Office of Scientific and Technical Information

Strategic Plan FY 2009–2013

U.S. DEPARTMENT OF ENERGY

Office of Science

OSTI works to ensure that U.S. scientists and engineers, educators, entrepreneurs, and the public find the right science information fast.

The mission of the Office of Scientific and Technical Information (OSTI) is to advance science and sustain technological creativity by making research and development (R&D) findings available and useful to U.S. Department of Energy (DOE) researchers and the public.

OSTI is headquartered within the DOE Office of Science and has operations in Oak Ridge, Tennessee.

As a matter of convenience to simplify this Strategic Plan, the term "science" is used throughout to mean all fields of science, technology, and development done by scientists, engineers, and technicians that work on projects supported by the DOE. To that end, the mission of OSTI is to capture and preserve the information generated by these and related projects and make it available in a mode that permits rapid, efficient and convenient access for those who need this information. This helps to advance their work in support of the missions of the DOE for the benefit of the American people. This process will contribute to the goal of accelerating the spread of knowledge that in turn accelerates the advancement of science.

Office of Scientific and Technical Information

Strategic Plan FY 2009–2013

Table of Contents

OSTI Corollary: Accelerating the sharing of scientific knowledge accelerates the advancement of science.

Director's Message

Science leads to better lives for people and underpins both a strong economy for our Nation and enhanced national security. What if we could speed the pace of science so that tomorrow's solutions arrive sooner? We would thus raise our standard of living, improve our quality of life, and increase the security of our Nation. These grand outcomes can be achieved by committing to openly sharing scientific research data and spreading the world's great scientific discoveries faster.

At OSTI, we are achieving this science knowledge imperative. Founded on the principle that science progresses only if knowledge is shared, OSTI is well positioned and prepared to speed the pace of science. The OSTI Corollary—accelerating the sharing of knowledge accelerates the advancement of science—takes our founding principle to the next level. This corollary is the premise behind everything that OSTI does today. The overarching goal is to ensure that those working or interested in scientific and technical fields get the right information fast. Today, OSTI manages a complex suite of tools and services to ensure that quality science content is readily findable for citizens everywhere—and that information is available in forms suitable to researchers' and citizens' needs. But we need to do more, as described in this Strategic Plan.

Historically, OSTI was established to capture the results of the work by scientists, engineers, and technicians working for the Atomic Energy Commission *(AEC)* at its facilities and at universities and industrial organizations. OSTI was responsible for preparing, reproducing, and distributing this information to those requiring it to advance certain aspects of our national security and economic growth. As the AEC evolved into the Energy Research and Development Administration *(ERDA)* and then into the Department of Energy, the information scope expanded from nuclear science to fossil, solar, and renewable energy. However, it was the subsequent development of information technologies and the capability to store massive amounts of information in digital form which led to a revolutionary change in the rate of the spread of knowledge and the access thereto. The National Library of Medicine *(NLM)* was one of the key pioneers in stimulating this revolution. Today millions of people around the world use the NLM website to learn about

diseases that might be afflicting them and what progress has been made to develop cures for their ailments. This same service also makes it possible for medical researchers to have essentially instant access to the latest research results developed almost anywhere in the world. This has resulted in an acceleration of knowledge in the life sciences that has benefited everyone.

With a similar approach in mind for a physical science information infrastructure, OSTI convened eight workshops to explore how it might better serve those for whom it has a mission responsibility by providing the same kind of accessibility to information needed by those working directly or indirectly for the DOE *(see Appendix 1, Strategic Planning: Basis for Action, p. 54)*.

OSTI made major changes in mode and manner of capturing information from paper to electronic formats in the late 1990s to provide information rapidly in forms that the public can readily find and use. The new electronic infrastructure reduced the time required to disseminate knowledge among scientists and engineers from weeks down to days, hours, minutes and in some cases seconds. The workshops held in recent years have stimulated ideas and resulted in improvements to systems, products, and services so that patrons can find the precise information needed from the staggering volume available.

OSTI is now poised and prepared to advance to the next steps of improving the capture and dissemination processes, creating an unprecedented level of openness in government. This Strategic Plan outlines steps that would improve the competitiveness of our scientists and engineers by providing them with fast and efficient access to national and international bodies of knowledge relevant to their fields. OSTI has made tremendous strides in achieving goals that are included in its mission, yet there are significant opportunities for improvements and many challenges that remain.

We believe that achieving these goals will accelerate the spread of knowledge in a manner that will benefit our Nation. Access to the rapid developments in many fields of science and technology will help to restore science to its rightful place. Failure to gain such access will put us at a serious disadvantage in our rapidly changing world.

Executive Summary

The web* should work better for science. Invented by scientists in 1989 to facilitate the sharing and transparency of intellectual capital, today the web contains major gaps that prevent it from realizing its full potential to accelerate the spread of science knowledge and, thereby, speed the pace of important discovery.

OSTI has set out to fill certain gaps in the web *(see Gaps To Be Filled, p. 19)* to fulfill its vision for global scientific discovery and its mission responsibilities for DOE. Using the foundation of web-enabled tools developed by OSTI in recent years, and building on stakeholder input, this Strategic Plan describes OSTI's potential to substantially improve the ability of scientists and engineers engaged in work related to the missions of the DOE to gain ready access to bodies of knowledge with greater speed and improved outcomes. Beyond the value to DOE, implementing the elements of this Strategic Plan will also contribute to our Nation's scientific and technical enterprises in a manner that enhances our national security and prosperity. None of this is to deny that progress has been made. More science is available on the web today than has ever before been available to the public.

OSTI, an early adopter of web technologies as they evolve, has championed an aggressive effort on a series of fronts to make authoritative science information more efficiently available to researchers and the public alike. In fact, the overarching OSTI strategy, briefly stated, is to use whatever tools necessary to make the web work better for science. To this end, OSTI developed a suite of cutting-edge web tools for sharing e-prints, technical reports, conference proceedings, and other forms of scientific and technical information *(STI)*. Because each form of STI comes from a distinct source, each requires a distinct pathway *(see Appendix 2, Regardless of Form, OSTI Search Tools Find Scientific and Technical Information, p. 55)*. The progress that OSTI is making is possible only by the most judicious use of resources.

Having these tools online, OSTI "federated" them so that all the databases could be searched simultaneously, in parallel, via a single query. The resulting tool is the

**The term "web" is used throughout the document to indicate the World Wide Web, the Internet, digitally networked environments, and other information technologies available.*

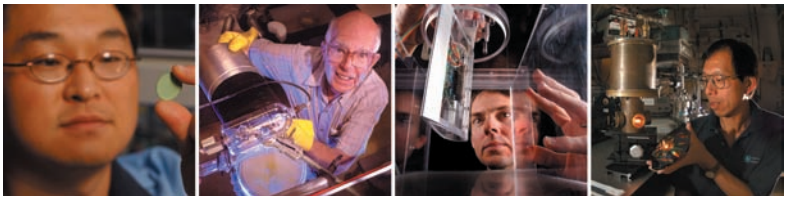
ScienceAccelerator.gov, which federates key DOE databases. But OSTI progress extends beyond DOE to include Science.gov, which federates U.S. federal agency science information, and WorldWideScience.org, which federates national databases and portals from around the globe. Differentiated from today's conventional search engines, these online gateways access only the most authoritative science information and research results available from government entity databases and websites. While OSTI has successfully advanced and deployed federated search technology, that technology is new and remains immature.

OSTI has utilized the resources available to it to provide information in ways that take advantage of new information technologies. Even so, the present circumstances offer an unprecedented opportunity to make dramatic improvements in the services OSTI provides consistent with its mission responsibilities.

This plan describes a set of six goals that would address identified gaps and take appropriate advantage of these opportunities in a cost effective and efficient manner. These goals, which are discussed in a later section of this plan, are:

1. Ensure Superior Access to and Preservation of Quality Scientific and Technical Content
2. Improve Education by Providing Enhanced Access to Scientific and Technical Knowledge
3. Bring the World's Science to U.S. Science and Research Communities
4. Increase Visibility and Understanding of DOE Contributions to Science
5. Contribute to National Security
6. Provide Specialized Support to DOE R&D Programs

By striving to make the web work better for science, OSTI will ensure that more of the world's scientific and technical information *(STI)* is available to researchers and the public to enable new scientific discovery.



"If I have seen further than others, it is by standing on the shoulders of giants."

Isaac Newton, 1676

Mission

To advance science and sustain technological creativity by making R&D findings available to DOE researchers and the public

OSTI supports the Office of Science mission to deliver the remarkable discoveries and scientific tools that transform our understanding of energy and matter and assists researchers to address DOE goals to advance the national, economic, and energy security of the United States.

OSTI works on the premise that speeding the sharing of knowledge, or "knowledge diffusion," will accelerate the advancement of science and thus accelerate benefits to DOE researchers, the Nation, and the world. OSTI works to achieve this by (1) making DOE R&D results openly accessible (2) preserving DOE R&D results for reuse, and (3) ensuring that DOE researchers and the public can find and access scientific discoveries from around the world.

U.S. scientists, researchers, and engineers will be at a distinct disadvantage if they do not have comprehensive access to the world's R&D knowledge. Research results of the present and past are the building blocks for future discovery.

The creation in 1947 of what is now OSTI signified a sea change in dissemination of government R&D. What was once shrouded in secrecy—basic nuclear science information—was made available and preserved for use for future discovery, thereby enabling free and open scientific inquiry.

The dissemination of scientific and technical information *(STI)* is a foundational mission for the Department, called out in the Atomic Energy Acts of 1946 and 1954, the Energy Reorganization Act of 1974, the Department of Energy Act of 1977, and, more recently, the Energy Policy Act of 2005 and the America COMPETES Act of 2007 *(see Appendix 3, Statutory Authorities, p. 60)*. The scientific and technical legacy of DOE and its predecessor agencies is preserved at OSTI, and the overwhelming majority of it is entirely open.

"The diffusion of such knowledge should help us stimulate new enterprises, provide jobs for our returning servicemen and other workers, and make possible great strides for the improvement of the national well-being." – President Franklin D. Roosevelt, November17, 1944, in his charge to Vannevar Bush, then the Director of the Office of Scientific Research and Development, to apply the experience of U.S. R&D war efforts to the "days of peace ahead."

"Openness will strengthen our democracy and promote efficiency and effectiveness in Government."

President Barack Obama, Memorandum for the Heads of Executive Departments and Agencies, January 21, 2009

Organizationally, OSTI is within the DOE Office of Science, a large supporter of basic research in the physical sciences in the U.S., providing more than 40 percent of total funding for this vital area of national importance. The Office of Science oversees research programs in high-energy physics, nuclear physics, and fusion energy sciences. OSTI fulfills a corporate function as stated in the Energy Policy Act of 2005:

> "The Secretary, through the Office of Scientific and Technical Information, shall maintain within the Department publicly available collections of scientific and technical information resulting from research, development, demonstration, and commercial applications activities supported by the Department."

"My goal is nothing less than to build research networks within the Department, across the government, throughout the nation, and around the globe."

Steven Chu, Secretary of Energy, statement before the Committee on Science and Technology, U.S. House of Representatives, Washington, D.C., March 17, 2009

To this end, OSTI coordinates the STI Program *(STIP)*, a complex-wide collaboration across DOE programs, field offices, national laboratories, and contractors to disseminate and preserve the Department's STI. In addition to text, OSTI recognizes the growing need to capture scientific research data and will pursue this mission responsibility as resources allow. OSTI builds on existing partnerships and seeks strategic alliances across a wide range of communities, including DOE, interagency, international, and private sector, to increase access to the world's research and development *(see Collaboration, p. 38)*.

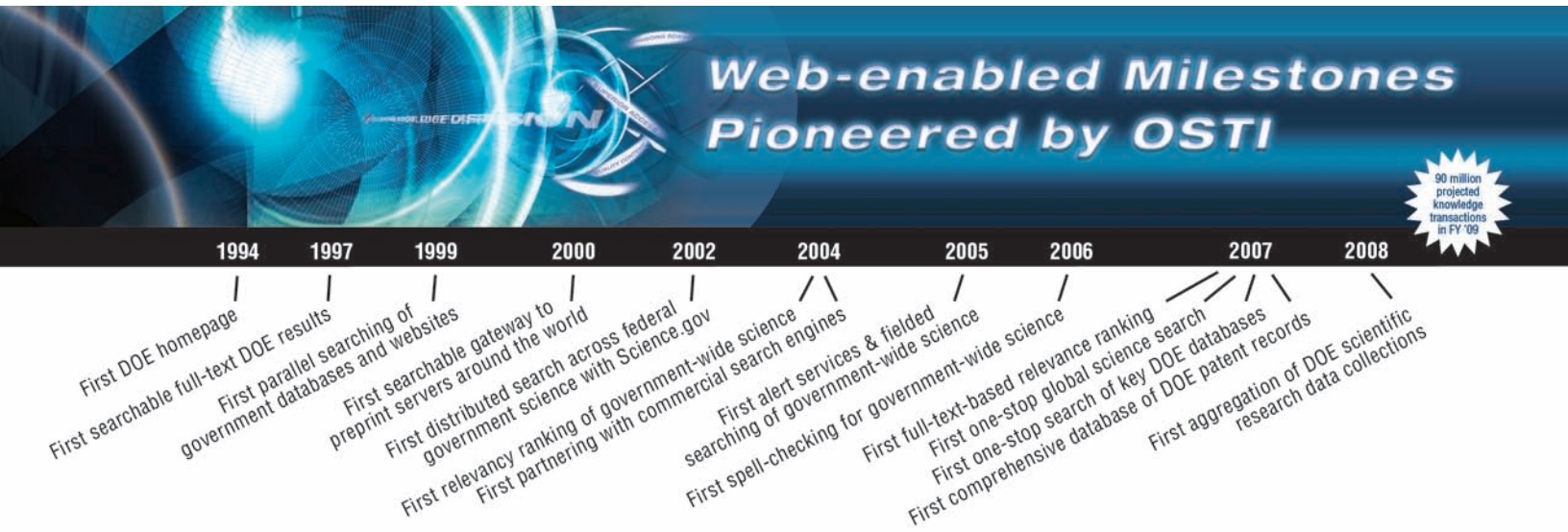
OSTI builds information dissemination networks that speed the sharing of research results within DOE, across federal agencies, and among sources worldwide. This helps nurture the scientific and engineering community and supports scientific discovery, a key to the Nation's economic prosperity, energy security, and global competitiveness.

