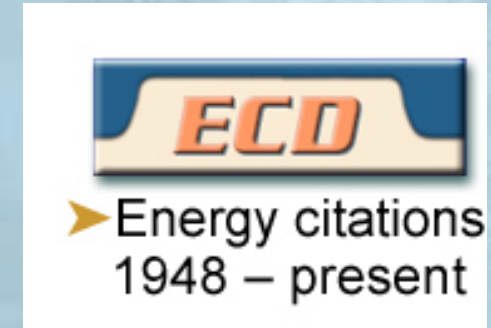
The Search for Hidden Treasure:

R&D Results in the Deep Web

Karen J. Spence

www.osti.gov

October 2004

The Partnership Continues

OSTI's Corporate Position

❖ OSTI's unique niche in the Department and in the global information community is the result of:
   ❖ OSTI's Departmental Program responsibility,
   ❖ A vast information resource that does not exist anywhere else, and
   ❖ OSTI's innovative information technology applications.
❖ Strategic use of this corporate position has resulted in providing the public state-of-the-art resources for finding DOE information through *GPO Access*.

Electronic Services Partnership

Then and Now

1998

- 23,323 full-text documents
- 23,323 bibliographic records
- 50,000 downloads annually
- EDB data format
- Partnership Web products
    - Reports Bibliographic Database
    - Information Bridge

2004

- 85,000 full-text documents
- 2.2 million bibliographic records
- 400,000 downloads annually
- Dublin-Core data format
- Partnership Web products
    - DOE Information Bridge
    - Energy Citations Database
- Additional Ancillary Services
    - GrayLIT Network
        - ~ 343,000 Federal documents
    - Federal R&D Project Summaries
        - ~ 500,000 Federal records

Advancing Public Access to Science through Federal Agency Collaboration
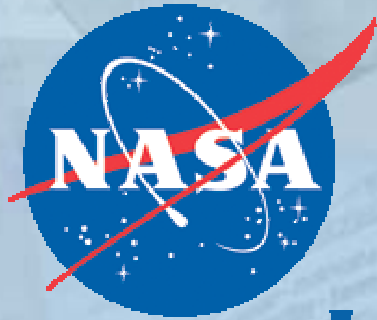
Customized Data Services for FDLP

OSTI is pleased to offer new technology to customize services to meet the individual needs of the FDLP.

- ❖ Data can be delivered in customized formats

- ❖ Data can be delivered in specific subject areas

- ❖ Data can be provided to meet specific library needs

Advancing Our Partnership

What is Science.gov?

An E-gov success!

With **National Archives and Records Administration** Support
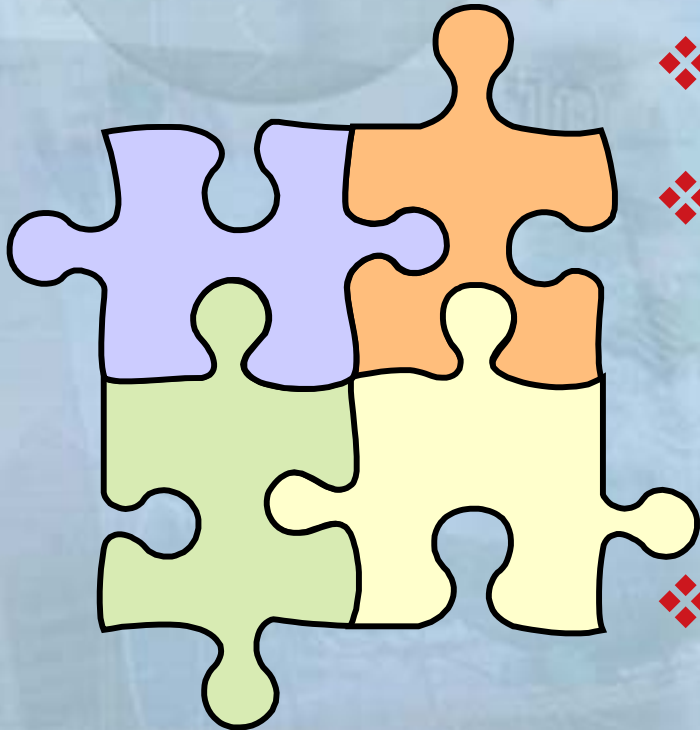
Science.gov Is Also...

- ❖ A cross-agency gateway to one of the government's most valuable resources — its R&D

- ❖ An Alliance of 17 information offices from 12 major science agencies

- ❖ A Web portal indexing over 1,700 resources

- ❖ A search tool, developed by OSTI, for 30 deep Web databases

- ❖ A creative way to maximize the U.S. return on its R&D investment

Behind the Scenes…

- ❖ **DOE convened principals**

- ❖ **To capitalize on new technology**

- ❖ **Vision: a new science information infrastructure**

Shared Premise

❖ **Web is the tool of choice**

❖ **Each science agency has information & services to fulfill its mission**

❖ **Science is not bounded by organization or geography**

❖ **Interagency collaboration is needed**

Science.gov Alliance is Born

- ❖ **Interagency science portal endorsed**

- ❖ **Extraordinary voluntary collaboration - without appropriated funding or legislative mandate!**

- ❖ **Supports e-Gov initiative**

- ❖ **Organizes major and often hard-to-find information collections**

Agency "Pot Luck"