From developing and publishing the world-renowned Nuclear Science Abstracts in 1948 to launching the online WorldWideScience.org in 2007, OSTI has been a first-adopter of promising technologies. Today OSTI is developing new and better ways to speed access to larger volumes of high quality, more relevant, and authoritative scientific content. Through its commitment to innovation and its focus on producing intuitive and easy-to-use applications, OSTI is laying a solid foundation for the future.

Foundations for the Future

Looking toward the future, OSTI continuously seeks to expose more of the world's scientific and technical information *(STI)* to researchers and the public to enable new scientific discovery. Science discovery is required to meet national and worldwide needs for major advances to power the Nation's economy, develop solutions to our energy challenges, and protect our environment. Sharing, or "diffusion," of innovative ideas and discoveries is central to advancing these national priorities.

OSTI has forged strategic alliances and partnerships *(see Collaboration, p. 38)* that help enable discovery. Over the last decade, numerous milestones have been achieved, including many "firsts" in government web search technology *(see Web-enabled Milestones below)*. OSTI led the way among federal science agencies in going beyond bibliographic information by bringing full text to scientists' desktops. OSTI was the first agency to provide full-text searching capability of R&D documents, i.e., DOE technical reports, through the DOE Information Bridge. OSTI also pioneered web-based knowledge diffusion products that foster transparency of e-prints, conference proceedings, patents, and R&D accomplishments. These products span both R&D results from DOE-funded work and R&D of others which is relevant to DOE and the U.S.



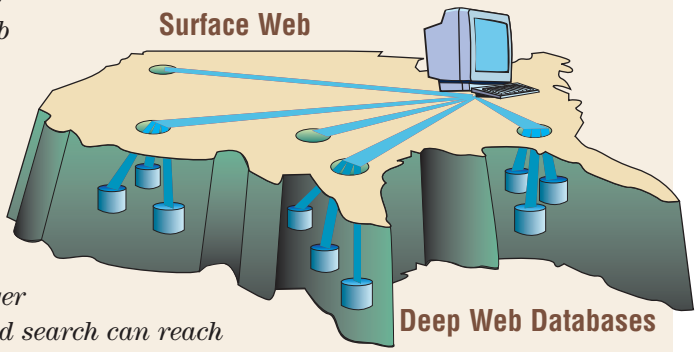
Having developed an impressive suite of key DOE collections, OSTI staff soon realized that the power of an innovative technology, federated search *(see Appendix 4, Federated Search and Directed Harvesting, p. 61)*, could be applied to integrate these holdings and make these collections searchable from a single point. This realization, combined with the understanding that knowledge sharing can be measured, modeled, and accelerated through increasing "transfer rates" among researchers *(see Appendix 5, Diffusion Research, p. 62)*, led to the creation of the DOE Science Accelerator. By simplifying the search experience and drastically reducing search time, OSTI made the web work better for science.

Working in partnership with an interagency alliance *(see Collaboration, p. 38)*, OSTI introduced federated search across the federal government agencies through the launch of Science.gov in December 2002 *(see Appendix 6, Science.gov: A Case Study in the "Art of the Possible," p. 64)*. Science.gov expanded the scope of searchable science content of interest to DOE beyond that produced by DOE. Science.gov makes available the R&D

Non-Googleable, Authoritative Content Is Critical to Science

OSTI has invested in federated search technology because most science information is out of reach of conventional search engines such as Google and Yahoo. Instead, this important but "non-Googleable" science information resides within databases in what is called the deep web. While conventional search engines are great for finding web pages ("surface web" content), such as www.osti.gov, *they typically cannot reach authoritative information inside a deep web database. Instead, this non-Googleable database content is retrieved through the search engine specific to that database.*

OSTI has pioneered a new class of search engine, the federated search engine. Federated search is specifically designed to access non-Googleable, distributed databases in the deep web. Federated search enables a single Google-like search box to launch searches across a discrete set of deep web databases. According to the OCLC (Online Computer Library Center) report, "Perceptions of Library and Information Resources," 84% of the general public begin their search using search engines. Only one percent of the general public begins with online databases.



By using federated search technology, it no longer matters where the information resides; federated search can reach it. Now, a new challenge has emerged—ramping up to search across larger numbers of deep web databases. The associated technological barriers need to be overcome.

results from 14 federal science agencies, representing 97 percent of the U.S. Federal R&D budget. More than 200 million pages of science are searchable through this single federated search application. Science.gov has become a model for knowledge sharing that has inspired other OSTI applications.

One such application is WorldWideScience.org. Through international collaboration, OSTI enlarged the scope of searchable documents from those produced solely by U.S. science agencies to those produced worldwide. WorldWideScience.org brings together approximately 375 million pages of scientific information from important databases and portals from around the world *(see Volume of Content Made Searchable by OSTI, p.14)*.

Through these state-of-the-art products and services, OSTI brought Google-like capabilities to the deep web, the part of the web where most science information resides, but where conventional search engines generally cannot go. This is an important step in making the web work better for science, as analysts estimate that perhaps 99 percent of all the web-accessible scientific documents are in deep web databases* and essentially "non-Googleable" *(content that cannot be searched by Google and other conventional search engines, see Non-Googleable, Authoritative Content Is Critical to Science, facing page)*.

OSTI has brought together islands of information dispersed within DOE, across the Nation, and around the world and made the information findable .The OSTI virtual colle ctions and the tools that make the mavailable are the gold standard of STI se arch, and they are possible only by the deployment of groundbreaking technology.

* *The Deep Web: Surfacing Hidden Value, Michael K. Bergman, BrightPlanet, http://brightplanet.com/pdf/deepwebwhitepaper.pdf (September 2001)*

Volume of Content Made Searchable by OSTI

WorldWideScience.org:
375,000,000 pages of Global Scientific and Technical Information (STI)

These web-available pages would fill 62,000 traditional 2-feet deep file drawers.

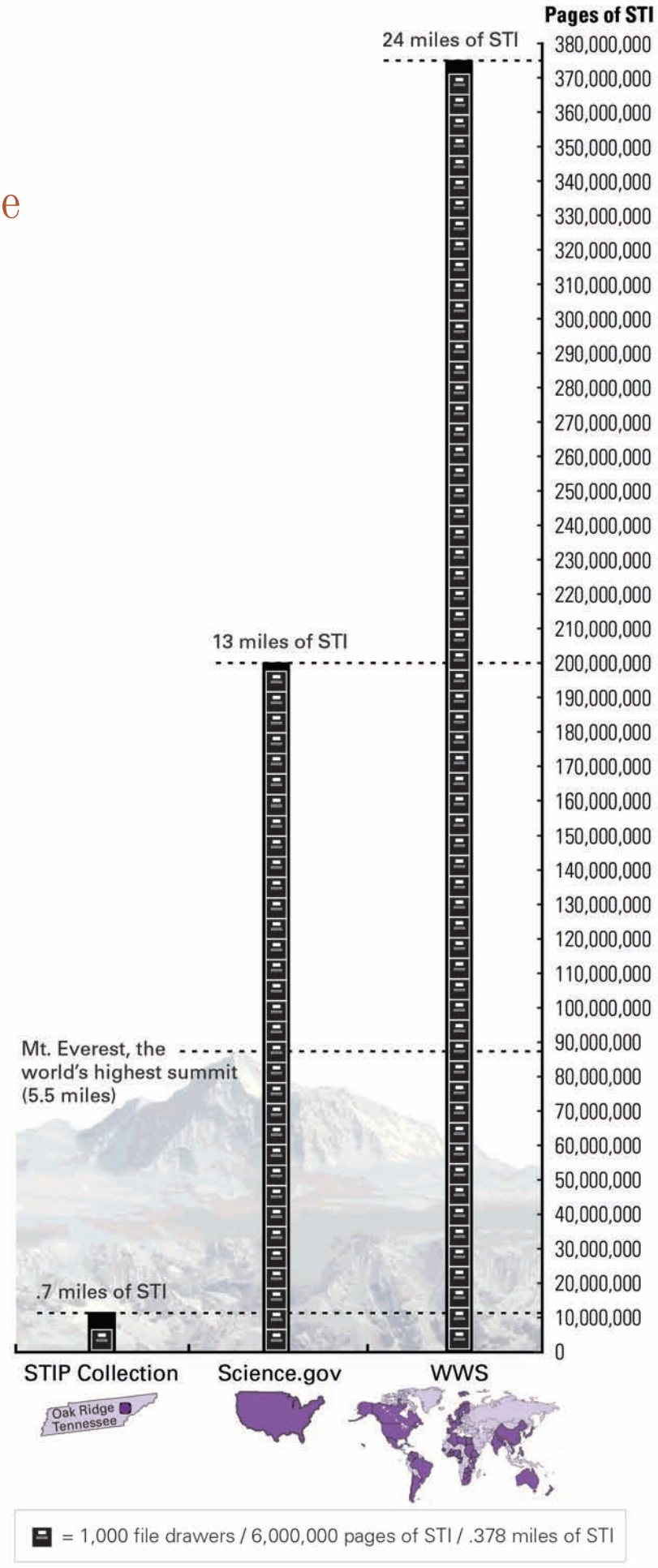
Science.gov:
200,000,000 pages of U.S. Government STI

These web-available pages would fill 33,000 traditional 2-feet deep file drawers.

STIP Collection:
11,400,000 pages of U.S. Department of Energy STI

These web-available pages would fill 1,900 traditional 2-feet deep file drawers.

Amount of Data Transferred in FY08: 9.95 terabytes



The WorldWideScience.org page count is comparable and complementary to the science content of Google, but does not duplicate it. WorldWideScience.org finds content from the deep web, while Google searches primarily the surface web *(see Non-Googleable, Authoritative Content Is Critical to Science, p. 12)*. OSTI staff is aware that although WorldWideScience.org and Google are huge, they still overlook much of the world's research and development findings. What do we mean by huge? At OSTI alone, researchers and the public can gain free electronic access to what would be *(if the documents were in paper format)* the equivalent of approximately 1,900 file drawers, each 2-feet-deep and full of 60-page documents. That's almost three-fourths of a mile of file drawers filled with important R&D findings.

File Drawer Calculation

Our "file drawer" is a metaphor that emphasizes the volume of content managed by different entities: to be clear, our applications search computer-based files.

We estimate an average scientific document to be 60 pages. There are approximately 250 pages per inch. 1 file drawer is 24" deep.

One file drawer holds 24 inches x 250 pages/inch = 6,000 pages

DOE has just such a collection through a dedicated program called the Scientific and Technical Information Program *(STIP)*, a collaboration that ensures DOE research findings are preserved in a central location and appropriately disseminated. For example, OSTI disseminates holdings resulting from DOE work within such collections as Information Bridge, DOE R&D Accomplishments, DOepatents, and DOE R&D Project Summaries. As large as the STIP collection may seem, the documents managed within STIP comprise only a tiny proportion of science documents available for OSTI to disseminate to researchers and the public.

OSTI also disseminates STI that is of interest to DOE but, for the most part, is not produced by DOE. These citations and full-text documents reside on the Internet in servers all over the world. OSTI has identified the locations of this important STI and provides a means to search these far-flung collections by deploying tools such as Science.gov, the E-print Network, Science Conference Proceedings, and EnergyFiles. As one example of the volume of information this entails, the E-print Network collection alone is over 5 million technical e-print documents. Unlike the STIP-managed technical reports that all originate in one agency *(DOE)*, the e-print documents must be gathered from a large number of worldwide sources.

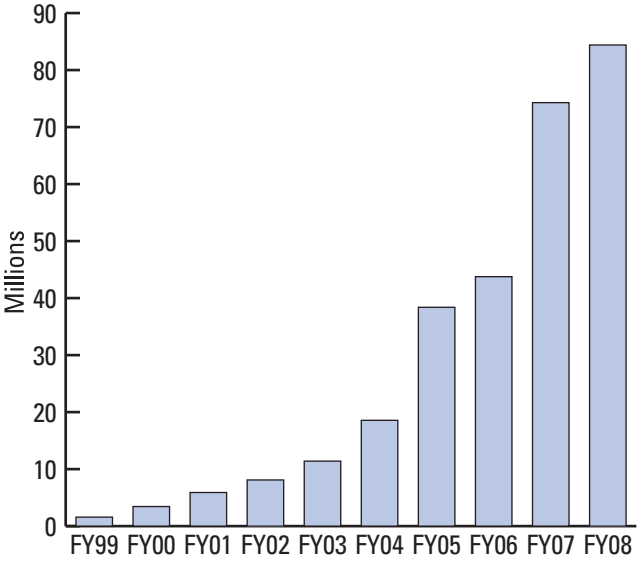
OSTI is not the only government office tasked with acquiring and sharing vital science information. Other U.S. federal agencies, other scientific organizations, and other countries share this mission. In collaboration with 13 other agencies, OSTI hosts Science.gov, the gateway to 200 million pages of U.S. science information *(33,000 file drawers with 13 miles of information)*. And WorldWideScience.org, the gateway to more than 375 million pages of global science information, would fill 62,000 file drawers with 24 miles of information.

Volume of content made searchable by OSTI is huge, and so is Google's; yet there are still gaps to be filled.

There are many, many more "file drawers" of science information that are not readily searchable by conventional search engines such as Google or by OSTI. The challenge is to find and deploy ways to search all "file drawers" of science information efficiently, thus making the information useable for the researcher and the U.S. public so that science can advance. OSTI is committed to this challenge.

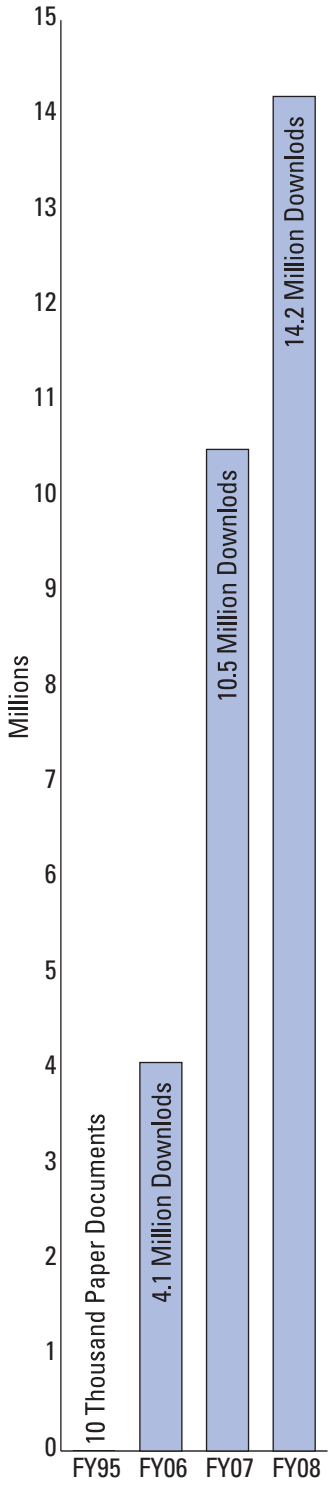
Measures of Progress

OSTI tools and services are playing an undeniable role in accelerating scientific progress.



Increases in Transactions *(Chart 1)*

In the aggregate, OSTI web tools receive 80 million transactions annually. A transaction is equivalent to a single page display in a given session.



Access to Full-Text Reports *(Chart 2)*

OSTI has dramatically increased access to information from full-text reports, from the pre-web environment when a few thousand documents were delivered in paper upon request to today as millions of full-text reports are downloaded from OSTI each year free of charge.

Vision

Thought Transfer: Researchers agree that science advances only when knowledge is shared.

OSTI accelerates discovery by finding new ways to share scientific and technical knowledge, seeking to make results increasingly relevant to search queries, and by making science information more usable and accessible to more people.

To enable global scientific discovery by making the web work better for science

OSTI envisions U.S. scientists and engineers, educators, entrepreneurs, and the public having superior access to find the quality scientific content they need. This vision, shaped by stakeholders during eight workshops, will ensure that scientific and technical content from sources around the world and in various formats will be made available to advance discovery on a global scale. By providing information rapidly in forms that can be readily found and used, OSTI will accelerate communication of DOE R&D results as well as worldwide discoveries, speeding up the sharing of new knowledge. OSTI calls this result global scientific discovery.

OSTI is working to achieve global scientific discovery through innovation such as federated search technology *(see Appendix 4, Federated Search and Directed Harvesting, p. 61)*, which allows simultaneous search of often hard-to-find scientific collections. The current challenge is the limitation in the number of collections that can be searched simultaneously.

In pursuit of global discovery, OSTI will ensure efficient access to authoritative scientific results across distant R&D communities, thus making the web work better for science. Such access will better equip DOE and the Office of Science to strengthen America's role as the world leader in science and technology and address the Department's priorities for:

- Science and Discovery
- Clean Energy
- Economic Prosperity
- National Security and Legacy
- Climate Change

Global discovery means making each original scientific discovery globally available and findable. It means transforming local discovery into globally known discovery. It means making the web work for science. That is the OSTI vision. Many challenges lie ahead.

The Challenge

To make the web work for science by bringing together authoritative science from all corners of the earth and making it easily findable.

Although the web was invented for scientists to share research findings, ironically these beginnings have been largely forgotten. Today the web does not work well enough for science because research findings are often difficult to locate and access on the web. OSTI has made great strides in developing new ways to expose science information from DOE (ScienceAccelerator.gov), from U.S. federal agencies (Science.gov), and from national sources worldwide (WorldWideScience.org) *(see p. 15)*. However, much more needs to be done.

There are gaps that need to be filled to make the web work better for science *(see Gaps To Be Filled, facing page)*. Much of the valuable science content on the web has not been integrated into the well-known science search applications, nor is it well organized for searching. Non-English documents need to be made searchable. Access is needed to websites that require authentication. The number of important science collections that can be searched simultaneously is limited. Non-textual media such as images, audio, video, and numeric data often remain difficult to find. Valuable science collections and content need to be identified and made efficiently findable.

To accelerate the advancement of science, these gaps need to be filled. Each objective in this Strategic Plan is designed to fill a gap to make the web work better for science. It is a significant undertaking, but the challenges are worthwhile.

Gaps were identified in three areas: content, i.e., valuable scientific information not easily findable on the web; information technology, including tools for searching and finding the valuable information; and communications/outreach, to ensure that patrons are aware of available resources.

Gaps To Be Filled

These gaps were discussed by stakeholders as steps OSTI should take to make the web work better for science.

SCIENTIFIC AND TECHNICAL CONTENT

- Stranded collections of DOE R&D results are not preserved and made accessible
- Non-textual forms of STI *(audio, visual, image, numeric data, and illustrations)* are not searchable and accessible
- Hundreds of thousands of DOE technical reports have not been digitized
- Science education resources have limited searchability and are often hard to find
- Many global information sources remain inaccessible *(information exchange agreements and search tools are absent)*
- Many developing countries are not engaged in information exchanges
- Comprehensiveness and currency of weapons R&D information collections are lacking

INFORMATION TECHNOLOGY

- Great quantities of STI are non-Googleable
- Precision search technology is outpaced by the volume of STI
- A new class of web tools is needed to help researchers collaborate and access, analyze, and manipulate data
- Research of the spread of knowledge is limited
- Federated searching technology and interoperability is limited
- OSTI's internet connectivity, capacity, and reliability are inadequate
- Technical global communications challenges exist *(e.g., language barriers and format standards)*
- The special challenge of sharing classified information via the web has not been met
- The R&D project life cycle still relies on outdated systems and procedures

COMMUNICATIONS/OUTREACH

- University research programs are not equipped with knowledge of DOE STI resources
- Public awareness of DOE and Office of Science researchers and their scientific contributions is lacking
- OSTI is not well known among DOE programs

Strategic Goals

1. *Ensure superior access to and preservation of quality scientific and technical content*
2. *Improve education by providing enhanced access to scientific and technical knowledge*
3. *Bring the world's science to U.S. science and research communities*
4. *Increase visibility and understanding of DOE contributions to science*
5. *Contribute to national security*
6. *Provide specialized support to DOE R&D Programs*

Researchers, engineers, and the public need authoritative science information fast. OSTI will build on a record of results and work to fill gaps in order to make the web work better for science.

OSTI Envisions Three Kinds of Growth

Content: Incorporate new STI into search applications to make it findable

Functionality: Build more sophistication, power, and functionality into products as technology evolves

Products: Create new types of products using web-based innovation to serve user needs

The pursuit of OSTI's mission is powered by rapid advances in information and communication technologies. OSTI continually seeks "the art of the possible" and is an early adopter of new information technologies to pioneer new products and services.

Because science and discovery are at the core of achieving national priorities, OSTI will combine its expertise in providing access to volumes of scientific and technical content with its drive for innovation. By achieving these goals, OSTI will advance the Nation's science and technology base.

The achievement of our mission over the next six years will be focused on the strategic goals shaped by our stakeholders and described in the following sections.

OSTI will ensure efficient access to authoritative scientific results across distant communities, thus making the web work better for science. Such access will better equip DOE and the Office of Science to strengthen America's role as the world leader in science and technology and address the Department's priorities for:

- Science and Discovery
- Clean Energy
- Economic Prosperity
- National Security and Legacy
- Climate Change

Goal 1

Ensure Superior Access to and Preservation of Quality Scientific and Technical Content

The overarching goal—identified and shaped by OSTI's stakeholders—is to ensure that scientific and technical content is easily findable and preserved for reuse. Achievement of Goal 1 includes objectives dealing with scientific content, innovative and intuitive search and retrieval tools, and the information infrastructure required to close these gaps. The availability of scientific and technical content is the primary focus, as databases and bodies of knowledge in various existing and emerging fields of science and technology continue to expand. Innovative systems will need to be developed and enhanced to provide the correct level of superior access to quality content.* While technology offers increased availability of scientific and technical information on the surface web, gaps exist *(see Gaps To Be Filled, p. 19)* because volumes of important scientific information are non-Googleable and therefore are not easily found.

OBJECTIVE 1.1

Increase Volume of Searchable DOE R&D Results

Gap: Not all DOE R&D findings are captured and preserved through the current systematic processes used by the DOE-wide STI Program. Some R&D results are dispersed at laboratory and university websites and in forms other than currently captured by STIP. Comprehensive coverage and preservation of DOE STI is needed.

Opportunity: OSTI will seek out these dispersed, stranded collections in new ways to identify and harness the scientific content for OSTI products and services.

Actions:

- Work with STIP community to identify new or obscure STI resources.
- Conduct pilot of new data mining methodology for capturing STI for OSTI's central collection.
- Explore use of site map protocols at DOE labs/facilities to identify STI suitable for access.
- Increase percentage of current volume of searchable DOE R&D results by 20 percent per year to make DOE R&D results accessible and findable for contributing to the advancement of science.

OBJECTIVE 1.2

Increase Digitization of Legacy Holdings

Gap: Only a small portion of past DOE R&D output is digitized. Approximately 800,000 DOE technical reports generated between 1946 and 1990 are available only in paper format, leaving two-thirds of the DOE legacy output unsearchable electronically.

* *Identified by workshop panel participants in the Workshop Panel Report on Accelerating the Spread of Knowledge About Science and Technology: An Examination of the Needs and Opportunities, http://www.osti.gov/publications/2007/workshop.pdf (April 2007).*

"Information maintained by the Federal Government is a national asset. My administration will take appropriate action, consistent with law and policy, to disclose information rapidly in forms that the public can readily find and use."

President Barack Obama, Memorandum for the Heads of Executive Departments and Agencies, January 21, 2009

Opportunity: To enable public access to the large majority of DOE and predecessor agency R&D outputs currently unavailable in electronic format, collaborative digitization efforts and other partnerships will be sought.

Actions:

- Negotiate partnership*(s)* with private and/or public sector organizations for accelerating document digitization activities.
- Digitize 100 percent of non-electronic full-text R&D reports dating from 1940s to early 1990s.
- As documents are digitized, add full-text documents to DOE Information Bridge as available until the entire legacy collection is searchable and made available through that database to federated portals.

OBJECTIVE 1.3

Ensure Access to Deep Web Science Content via Federated Search

Gap: Most scientific content cannot be found by conventional search engines. The majority of scientific research results resides in databases that the conventional web crawlers cannot find. Many of these databases, which include government databases, are free and publicly accessible.

Opportunity: To make the web work better for science, OSTI will continue to ensure access through single-search interfaces to science information that cannot be found via popular search engines, e.g., Google.

Actions:

- Add databases/resources into each federated portal, such as adding the DOE Data Explorer database to the DOE Science Accelerator.
- Add DOE as well as other U.S. science agency databases and portals into the federated search of Science.gov, such as the seven new sources added in Science.gov Version 5.0.

Today, 50 million pages from DOE technical reports generated between 1946 and 1990 are not digitized and are thus very difficult to retrieve. Lack of budget has kept OSTI from filling this gap more quickly. It would cost $5 million to completely fill this gap.

- Identify database sources that are difficult to access and make them findable.
- Expand federated searching tools to incorporate national and international information and data sources, enhancing distributed and federated searching technologies with the capability to quickly search a greater number of databases simultaneously and deliver real-time results.

OBJECTIVE 1.4

Make Non-Text Information Findable

Gap: Non-text scientific research data (such as images, illustrations, audio, video, and numeric data) are largely not searchable, making data difficult to locate. While DOE research results are well documented in journal and conference literature, technical reports, and full-text databases, there are

also many available collections of supporting data in non-text formats. Until recently, researchers, scientists, and the public could not easily locate non-text data generated from DOE-sponsored research.

Opportunity: OSTI is narrowing the science content gap initially by creating a catalog of the sources of non-text collections in a database, the DOE Data Explorer. The next step is to make this scientific research data even more findable, ensuring that scientific data is never distorted or concealed.