❖ **Agencies brought to the Internet table their unique information specialties**

❖ **Arranged resources on a simple Web site**

❖ **Powered by impressive collaborative intellectual work, computer systems, and software**

STRENGTHENING THE PUBLIC INFORMATION INFRASTRUCTURE FOR SCIENCE

Report of the Workshop

Held April 18 - 19, 2001

At the National Institute of Standards and Technology

Issued by the Science.gov Alliance

First ".gov" Street in America

Oak Ridge, Tennessee, renames roadway

Nov. 7, 2003

Science.gov 1.0: Launched 12/02

Initial Features:

❖ Browse by Subject (1700 sites); broader/narrower categories listed

❖ Website search

❖ Deep Web search full content of 30 databases using Distributed Explorit

❖ Search results sorted by database with links to the websites

❖ Special section on Science Internships and Fellowships; special link with logo to Kids.gov; Featured Sites

Science.gov 2.0: Launched 5/11/04

Current Features:

- ❖ QuickRank – a title-based mechanism for quickly ranking returns based on a two-step relevancy algorithm
- ❖ Results filtered/sorted by relevancy; can be separated by database
- ❖ Fielded or Advanced Search option; Author, Title, Full Record, Data Range
- ❖ Expansion of the number of databases searchable at one time: 10 databases expanded to 30
- ❖ Search page organized by broad subject
- ❖ "MORE" button for additional returns displayed
- ❖ Regulations section provides metasearch of CFR and the Federal Register

Science.gov 3.0:
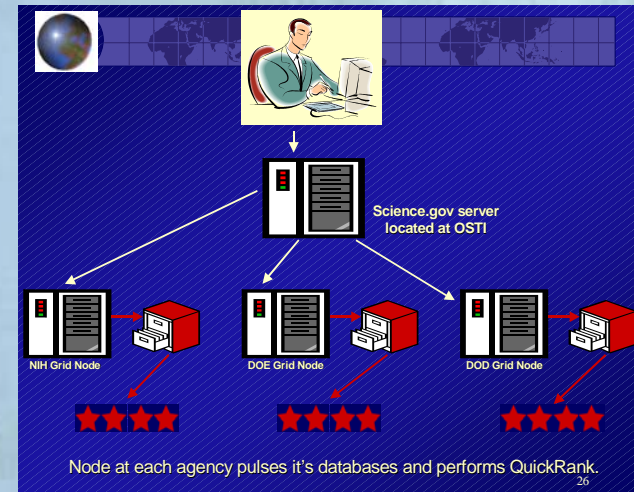# Proposed Launch Summer 2005

Proposed New Features:

- ❖ **MetaRank – a metadata-based ranking of returns based on a complex relevancy algorithm; may sort returns by database, by QuickRank, by MetaRank**
- ❖ **Fielded search will be integrated into the Advanced Search**
- ❖ **Boolean search will be integrated into the Advanced Search**
- ❖ **Distributed Alerts, a site-wide alert system for notifying users of new material added in their areas of interest; slated as first new function for FY05**

Science.gov 4.0:
Proposed Launch 2006

Proposed New Features:

❖ **Deep Rank (full-text relevancy ranking)**

❖ **Grid computing; parallel processing**

Node at each agency pulses it's databases and performs QuickRank.

The DeepRanked results are then merged at the OSTI server and rapidly returned to the user.

Yahoo/Google Collaboration

❖ **Partnership with commercial search engines**

❖ **Objective of DOE R&D visibility**

❖ **Success of project demonstrated through usage statistics**

What's New

Rolling Window Data Service

❖ **DOE and GPO Started the Rolling Window Data Service as a Pilot Project**

❖ **Customized Data Service**

❖ **Ready to Broaden Availability**

Rolling Window Data Service

Functions/Features

❖ **Manage Profiles via a Web-Based Interactive, User Profile Development Application**

❖ **Customized Email Alerts**

❖ **Single Window Requests**

❖ **Configurable Date Ranges**

❖ **XML, HTML, and Counts of Records Matching Specified Criteria**

Rolling Window Data Service

Uses

- ❖ **Advertise "What's New" in Your Library**

- ❖ **Capture Metadata to Include in Your Library Catalogs**

- ❖ **Other**

Log In Screen

Partners in Innovation

Rolling Window Data Service

DOE/GPO Alert Services

Log In   Single Window Request   Help

Log In to Rolling Window Data Service

User Name:	test
Password:	******

LOGIN

The **OSTI Rolling Window Data Service** is a web-based application that allows interactive user profile creation. Users can receive individual email alerts based on a configurable date range (window) of DOE's bibliographic records and full-text links (when available) for documents/records made available to the public via the Internet based on data matching their specific interest and have those records delivered to them in the alert. For more information, contact tacketts@osti.gov. Download the User Guide (Adobe PDF).

Manage Profile Screen

Email Alerts Screen

Single Window Request Screen

Rolling Window Data Service

Status

❖ **Now Open to all GPO FDLP Libraries**

❖ **URL: www.osti.gov/window**

❖ **Registration Required**

❖ **Initial Username Guest**
      Password Anonymous

❖ **Contact: Susan Tackett, DOE**
   Office of Scientific & Technical Information
   Email: tacketts@osti.gov