Actions:

- Develop searchability of the websites where DOE digital data sources reside, including DOE Data Centers and other recognized sets of numeric data at DOE labs by directed crawling, harvesting and indexing these sites.
- Identify databases and websites to continually expose more non-text data to those seeking to find useful information fast.
- Deliver information from data repositories, video, images, and other non-text sources by providing federated searching capabilities such as those available via the DOE Science Accelerator.
- Establish method for linking datasets to published papers and reports.
- Explore creation of a new product to make multimedia information, such as videos of scientific presentations, searchable.
- Make illustrations, tables, and charts within text searchable.

OBJECTIVE 1.5

Create Next-Generation OSTI Infrastructure

Gap: OSTI infrastructure and internet capacity are inadequate and need to be improved to meet the information needs of tomorrow.

Opportunity: OSTI will work to improve the sophistication and power of federated searching technology and reduce limitations to interoperability. Redundancy in connectivity will ensure that online capabilities remain in case of a single failure of a router, cable, or circuit.

Actions:

- Improve network and computing speeds and incorporate architecture changes to promote development and hosting of next generation search technologies.
- Overcome limitations of interoperability to perform deep web searches of full-text databases and return complete and relevant results.
- Actively monitor system performance and incorporate architecture changes as needed to improve load balance.
- Continue to ensure network capacity and computing resources support exponential increases in the volume of information.
- Provide a redundant off-site processing capability which, in the event of a building disaster or long-term communications failure, would ensure the discovery of knowledge and continued advancement of science supported by the significant STI holdings of OSTI and its search/retrieval services.
- Procure an ESNet-independent T3 line and the network infrastructure necessary to automatically operate to provide required service in the case of an ESNet connection failure.

OBJECTIVE 1.6

Enhance and Deploy Precision Search Tools

Gap: Precision search technology has not kept pace with the volume of scientific information available to search engines. As more and more information becomes available on

the web, the need to refine and hone search results becomes more critical. Without precision, searchers will be inundated with non-relevant information.

Opportunity: OSTI has deployed a number of federated search tools with precision search capabilities. The volume of information calls for search tools that facilitate discovery of federal government science information to be enhanced to deliver accurate and timely results that are increasingly more relevant to a patron's queries. OSTI will continue to provide the most relevant, real-time information through precision searching using an efficient, intuitive interface *(see Appendix 7, Precision Search, p. 66)*.

Actions:

- Work with the Science.gov Alliance to define and develop next-generation algorithms for more sophisticated and faster relevance ranking across large retrieval sets.
- Deploy improved search and clustering of results features, such as in Science.gov Version 5.0 and WorldWideScience.org, allowing for faster relevance ranking and easier navigation and review of search results.
- Continue to implement more precise search algorithms so that a thousand or more sources can be searched in real time with greater flexibility for customization by the user.

OBJECTIVE 1.7

Implement Web Tools with the Researcher in Mind

Gap: Web tools do not adequately serve the needs of researchers to collaborate across distant communities and to access, manipulate and reuse the world's research findings.

Opportunity: OSTI will continue to actively seek ways to better serve the research community by increasing visibility of, access to and reuse of scientific research. OSTI will continue to develop where necessary and broadly deploy state-of-the-art information tools and services to enhance scientist-to-scientist communication.

Actions:

- Expand on a wide range of initiatives to ease access to research results, such as Reference Linking and Author Notification as well as collaborative efforts with private sector entities such as CrossRef, SPARC *(Scholarly Publishing and Academic Resources Coalition)*, and WorldCat *(see Collaboration, p. 38)*.
- Seek partnerships and collaborative arrangements to ensure continued visibility and accessibility of DOE R&D.
- Enhance products and services to support multiple channels of communication capabilities, making use of Web 2.0 services and other developing, next-generation technological capabilities.

OSTI Helps Increase Researcher Networking

"Reference Linking" is a research networking tool that leads to knowledge sharing of all kinds, encouraging the exchange of ideas and ensuring that more research information is just a click away. OSTI supports the practice of reference linking to technical reports and journal articles by hyperlinking the references at the end of a document. This greatly enhances the technical report and the accessibility of the references. This service is provided at no cost to the author or their organization. Working with CrossRef (see Collaboration, p. 38), OSTI is using Digital Object Identifiers (DOIs) to facilitate access to DOE's vast collection of science research reports. Technical reports are assigned a DOI so that journal publishers can easily link to those reports.

This includes deploying information tools and social networking services, such as the review and commentary capability, Document Discussion, for technical reports through the DOE Information Bridge.

- Enhance the DOE E-print Network by including a capability for researchers to create and share personal library collections. Introduce the use of blogs for various OSTI database products and services to enhance researcher communication efforts.

OBJECTIVE 1.8

Update STI Best Practices

Gap: STI Program practices in place across the DOE complex have not kept pace with technology or new types of content and formats of STI.

Opportunity: OSTI will work to ensure that STI practices and the collaborative infrastructure in place across the DOE complex remain best in class and continue to make the DOE STI Program a success, facilitating the innovative and creative delivery of R&D results.

Actions:

- Develop and implement DOE-wide STI practices to make additional formats of STI accessible, such as multimedia.
- Support pilot projects for innovation in STI processes across the laboratory complex, as funding allows.
- Implement additional measures to ensure originating sites conduct appropriate reviews prior to public release of STI, e.g., to identify Protected PII (Personally Identifiable Information) for redaction prior to posting on publicly accessible web systems.
- Identify and implement improvements in practices and infrastructure that create no burden to the laboratories and other information originators but that ensure the DOE-wide STI program maximizes access to DOE R&D findings, regardless of the form of the information.
- Work with STI organizations in other U.S. federal agencies, such as CENDI, to explore and share practices in emerging fields.

OBJECTIVE 1.9

Perform Diffusion Modeling Research

Gap: There is a lack of applied research on the spread of scientific knowledge.

Opportunity: By conducting applied research of the knowledge diffusion process, OSTI is exploring ways to accelerate the "transfer rate" among researchers to speed the sharing of new scientific knowledge *(see Appendix 5, Diffusion Research, p. 62)*. OSTI will seek out mechanisms to quickly spread important ideas among researchers and foster collaboration using concepts related to the well-developed mathematics of the spread of infectious disease. By measuring and modeling the spread of ideas, steps can be taken to accelerate the "rate of infection" for future scientific productivity.

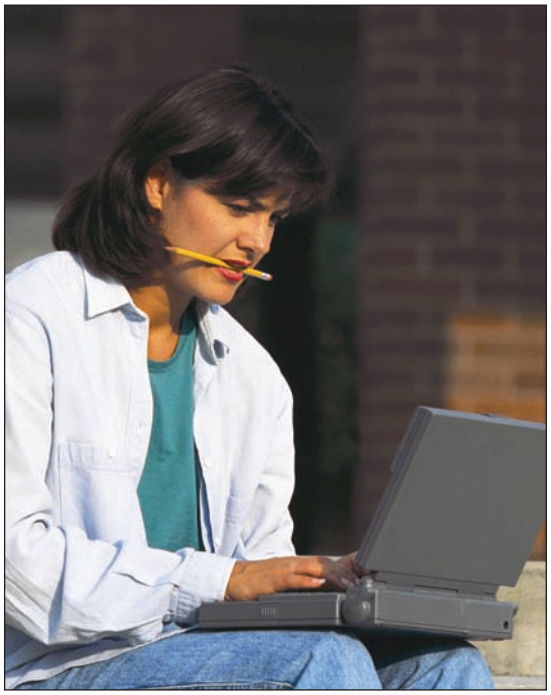
Actions:

- Study the modeling of diffusion to gain additional insight on methods to speed the rate of knowledge transfer.
- Continue to explore ways that knowledge diffusion can be accelerated through services/tools developed or applied by OSTI.

Goal 2

Improve Education by Providing Enhanced Access to Scientific and Technical Knowledge

As a means of sustaining American competitiveness in science and technology, OSTI will enhance access to high-quality education resources available at DOE national laboratories and other DOE-funded institutions. Additionally, valuable education resources reside beyond DOE in various agencies which could be included in a federated portal of Science, Technology, Engineering, and Mathematics *(STEM)* resources in support of the *America COMPETES Act.*



OBJECTIVE 2.1

Provide Web-Searchable Access to DOE's Education Resources

Gap: Teachers and students—i.e., the Nation's emerging scientists—have limited web-searchable access to world-class science education materials. While commercial search engines are heavily used by students and teachers, a wealth of authoritative STEM education resources resides on websites that are unfamiliar to the general public.

Opportunity: OSTI is working to help teachers and students find STEM educational resources from DOE programs and national laboratories through a web-based portal.

Actions:

- Develop web-searchable access to STEM resources by enabling automated directed crawling and indexing of the DOE programs' and labs' education materials that are accessible via the web.
- Create a "learning level" selection based on education terms to yield more precise and useful materials for students or teachers. Content will be searched and results will be displayed by degree of difficulty.
- Identify DOE education materials not yet online and convert to web accessible formats.

America COMPETES Act

SEC. 3191. NATIONAL ENERGY EDUCATION DEVELOPMENT.

(a)IN GENERAL—The Secretary, acting through the Director and in consultation with the Director of the National Science Foundation, shall establish a program to coordinate and make available to teachers and students web-based kindergarten through high school science, technology, engineering, and mathematics education resources relating to the science and energy mission of the Department, including existing instruction materials and protocols for classroom laboratory experiments.

(b)ENERGY EDUCATION—The materials and other resources required under subsection (a) shall include instruction relating to—

(1) the science of energy;

(2) the sources of energy;

(3) the uses of energy in society; and

(4) the environmental consequences and benefits of all energy sources and uses.

(c)DISSEMINATION—The Secretary, acting through the Director, shall take all steps necessary, such as through participation in education association conferences, to advertise the program authorized under this section to K–12 teachers and science education coordinators across the United States.

OBJECTIVE 2.2

Develop a Portal to World-Class STEM Resources Beyond DOE

Gap: Science education resources have limited searchability and are scattered across the websites of various U.S. federal agencies in isolated pockets.

Opportunity: By integrating the STEM resources of all U.S. federal agencies and improving their usefulness, OSTI will ensure that students and teachers are able to more easily find the materials most relevant to their needs.

Actions:

- Collaborate with other federal agencies to develop distributed searching of education materials available across U.S. science agencies so students and teachers can find relevant materials via a single search.
- Enhance functionality so users can find materials of specific interest, such as lesson plans for teachers.
- Provide collaborative online workspace for educators and students to foster learning among the Nation's emerging scientists.

Goal 3

Bring the World's Science to U.S. Science and Research Communities

To strengthen America's role as a world leader in science and technology, U.S. scientists and engineers must keep abreast of scientific discoveries worldwide. OSTI will multiply the volume of international science information available to U.S. science and research communities, thereby contributing to the acceleration of scientific progress.

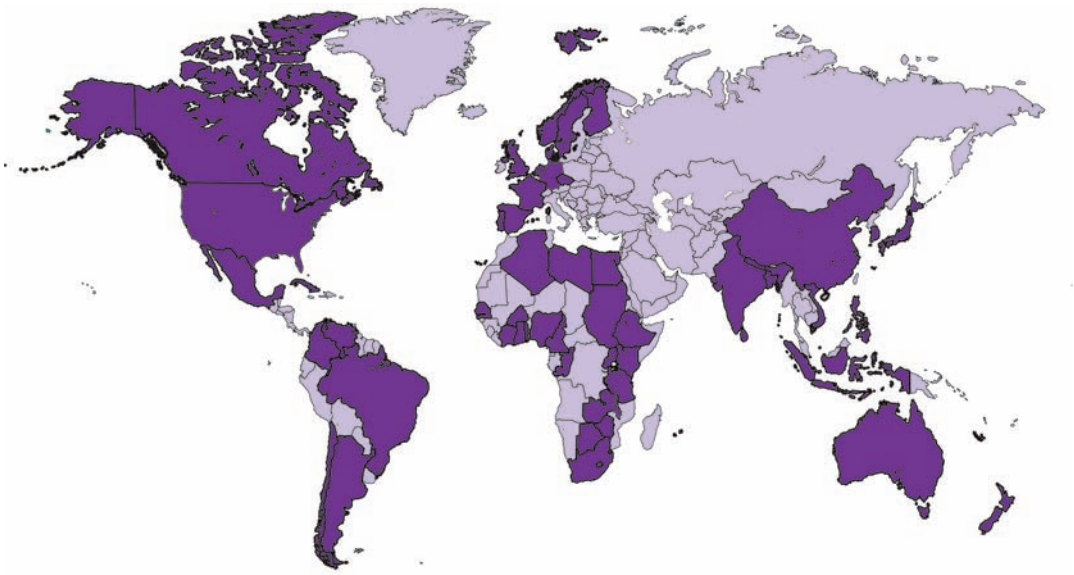
OBJECTIVE 3.1
Leverage International R&D Results

Gap: Many global research sources remain inaccessible through traditional search engines.

Opportunity: OSTI will enhance America's centralized and competitive position in science by leveraging international information exchange agreements and global search tools to provide researchers with access to far-flung science information.

Actions:

- Increase the number of countries and information collections represented in WorldWideScience.org and the Energy Technology Data Exchange *(ETDE)*, ultimately enlarging the global body of information available to the U.S. science community.



- Build technical and interactive enhancements into international information tools *(e.g., simplified authentication, translations, web functionality)* to facilitate access to higher quantities of information and to provide a forum for cyber communications among distant research communities.
- Build on OSTI's standing as a player in international scientific information access by expanding and enhancing information exchange and global search capabilities.

OBJECTIVE 3.2

Develop and Enhance Global Partnerships to Address Technical Challenges

Gap: Language differences and information quality and format standards vary widely across countries.

Opportunity: OSTI will partner with its counterpart agencies in other countries which have specialized expertise in addressing certain challenges such as multilingual translations, thereby addressing communications and other technical challenges.

Actions:

- Use OSTI's well-established relationships with world-class STI organizations to bring leading-edge capabilities and functionality to OSTI's information products.
- Engage in joint systems development and other projects with leading libraries in Europe and Asia to promote maximum access to full-text literature and numeric data.

OBJECTIVE 3.3

Promote G8 Energy-Related Developing Country Goals

Gap: The role of the U.S. and other developed countries in sharing science knowledge with the developing world has not been fully realized.

Opportunity: Providing access to science and technology in the developing world, combined with leveraging their own expertise, will be critical to developing country economic and environmental progress as well as to their contributions to meeting global scientific challenges.

Actions:

- Support G8 energy ministerial goals of promoting energy technologies in the developing world by maximizing access and use of global R&D information.
- Demonstrate increased use of OSTI global discovery tools among developing countries.
- Engage developing countries in dissemination of and representation in global scientific communication strategies.

Goal 4

Increase Visibility and Understanding of DOE Contributions to Science

As the organization within the Office of Science responsible for communications and public affairs, OSTI will work to ensure that the DOE contributions to science are shared with a wide range of audiences, including the general public, news media, scientific and research communities, business/industry, and government, including Congress, Office of Management and Budget *(OMB)*, and Office of Science and Technology Policy *(OSTP)*. Emphasis will be placed on the achievements funded by the Office of Science.

OBJECTIVE 4.1

Increase Understanding by Public and Stakeholders

Gap: The U.S. public does not fully recognize contributions to the advancement of science made by DOE and the Office of Science.

Opportunity: Through multiple outreach initiatives, OSTI will increase understanding of the public and key stakeholders of DOE and Office of Science researchers and mission accomplishments.

Actions:

- Efficiently communicate Office of Science contributions to science by conducting communications activities for all Office of Science Headquarters, including the Office of Science Home Page, press releases, speeches, and announcements as well as coordinating with DOE laboratories in support of Office of Science mission areas.
- Broadly promote and publicize the DOE National Science Bowl for high school and middle school competitions, highlighting the best in U.S. science education.
- Raise public awareness of the contributions made by Office of Science, its programs, national laboratories and scientific user facilities and effectively share the results of the Nation's R&D investment through exhibits, press articles, and other public information outlets.

OBJECTIVE 4.2

Raise Visibility of Research Achievements

Gap: Communication tools do not currently exist to highlight specific accomplishments of researchers funded by the Office of Science or to recognize major achievements of Office of Science projects.

Opportunity: OSTI will utilize advanced communication tools to spread awareness about the Office of Science record of accomplishments and progress on key

Transformative Solutions

"Some of these technologies are already within our grasp, while others are yet to come. As the premiere science agency in the field, the Department of Energy will play a crucial role in developing those solutions and in helping them take hold in our Nation's economy."

Energy Secretary Steven Chu, Department of Energy, February 5, 2009

scientific challenges and its leadership role for large-scale scientific infrastructure. Various tools and services, such as social media and partnerships, will be explored to build and maintain support for sustained investments in Office of Science research.

Actions:

- Produce a monthly electronic newsletter and an annual report to highlight the progress and outcomes from Office of Science research projects, for the scientific community, industry, academia, and policymakers.
- Recognize national scientific achievement and merit through a major science awards program and use high-visibility forums to promote science through exhibits, distinguished lectures, and science news services.
- Expand ongoing initiatives to help the Office of Science program offices improve their communications of science contributions.

OBJECTIVE 4.3

Raise Visibility Across University Research Programs

Gap: Many of the Nation's university research programs and their supporting libraries are not aware of various DOE information resources that are available for their use.

Opportunity: OSTI will continue to work more directly with university libraries and research departments to increase their awareness of valuable DOE scientific and technical information resources. In the process, OSTI will gain insight into ways to improve access by this community of information consumers and enhance reuse of the R&D information.

Actions:

- Make the DOE STI collections more useful to universities by making them suitable to be acquired/utilized by university libraries through their routine acquisition programs and services, such as was done in 2008 by making STI available in MARC format.
- Educate university libraries and research departments about the availability and value of the DOE scientific and technical information products and facilitate access to these resources by this community from their own websites and information resources.
- Reach out to all major university research programs that do business with DOE, providing educational programs, outreach materials, and site visits to assure that they have resources needed to utilize the results of DOE's R&D investments.

Goal 5

Contribute to National Security

In addition to OSTI's responsibility for collecting and disseminating DOE's unclassified research outputs, OSTI has responsibility for the collection, preservation, and protection of classified research information as well as providing secure access to those with a programmatic need for it. OSTI will provide secure access to DOE's historic classified R&D information collection, thereby contributing to DOE priorities and National Nuclear Security Administration *(NNSA)* national security responsibilities.

OBJECTIVE 5.1

Provide Secure Access to Weapons and Other Classified Research Information

Gap: The comprehensiveness and currency of a centralized, accessible classified weapons R&D information collection have declined following classified electronic media concerns.

Opportunity: OSTI will place renewed emphasis on the importance of a comprehensive weapons R&D information collection as a critical component of the DOE/NNSA national security responsibility.

Actions:

- Build on OSTI's leadership and integration of classified R&D information management environment for DOE/NNSA.
- Conduct annual site visits to one or more weapons laboratories to increase awareness of OSTI's and sites' responsibilities in collecting and protecting weapons research information.
- Increase interaction with DOE/NNSA program offices and facilities to ensure total site compliance with DOE requirements for STI submissions to ensure a corporate approach to information management and accessibility.
- Continue to work with subject-matter experts well-versed in classified and sensitive information categories.
- Ensure OSTI's automated systems recognize markings and protect or disseminate information accordingly.

OBJECTIVE 5.2

Enhance Information Protection and Security Technologies

Gap: Web-based sharing of full-text information has not been fully realized in the classified arena because of the requisite security and need-to-know requirements.

Opportunity: OSTI will integrate these capabilities while effectively safeguarding classified and sensitive information.

Actions:

- Engage in DOE/NNSA integrated cyber security initiatives to enable secured sharing of classified research information.
- Establish OSTI's operational status on the NNSA Enterprise Secure Network as an electronic means for classified information access.
- Integrate encrypted and need-to-know technologies into electronic classified information systems.

Goal 6

Provide Specialized Support to DOE R&D Programs

Fulfilling Departmental requirements related to research and development *(R&D)* reporting is a significant area where OSTI systems and information management expertise are called into service. By ensuring that information resulting from research, development, demonstration, and commercial applications activities supported by the Department is accessible and findable, OSTI is not only fulfilling a legislative mandate *(see Appendix 3, Statutory Authorities, p. 60)* but also supporting each DOE program in meeting its specific mission needs. In managing the corporate framework for ensuring that STI emanating from DOE-funded work is made broadly yet appropriately available, OSTI must develop and maintain close ties with both research programs that fund work resulting in STI outputs and staff offices that impact STI policy.

OBJECTIVE 6.1

Enhance Use of OSTI's Resources by DOE Programs

Gap: OSTI is not well known among DOE programs.

Opportunity: OSTI's products and services related to the collection, preservation and dissemination of output from DOE programs provide opportunities for OSTI and the programs to collaboratively meet DOE mission needs. OSTI will actively pursue ways to ensure transparency of information and work with DOE programs to become more familiar with their needs and thereby to enhance the application of OSTI's resources and capabilities in meeting those needs.

Actions:

- Convene meetings between program managers and OSTI staff, providing an opportunity to develop relationships and enhance understanding about OSTI resources directly related to individual programs.
- Pursue regular communication channels with DOE programs to help programs understand how OSTI can support their needs.
- In collaboration with individual DOE offices, continue to develop and maintain systems and customized information tools that provide R&D results and other forms of scientific communication.

OBJECTIVE 6.2

Expand Transparency of R&D Project Data

Gap: DOE R&D Project Summaries does not accurately reflect the overall scope of R&D being carried out nor the latest status of research.

"Executive departments and agencies should harness new technologies to put information about their operations and decisions online and readily available to the public."

President Barack Obama, Memorandum for the Heads of Executive Departments and Agencies, January 21, 2009

Opportunity: To improve transparency of DOE R&D, OSTI will work with various DOE corporate offices and specific programs to enhance and expand comprehensiveness and timeliness of content in the web-accessible DOE R&D Project Summaries database. OSTI will also expand agency coverage in the federated portal, Federal R&D Project Summaries, to ensure transparency of U.S. R&D.

Actions:

- Contribute to DOE's accountability of R&D projects by managing a corporate system in support of Departmental requirements for reporting R&D projects.
- Coordinate collection of R&D project data from laboratories and other facilities to provide open access and to make the data findable with one query.
- Enhance comprehensiveness of the publicly viewed R&D Project Summaries by obtaining timely and accurate data from HQ Programs, other offices and labs, thus improving DOE's information in Federal R&D Project Summaries, the federated portal that has grown from three to nine agencies' R&D data since its inception.
- Continue to expand comprehensiveness of federated search in the Federal R&D Project Summaries by working with other agencies to add more sources from U.S. federal science agencies.

OBJECTIVE 6.3

Improve Efficiency and Effectiveness of R&D Related Business Processes

Gap: Office of Science and other DOE R&D-related business processes do not take full advantage of web-based applications. Information technology has made great strides in recent years, but some R&D management systems remain dependent on outdated software and procedures.

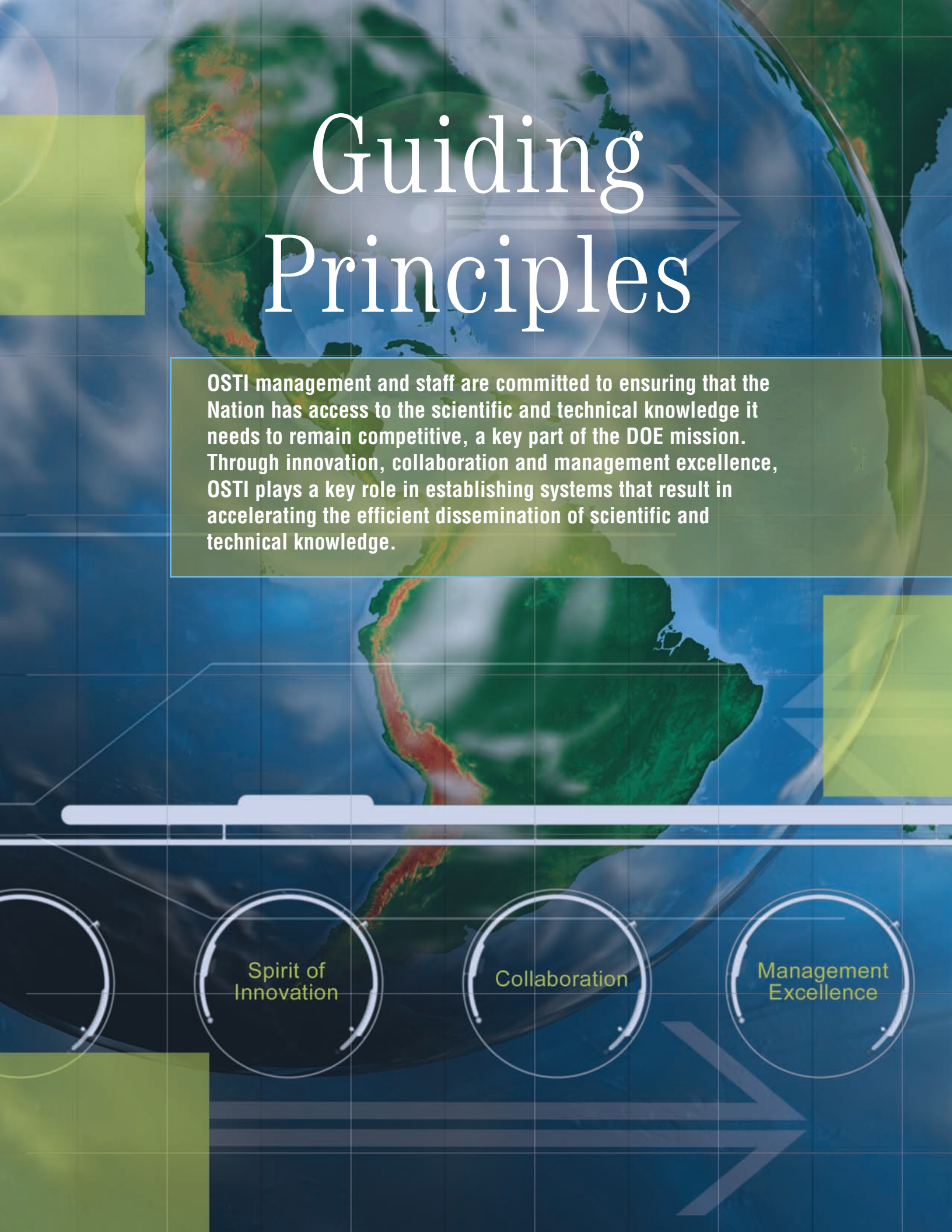
Opportunity: By implementing new web-based functionality and retooling management practices, such as Searchable Field Work Proposal (FWP) submission, access and analysis, many opportunities exist to improve efficiency and effectiveness of the R&D project cycle from proposal to completion. OSTI will explore means for making possible interactive links between project award and R&D results.

Actions:

- Provide Office of Science program offices, integrated service centers, and DOE laboratories with an enterprise-wide Searchable FWP system to support R&D funding requests, facilitating online submission, tracking, and searching.
- Based on an interim Searchable FWP system developed in FY08, maintain a new operational system in FY09 and beyond, with improved functionality and enhancements.
- Extend application to university grant submissions and the wider DOE program community to meet enterprise-wide needs where feasible.

Guiding Principles

OSTI management and staff are committed to ensuring that the Nation has access to the scientific and technical knowledge it needs to remain competitive, a key part of the DOE mission. Through innovation, collaboration and management excellence, OSTI plays a key role in establishing systems that result in accelerating the efficient dissemination of scientific and technical knowledge.



Spirit of Innovation

The spirit of innovation emerges in every aspect of OSTI work. It is the "art of the possible" view which takes technology to the next step in every area—from information acquisition, to preservation, to access and intuitive presentation of search results.

OSTI recognizes that knowledge sharing accelerates the advancement of science and that innovative deployment of information technology accelerates knowledge sharing. OSTI plans to continue pioneering approaches and methods to enhance such sharing of knowledge. Panelists of the *2007 Workshop Panel Report on Accelerating the Spread of Knowledge About Science and Technology: An Examination of the Needs and Opportunities* found that progress in science and technology advances most efficiently when there are mechanisms that enhance the rapid sharing of knowledge to those who can most effectively use it. To advance capabilities in a manner that provides the needed services and support, OSTI will continue to develop and improve systems and tools for gathering, processing, reformatting, describing and packaging information in ways that will enhance its current usefulness and its ability to be reused in the future. Specifically, OSTI is committed to take action to:

- Increase its technical and systems capacity to accommodate a five- to ten-fold surge in DOE-generated R&D information.
- Integrate access to non-textual data through textual information systems.
- Develop the mechanisms to retrieve and make accessible all DOE-sponsored, published R&D literature.
- Introduce translations technology that could search Chinese, Russian, and other languages in future versions of products.
- Pioneer solutions to take federated searching beyond its current capacity limitations and enhance precision searching and Web 2.0 applications, accommodating exponential increases in information while still delivering results in seconds.
- Build on a well-managed, unified information infrastructure to produce superior access to quality content for diverse scientific communities.
- Increase sharing of information on an international basis as a key component of accelerating knowledge diffusion; use collaborative agreements, such as the one begun with the British Library that grew to the WorldWideScience Alliance to advance scientific research.
- Make DOE-funded and other unclassified research results broadly available as soon as possible in dynamic, publicly accessible systems with a goal of making good information available fast.

Implementing Technology Developed through SBIR

OSTI has implemented several innovations that were researched and developed through the DOE Small Business Innovation Research (SBIR) program. Outcomes include federated searching enhanced with powerful relevancy ranking, technology that allows field-specific searching across disparate databases, email alerts, customized content management, and clustering which allows for more intuitive navigation by grouping related results.

Collaboration

OSTI has a track record of success in forging DOE, interagency, international, and private sector collaborations. These collaborations increase communications across communities, helping open access and accelerating the sharing of scientific knowledge.

OSTI will continue to build on existing partnerships and strategic alliances across a range of communities to explore new ideas and solidify scientific exchange. OSTI will also seek opportunities to forge new relationships with a goal to open avenues for researchers and the U.S. public to gain appropriate and ready access to both historic and ongoing R&D information.

DOE Collaboration

Scientific and Technical Information Program (STIP)

Throughout its history, OSTI has worked with representatives across the agency—in laboratories, field offices, Headquarters—to facilitate access to information and to promulgate policy and best practices for STI management. OSTI will work to broaden its complex-wide collaboration by increasing communications beyond the designated STI points of contact and engaging others across DOE in projects that advance STIP goals.

DOE Numeric Data Community

Scientific and technical information *(STI)* includes not only text documents but also numeric scientific data. In 2004, OSTI began placing a special emphasis on working with the scientific data community and hosted a meeting with the DOE Data Centers to discuss access and long-term preservation practices. OSTI facilitates communications on these issues both within DOE and with interagency groups, and has worked toward integrating access to data with text in the framework of the STI Program. In 2008, OSTI developed the DOE Data Explorer, an information tool that can be used to find collections of scientific research data, such as computer simulations, figures and plots, interactive maps, multimedia, numeric files, and scientific images. OSTI will continue to collaborate with DOE Data Centers and other data collection owners to identify DOE data collections wherever they reside and work towards integrated searching of DOE Data Center websites. OSTI will continue to support and raise the visibility of DOE Data Centers and will actively work toward linking published research to its supporting data.

Small Business Innovation Research (SBIR)

DOE's SBIR/Small Business Technology Transfer Research (STTR) program defines topical areas in which both the federal and commercial sectors could benefit from

innovation. With the connection between scientific progress and the dissemination of scientific information, OSTI can use SBIR/STTR information-related innovation to advance science. OSTI will specifically look to deploy technologies and innovation emanating from SBIR/STTR grants related to scientific knowledge diffusion.

DOE Scientists at Labs and Universities

Many OSTI products, though usable by all communities, have historically been developed to serve information intermediaries rather than directly involving the research community. Current and new OSTI products are being developed to serve individual scientist and researcher needs at labs and universities. These products will include the collaboration tools that are often critical to scientists in their peer-to-peer communications.

Interagency Collaboration

CENDI

OSTI was a charter member in 1985 of CENDI and continues to be an active member of the community of federal agencies and organizations responsible for handling the Nation's federally funded STI. This interagency working group of senior STI Managers from 13 U.S. federal agencies represent approximately 97 percent of the federal R&D budget. While OSTI continues to fulfill statutory responsibilities of sharing DOE's research information with the National Technical Information Service and the Government Printing Office, OSTI will continue to collaborate with other federal agencies in support of our mission, specifically to seek partnerships on innovations that advance knowledge diffusion.

Science.gov Alliance

OSTI championed the concept of a U.S. science portal and convened the 2001 workshop where the Science.gov Alliance was founded. OSTI hosts Science.gov, an interagency portal with 18 U.S. government science organizations within 14 federal agencies that form the voluntary Science.gov Alliance. OSTI will continue to collaborate with other federal agencies to increase the amount of science information readily findable and to leverage technology innovations to accelerate discovery.

International Collaboration

WorldWideScience Alliance

Initially established as a bilateral agreement between the U.S. and the U.K., WorldWideScience.org became a multilateral alliance.

OSTI will continue to play a key role in the WorldWideScience Alliance by hosting and maintaining the WWS.org search engine and using its international expertise to maximize U.S. access to global science information.

Energy Technology Data Exchange (ETDE) and International Nuclear Information System (INIS)

Since 1987 and 1970, respectively, OSTI has represented DOE and the U. S. in two multilateral information exchange agreements under the auspices of the International Energy Agency *(IEA)* and the International Atomic Energy Agency *(IAEA)*. Through these agreements, OSTI shares U.S. scientific output in exchange for other countries' R&D literature. As a result, the U.S. science community gains access to 100,000 foreign research

records annually. OSTI will continue these partnerships as the foundation of its international R&D information collection, while offering additional capabilities to enhance the use and visibility of the ETDE and INIS collections.

International Council for Scientific and Technical Information (ICSTI)

Just as CENDI provides a national forum for federal STI organizations to share and benefit from common experiences, ICSTI is the international body that brings together both governmental and non-governmental STI stakeholders globally. This forum provides OSTI with valuable insight into international trends and STI practices, including open access, digital preservation practices, intellectual property law, and web innovations. OSTI will use ICSTI as a primary vehicle to support its efforts to maximize U.S. access to international scientific information. Specifically, ICSTI will serve as a chief sponsor to the WorldWideScience Alliance and OSTI's strategic push to develop federated searching of globally dispersed STI collections.

Private Sector Collaboration

Library Associations

Through participation in the American Library Association *(ALA)* and Special Libraries Association *(SLA)* activities, OSTI has increased awareness and helped expand dissemination of STI through the librarian community. OSTI has been visible at conferences to exhibit and answer in-depth questions about specific, key scientific databases and portals. Participation in the 2008 Government Documents Roundtable, an ALA group specializing in government information, showcased the Department's research findings. OSTI reaches out to the librarian community by facilitating access to research results by making available MARC *(MAchine-Readable Catalogue)* records and Open Archives Initiative records for reports in the Information Bridge. These records, freely accessed from the OSTI web page, can be added to library catalogs, thus greatly increasing public accessibility to DOE STI.

Conventional Search Engines

OSTI made a special effort to open its government databases to conventional search engines. OSTI databases have enormous content and powerful search features. Google and Yahoo! attract huge numbers of searches, but conventional search engines cannot search the OSTI content because they are surface web crawlers and not federated search applications. Recognizing the value of government information and the challenge of exposing database content led Google to produce the Sitemap Protocol. The Sitemap Protocol provides a mechanism for content owners to permit web crawlers to access and index content that is normally only available to federated search. With the Sitemap Protocol, DOE research results are more readily available through conventional search engines Google and Yahoo! Both search OSTI's deep web database collection, sending to public desktops a vast amount of DOE science information. Prior to these partnerships, documents inside the deep web were not easily

"…the Department of Energy's Office of Scientific and Technical Information operates a large database that makes research and development findings available to the public. OSTI developed a Sitemap for its Energy Citations and Information Bridge services in just 12 hours, opening up 2.3 million bibliographic records and full-text documents to crawling by search engines. After its implementation of Sitemaps, OSTI saw a dramatic increase in traffic to its services…"

J.L. Needham, head of Google's public sector content partnership, testimony to the Senate Homeland Security and Governmental Affairs Committee, December 11, 2007

accessible by the public through these conventional search engines.

CrossRef

In 2005, OSTI entered into an agreement with CrossRef, a nationally recognized reference-linking service. OSTI was the first government member of CrossRef. OSTI agreed to join CrossRef in order to facilitate access to DOE's vast stores of STI via references in journal articles. OSTI and CrossRef use Digital Object Identifiers *(DOIs)* to facilitate the access to and reuse of science research reports available electronically in OSTI's Information Bridge.

NFAIS

In 1958 President Eisenhower directed the National Science Foundation to ensure the provision of indexing, abstracting, translation, and other services leading to a more effective dissemination of scientific information.

Fourteen organizations, including OSTI, were convinced of the value of mutual interaction and the interchange of ideas and expertise, thus, a new organization—the National Federation of Science Abstracting and Indexing Services *(NFSAIS)*—was formed. Over the years, the organization has adjusted its name, and NFAIS membership has come to include international scholarly associations, public and private companies, libraries, major corporations, and government agencies. Today, OSTI participates in NFAIS to keep abreast of developments in this area.

SPARC

OSTI works in partnership with the Scholarly Publishing and Academic Resources Coalition *(SPARC)* to broaden access to STI. SPARC encourages expanded dissemination of research and reduced financial pressure on libraries and touts a worldwide membership of over 800 institutions and organizations. The OSTI E-print Network was selected as a Scientific Communities Partner. The partnership aims to enhance search, expand audience, ease scientific collaboration, and advance science.

WorldCat

OSTI made available for the first time DOE STI to library catalogues around the world in 2006. Records for DOE STI are now available in WorldCat, a bibliographic database used for cataloging by more than 50,000 libraries in 96 countries and territories. The WorldCat database was built and is maintained collectively by member librarians of the Online Computer Library Center *(OCLC)*. The OCLC is a nonprofit, computer library service and research organization dedicated to the public purposes of furthering access to the world's information and reducing information costs.

Management Excellence

Management excellence is the driving force behind OSTI innovation.

The key to OSTI's success and the driving force behind making the web work better for science is an unrelenting commitment to excellence. OSTI is a results-driven organization, relying on stakeholder feedback for guidance, collaborations for expansion of innovation, and effective, efficient management. OSTI aggressively works to capitalize on current technologies to provide the tools and services needed to accelerate the sharing of scientific and technical information to speed the advancement of science.

OSTI consistently maintains performance metrics demonstrating its value and cost-effective operations *(see Measures of Progress, p. 16)*. This performance is aided by an entrepreneurial spirit among employees, who have a sense of urgency about innovatively serving our patrons' needs and a drive to remain best-in-class in electronic government, particularly government-to-citizen initiatives. OSTI has a flat management structure, which facilitates decision making and maximizes organizational flexibility. OSTI values diversity and performance excellence, and treats its contractors as partners, which results in the entire organization working toward common goals.

In terms of human resources *(HR)*, OSTI maintains lean federal staffing resources and uses contractors to adapt to increases and reductions in workload. Like most other federal organizations, a large segment of OSTI's workforce is now at or nearing retirement eligibility. OSTI will meet the challenge of staffing succession by taking advantage of a full range of HR flexibilities, including co-op arrangements, cross-training, and contractor support.

Likewise, the OSTI "workplace" is also important in contributing to successful mission performance, and OSTI's primary facility in Oak Ridge, Tennessee, is continually adapted to meet technical and staffing requirements. Still, the facility, built in 1947, requires expensive capital replacements and improvements, and OSTI will use the DOE budgeting process to request funding needed to maintain the facility in a high-quality condition.

Because OSTI is dedicated to quality service to its patrons, OSTI works to assure that anticipated exponential increases in the volume of information available and the volume of traffic utilizing that information can be accommodated.

Quality and innovation are also achieved by working with external panels, committees of visitors, community members, and other expert individuals and groups *(see Appendix 1, Strategic Planning: Basis for Action, p. 54)*. In managing the DOE-wide program for information collection, preservation and dissemination, OSTI regularly relies on a multitude of partners, including subject matter experts in science and technology, classification and review, document tracking and retrieval systems, library services and records management.

Budget

Science in America faces a crisis—a crisis of awareness and support among the public. Such support is essential to secure increased and sustained funding for science as the foundation for future solutions to energy, environmental and national security challenges. Yet, we live in an age when the average American, unlike in previous generations, would have difficulty identifying any of our Nation's leading scientists.

This lack of awareness is not a result of U.S. antipathy toward science. Quite the opposite: America has a long and proud tradition of leadership in science and the underpinnings of science, technology, math, and engineering education. Instead, it is a result of not demonstrating or communicating the linkages between scientific achievement and our standard of living.

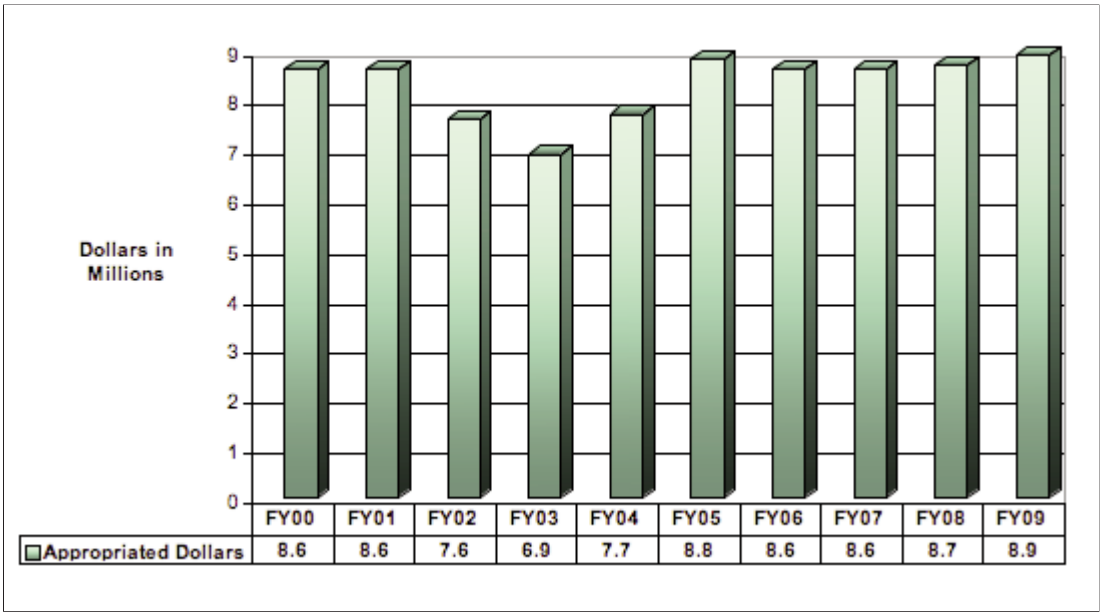
Moreover, while the pace of scientific achievement is increasing, the promise of truly accelerating progress in unimagined ways—ways that will capture the minds and attention of Americans—hinges on our ability to accelerate the rate at which scientists build on each others' work. This linchpin role belongs to OSTI. Just as the National Library of Medicine *(NLM)* is the conduit for communicating scientific discovery for the National Institutes of Health *(NIH)*, OSTI is the vehicle for communicating and accelerating the scientific achievements of DOE.

Specifically, implementing the OSTI "Accelerated Science" strategies will involve a complex and multi-pronged effort, including accommodating major increases in the output of scientific content as well as new non-textual data types. It will involve extending the sophistication of federated search technology to both worldwide scientific information and science education resources; enhancing the ability of computers to make sense of and perform precision search and ranking of massive quantities of R&D information; and then strengthening the technical infrastructure to support uninterrupted performance of more complex computer functions and increased web transactions. It will involve communications systems and media to inform the U.S. public about the dividends from their investment in science.

A significant investment in new tools and capabilities is essential to the Office of Science's demonstration of the return on an increased R&D budget. The knowledge that will be created from this investment will only be useful if it is accessible; if it can be precision searched; if it is available across scientific communities; and if it has a facility to move easily between data and text.

On its limited resources and with extreme efficiency, OSTI has had an undeniable impact on access to and discovery of scientific knowledge over its more than 60-year history and, particularly, over the last 10 years—now performing 84 million web transactions annually. OSTI welcomes the opportunity to demonstrate the possibilities in elevating science through the strategies outlined in this plan.

Budget by Fiscal Year *(Chart 3)*



External Reviewer Comments

OSTI's "Strategic Plan: FY2009–FY2013" is based on findings and evolved in large part from eight panels of experts convened in recent years to explore best pathways forward in mission responsibilities *(see Appendix 1, Strategic Planning: Basis for Action, p. 54)*. Once the plan was drafted, OSTI invited a broad range of stakeholders to provide feedback.

Three Focus Group meetings were held December 8–10, 2008, at DOE Headquarters in Washington, D.C., with participants from DOE programs, other government agencies, universities and the private sector. Participants provided verbal feedback as well as written feedback in the form of a decision-support tool developed by OSTI. Other stakeholders who could not participate in the meetings provided online feedback to the decision-support tool, which was available until January 13, 2009. Additionally several people offered individual comments or suggestions but did not provide responses using the tool.

Stakeholder input is reflected throughout the plan. The statements in the front matter, such as those dealing with OSTI's vision, mission, foundation, challenge, and the OSTI corollary, were agreed upon by an overwhelming majority of respondents. Among stakeholders, those within the DOE and DOE contractor community, being more familiar with the Department's long-standing STI program, provided many helpful comments and gave OSTI's goals the highest ratings, indicating they highly valued OSTI's efforts for filling the gaps. The few respondents who did not agree with certain statements came from the interagency or private sector, with indications of either misunderstanding about the statements or questions about OSTI activities or other aspects alluded to in the text rather than disagreement of the plan itself. Those comments were very valuable in refocusing or adjusting the goals and objectives, and their feedback has been accommodated throughout this document. The Strategic Plan now reflects the culmination of suggestions and comments of many stakeholders.

Common Themes

A total of 19 gaps that need to be addressed by OSTI are identified in this plan, within three areas: Information Technology, S&T Content, and Communication/Outreach *(see Gaps To Be Filled, p. 19)*. While a wide range of feedback indicated differences among the stakeholders' perspectives, some themes emerged of mutual interest among the focus groups.

The two technology gaps that were ranked as High Importance by the majority of stakeholders are:

- Great quantities of scientific and technical information *(STI)* are non-Googleable.
- A new class of web tools to help researchers collaborate and access, analyze and manipulate data is needed.

The two content gaps that gained the greatest number of High rankings are:

- Non-textual forms of STI *(audio, visual, image, numeric data, and illustrations)* are not searchable and accessible.
- Hundreds of thousands of DOE technical reports have not been digitized.

The two communications/outreach gaps that were ranked the highest by the majority of respondents are:

- University research programs are not equipped with knowledge of DOE STI resources.
- Public awareness of DOE and Office of Science researchers and their contributions is lacking.

All six goals in this plan received a good rate of agreement by most of the stakeholders. The goal that ranked as Highest importance by the greatest majority was Goal 1, which was stated:

- Accelerate the spread of scientific and technical knowledge. Due to stakeholder feedback, this goal has been re-stated: Ensure superior access to and preservation of quality scientific and technical content.

The key points from the stakeholder group have been incorporated within the plan, including clarifying throughout the document these key points made by our stakeholders:

- The content collected, preserved and disseminated by OSTI—the output of DOE programs—is a public good.
- Content is the primary focus, with information technology second, as the enabler or means.
- The overarching goal is to ensure that patrons "get good information fast."
- The "web" is a "catch all" term meant to indicate the broader information infrastructure or digital networked environment.
- Partnerships and alliances of OSTI's add value to OSTI and to the work of others, especially in niche areas.
- OSTI uses today's tools of choice for sharing quality content as efficiently as possible.
- OSTI develops technology in narrow niche areas to fulfill special mission needs but deploys, and is often the first-adopter of, technologies to advance the OSTI mission.

DOE OSTI Strategic Plan Stakeholders

OSTI gratefully acknowledges the participation and feedback provided by various stakeholders who offered constructive comments and especially appreciates the enthusiasm and contributions of those stakeholders who participated in the Focus Group discussions.

Department of Energy

Headquarters Programs

Justine Alchowiak	*Office of Environmental Management*
Lauren Boyd	*Office of Energy Efficiency and Renewable Energy*
Aleta Boylan	*Office of Headquarters Procurement Services*
Jay Braitch	*Office of Fossil Energy*
Vince Dattoria	*Office of Science*
Walter Dykas	*Office of Science*
Tim Fitzsimmons	*Office of Science*
Paul Gottlieb	*Office of General Counsel*
Wendolyn Holland	*Office of Energy Efficiency and Renewable Energy*
Allan Jelacic	*Office of Energy Efficiency and Renewable Energy*
Tom Malec	*Office of Intelligence and Counterintelligence*
Darren Mollot	*Office of Fossil Energy*
William Turnbull	*Office of the Chief Information Officer*
Bill Valdez	*Director, Workforce Development for Teachers & Scientists, Office of Science*
Linda Wilbanks	*Chief Information Officer, National Nuclear Security Administration*
Collis Woods	*Office of Science*

DOE Field and National Laboratories

Michael Angulo	*Pacific Northwest Site Office*
Luis Bettencourt	*Los Alamos National Laboratory*
Miriam Blake	*Los Alamos National Laboratory*
Gayla Bratton	*Hanford Site*
Leslie Brown	*Washington Closure Hanford*
Nicole Carson	*Nevada Test Site*
Debbie Caver	*Savannah River Operations Office*
Dale Claflin	*Idaho National Laboratory*
Theresa Connaughton	*Los Alamos National Laboratory*
Tony DeMeo	*Princeton Plasma Physics Laboratory*
Mary Donahue	*National Renewable Energy Laboratory*
Nancy Doran	*Pacific Northwest National Laboratory*

Alison Easter	*Ames Laboratory*
Marsha Fish	*Washington Closure Hanford*
Janine Ford	*NNSA Service Center*
Scott Gibson	*Oak Ridge Associated Universities*
Dennis Gound	*Oak Ridge National Laboratory Site Office*
Donald W. Guy	*Sandia National Laboratory*
Mary Hale	*Argonne National Laboratory*
Mona Higgins	*Nevada Site*
Dave Hamrin	*Oak Ridge National Laboratory*
Isom Harrison	*Lawrence Livermore National Laboratory*
Judy Hulstrom	*National Renewable Energy Laboratory*
Bob James	*East Tennessee Technology Park*
Kimberly Johnson	*Argonne National Laboratory*
Kim Kindrew	*Thomas Jefferson National Accelerator Laboratory*
Donna Lawson	*Y-12 National Security Complex*
Miriam Legan	*Chicago Operations Office*
Rosemary Lowden	*Lawrence Berkeley National Laboratory*
Kathy Macal	*Argonne National Laboratory*
Jerry Massee	*Los Alamos Site Office*
Heath O'Connell	*Fermi National Accelerator Laboratory*
Clayton Olgilvi	*Idaho Operations Office*
Danielle Pollock	*Sandia National Laboratories*
Mary Petersen	*Brookhaven National Laboratory*
Bill Prymak	*Golden Field Office*
Nicole Rantz	*Lawrence Livermore National Laboratory*
Betty Robinette	*Y-12 National Security Complex*
Annanaomi Sams	*Pacific Northwest National Laboratory*
Bonna Savarise	*Yucca Mountain Site Characterization Office*
Kevin Schmidt	*Savannah River Site*
Jessica Shaffer-Gant	*Sandia National Laboratories*
Donna Sharpe	*East Tennessee Technology Park*
Jane Tierney	*Lawrence Berkeley National Laboratory*
Sharon West	*SLAC National Accelerator Laboratory*
Cathy Wright	*National Energy Technology Laboratory–Albany*

Interagency

Nancy Allard	*National Archives and Records Administration*
JoAnne Calhoun	*National Aeronautics and Space Administration*

Blane Dessey	*Justice Management Division, U.S. Department of Justice*
Tamas Doszkocs	*Office of the Director, National Library of Medicine*
Michael Frame	*Center for Biological Informatics, U.S. Geological Survey*
Eleanor Frierson	*Deputy Director, National Agricultural Library, U.S. Department of Agriculture*
Cita Furlani	*Chief Information Officer, National Institute of Standards and Technology*
Glenn Gardner	*Library of Congress*
Tina Gheen	*Library Director, National Science Foundation*
Bev Godwin	*Director, USA.gov , General Services Administration*
Chris Greer	*Director, National Coordination Office of Networking and Information*
Don Hagan	*Associate Director, National Technical Information Service, U.S. Department of Commerce*
Ellen Herbst	*Director, National Technical Information Service, U.S. Department of Commerce*
Thomas Lahr	*BIO Deputy Director, U.S. Geological Survey*
Julia Lane	*Science of Science and Innovation Policy (SciSIP), National Science Foundation*
Stephanie Madden	*General Services Administration*
Cora Marrett	*National Science Foundation*
Susan McCarthy	*National Agricultural Library, U.S. Department of Agriculture*
Gulnar Nagashybayeva	*National Oceanic and Atmospheric Administration*
Chuck Romine	*Office of Science and Technology Policy, Executive Office of the President*
Roberta Shaffer	*Executive Director, FLICC-FEDLINK, Library of Congress*
Elliot Siegel	*Associate Director, Health Information Programs Development, National Library of Medicine*
George Strawn	*Chief Information Officer, National Science Foundation*
John Sykes	*U. S. Environmental Protection Agency*

University Participants

Elizabeth Aversa	*University of Alabama*
Rick Borchelt	*Genetics and Public Policy Center, Johns Hopkins University*
Christine Borgman	*University of California, Los Angeles*
Mel DeSart	*Engineering Library, University of Washington*
Susan Fingerman	*RE Gibson Library and Information Center, Johns Hopkins University Applied Physics Laboratory*
Roland Horne	*Stanford University*
David Kaiser	*Massachusetts Institute of Technology*
Linda Kennedy	*Peter J. Shields Library, UC/Davis*
Michael Lesk	*Rutgers University*
Sandra McAninch	*University of Kentucky*
Kirk McDonald	*Princeton University*

Maliaca Oxnam	*Science-Engineering Library, University of Arizona*
John Regazzi	*Long Island University*
Nancy Roderer	*Johns Hopkins University*
John Saylor	*Cornell University*
David Stern	*Associate University Librarian for Scholarly Resources, Brown University*
Peter Suber	*Earlham University*
Carol Tenopir	*School of Information Sciences, University of Tennessee*

Private Sector

Prudence Adler	*Federal Relations and Information Policy, Association of Research Libraries*
Gil Baldwin	*Retired, Government Printing Office*
Roberta Balstad	*Chair, US CODATA, Consultant to Columbia University*
Wally Finch	*Consultant; Retired National Technical Information Service*
Beth Fitzsimmons	*National Commission on Libraries and Information Science*
Mark Hyer	*ProQuest*
Richard Johnson	*Scholarly Publishing & Academic Resources Coalition*
Robert Kahn	*Corporation for National Research Initiatives*
Dick Kaser	*Information Today, Inc.*
Larry Lannom	*Corporation for National Research Initiatives*
Clifford Lynch	*Coalition for Networked Information*
Robert Marianelli	*Consultant*
Heather Joseph	*Executive Director, Scholarly Publishing and Academic Resources Coalition*
Jim McGinty	*Vice Chairman, Cambridge Information Group*
Dave Penniman	*Executive Director, Nylink*
Alvin Trivelpiece	*Consultant; Formerly Director, Oak Ridge National Laboratory and Director, DOE Office of Energy Research*
Paul Uhlir	*CODATA –The National Academies*
Rick Weingarten	*Director, Office for Info Technology Policy, American Library Association*

International

Pam Bjornson	*Canada Institute of Scientific and Technical Information (CISTI)*
Michael Ireland	*Director of Information Access and Delivery, Canada Institute of Scientific and Technical Information (CISTI)*

Glossary

AEC
Atomic Energy Commission

CENDI
An interagency working group of senior scientific and technical information *(STI)* managers from 13 U.S. federal agencies

Deep Web
Where most scientific research resides

DoD
The U.S. Department of Defense

DOE
The U.S. Department of Energy (www.energy.gov)

DOE Data Explorer
(www.osti.gov/dataexplorer)

DOE R&D Project Summaries
(www.osti.gov/rdprojects)

DOE Science Accelerator
(www.ScienceAccelerator.gov)
A federated search engine of key DOE databases

DOIs
Digital Object Identifiers

DTIC
Defense Technical Information Center, centralized scientific, technical and related defense information services to support the defense community

Energy Citations Database
(www.osti.gov/energycitations)

EnergyFiles
(www.osti.gov/energyfiles)

Energy Science and Technology Software Center
(www.osti.gov/estsc)

Enterprise Secure Network
DOE Enterprise Secure Network system, classified network connectivity between a number of DOE/NNSA sites

E-print Network
(www.osti.gov/eprints)

ERDA
Energy Research and Development Administration

ESNet
Energy Sciences Network, a high-speed network serving thousands of DOE scientists and collaborators worldwide

ETDE
Energy Technology Data Exchange https://www.etde.org/etdeweb/, an international energy information exchange agreement formed in 1987 under the International Energy Agency *(IEA)*

Federal R&D Project Summaries
(www.osti.gov/fedrnd)

Federated Search
The simultaneous search of multiple web databases in real time via a single search query

Feynman Diagrams
Invented by physicist Richard Feynman, pictorial representations of quantum field theory processes in terms of particle paths

Fielded Search
A type of search technology that specifies the range by metadata element (ex: author, title, date) of a search query to increase the relevance of the search results

Global Discovery
Bringing together the best science from all corners of the earth and making it easily searchable, regardless of the media in which it resides, to speed the diffusion of knowledge

Harvesting
Human expertise or machine guidance to direct the web crawlers to specialized collections or sets of knowledge, can be viewed as focused or directed web crawling

ICSTI
International Council for Scientific and Technical Information, the primary sponsor of WorldWideScience.org

Information Bridge
(www.osti.gov/bridge)

INIS
International Nuclear Information System (http://www.iaea.org/inisnkm/)

ISC
Integrated Support Centers, virtual combination of field offices to provide administrative, business, and technical services in support of the Office of Science complex

MARC
MAchine Readable Cataloging format (www.osti.gov/marcrecords)

NFAIS
National Federation of Science Abstracting and Indexing Services

NIH
National Institutes of Health

NLM
National Library of Medicine

NNSA
National Nuclear Security Administration

Non-Googleable
A term coined by OSTI, information which cannot be found through conventional search engines, such as Google, Yahoo! and MSN

OCLC
Online Computer Library Center

OSTI
Office of Scientific and Technical Information
(www.osti.gov)

OSTP
Office of Science & Technology Policy
(www.ostp.gov)

PII
Personally Identifiable Information

R&D
Research and development

Relevancy Ranking
The technology that allows search results to be returned in a ranked order relevant to the search query

SBIR
Small Business Innovation Research
(www.sbir.gov)

SC
Office of Science
(www.science.doe.gov)

SCejournals
A website providing the Office of Science electronic access to key peer-reviewed scientific journals

Science Conference Proceedings
(www.osti.gov/scienceconferences)

Science.gov
(www.Science.gov)
A federated search engine of U.S. government science agencies

Science knowledge imperative
Spreading the world's great scientific discoveries as fast as possible

Searchable FWP
Field Work Proposal***Sitemap Protocol***

A special method for informing conventional search engines of pages available for them to search and retrieve

STI
Scientific and technical information

STIP
Scientific and Technical Information Program
(www.osti.gov/stip)

STTR
Small Business Technology Transfer Research
(www.sc.doe.gov/sbir)

WorldWideScience.org
(www.WorldWideScience.org)
A federated search engine of science sources around the world

Appendix 1

Strategic Planning: Basis for Action

In recent years, OSTI has convened eight distinguished panels of experts in science, science policy, information science, and scientific publishing to explore best pathways forward in mission responsibilities. These workshops helped form and solidified OSTI mission goals and objectives and stimulated major changes in the mode and manner of capturing information to better disseminate scientific knowledge. The Strategic Plan is based on findings and evolved in large part from these workshops:

- In May 2000, the Workshop on a Future Information Infrastructure for the Physical Sciences at the National Academy of Sciences, chaired by Alvin Trivelpiece, was held to examine the need for an infrastructure to facilitate scientific communication and increase productivity of the scientific enterprise in the U.S. Two federated search tools were created in response.
- In April 2001, a Workshop on Strengthening the Public Information Infrastructure for Science at the National Institute of Standards and Technology *(NIST)* examined the infrastructure for public access to science information of the Federal agencies. This workshop led to Science.gov.
- In June 2002, a Board of Visitors convened at OSTI for an STI program review. Individual board members submitted reports noting OSTI's recent accomplishments, partnerships and contributions to diverse user groups. A number of recommendations to help OSTI position itself in the mainstream web were adopted and implemented.
- In July 2004, data management experts from DOE data centers and other DOE organizations discussed data access and preservation as integral to research programs and projects, issued in The State of Data Management in the DOE Research and Development Complex Report. As a result, the DOE Data Explorer was developed and launched.
- In June 2006, OSTI laid out the concept of accelerating science access by accelerating the sharing of science in the DOE Science Accelerator document, which was distributed widely to key stakeholders for comment. Subsequent to feedback from these stakeholders, a workshop was planned for 2007.
- In February 2007, a workshop panel chaired by Alvin Trivelpiece at the National Academy of Sciences examined strategies to accelerate the spread of knowledge about science and technology. An essential finding of the Workshop Panel on Accelerating the Spread of Knowledge about Science and Technology: An Examination of the Needs and Opportunities was: "Superior access to quality content is key to accelerating science and technology." The new Strategic Plan began to be developed soon after this workshop.
- In December 2008, OSTI gathered three focus group sectors *(DOE, Interagency, Private/University)* to provide feedback on this draft Strategic Plan: FY 2009–2013. Stakeholders provided feedback in a number of areas relating OSTI's plans to facilitate the sharing of science information. OSTI incorporated the feedback into the Strategic Plan: FY 2009–2013 document and is actively pursuing actions related their feedback.
- In March 2009, OSTI convened a meeting of science education representatives from federal government agencies at the National Academy of Sciences to examine the opportunity to provide a web-based interagency portal to federal STEM education resources in support of training the next generation of scientists and engineers.

Appendix 2

Regardless of Form, OSTI Search Tools Find Scientific and Technical Information

Scientific and technical information *(STI)* has many forms, such as journal articles, technical reports, patents and e-prints. Each form is generated in a unique way, has its own publication route and its own life cycle, which requires its own method of acquisition and dissemination. Thus, it is necessary for OSTI to create a distinct information tool to virtually collect each form and to make it efficiently accessible to users.

OSTI creates two different types of electronic collections. The first type is more like a traditional library in that OSTI compiles a collection of STI produced by or funded under the provenance of the DOE issued by an institution, and residing on an OSTI computer. OSTI controls what goes into these collections and in what format. The OSTI databases that are of this sort include the full-text documents in the Information Bridge and the bibliographic citations and summaries created for the Energy Citations Database, DOepatents, and the DOE R&D Project Summaries. The second type of electronic collection is a virtual collection of STI outside of DOE. These collections contain STI that is of interest to DOE, but, for the most part, is not produced by DOE. The citations and full-text documents in these virtual collections reside on the Internet in servers all over the world. OSTI has identified the locations of the STI and provides a means to search these far-flung collections. Searching such virtual collections requires a different type of search engine than searching a single collection. It requires knowing where to search, what to search and how deep to search. This is new technology that OSTI has had a hand in developing. Examples are Science.gov, the E-print Network, Science Conference Proceedings, and EnergyFiles. OSTI remedies the uncertainty about which of its various search tools one should use by providing the Science Accelerator, a federated search engine that covers ten OSTI databases. Entering a single query will bring back results from each OSTI-produced database containing relevant citations. Federated search helps the information consumer decide which of the OSTI databases might be appropriate for more detailed searching.

Technical Reports

This is a broad category that includes a wide variety of types of reports. For the most part, technical reports describe the progress or results of scientific or technical research and development. They often serve as a report of accountability to the funding organization. Technical reports fall into two categories: government sponsored or privately funded. Both categories include domestic and foreign reports written by researchers at universities, institutes, private companies, and government agencies and labs. The different types of technical reports include technical background reports, primary research reports, including progress reports and final reports, technical specifications, instructions, user guides, and reports from government advisory committees and national and international societies.

Discovering and locating technical reports can be difficult. They are usually not included in the standard subject databases covering books and journal articles. Prior to the advent of the web, technical reports were usually produced in very limited quantities with limited dissemination. Government technical reports were often distributed in microfiche. Only some of the larger research libraries have extensive collections of technical reports and usually these collections are uncataloged. Today, most recent technical reports and some older material are accessible via the web, if you know where to look. Most of these reports reside in the deep web and are not discovered by the standard web search engines.

OSTI provides access to full-text DOE technical reports through the Information Bridge from 1991 forward, with selected older reports. The Energy Citations Database indexes older DOE, ERDA and AEC technical reports back to the 1940s, as well as energy and nuclear science related reports from other government agencies and foreign governments. Virtual technical reports collections not under the provenance of DOE can be searched by Science.gov, EnergyFiles, the E-print Network and WorldWideScience.org.

Journal Literature

The publication of research in scientific journals started in the mid-seventeenth century. Before that and for some time after, STI was circulated via letters, printed tracts and books. Journals became a preferred medium because journal publishers worked to achieve wider dissemination and faster publication. Today, however, even with the tremendous growth in scientific journals in the latter half of the twentieth century, publishing in scientific journals is often not a speedy process. It can often take a year or more for an article to be published once it has been accepted by a journal. For this reason, many scientists and engineers also utilize other means to share their research. Options include technical reports, conference papers, preprints and a growing use of e-prints.

From 1948 to 1976, OSTI's predecessors under the Atomic Energy Commission published Nuclear Science Abstracts, providing comprehensive indexing of the international nuclear science literature, including journal literature on a worldwide basis. This literature can now be found using Energy Citations Database *(ECD)*. ECD also indexes articles written by DOE employees and contractors since 1977. So while OSTI does provide some coverage of journal literature, for the most part, scientific and technical journal articles are extensively covered in discipline-related databases produced by commercial publishers or scientific societies. Preprints of journal articles can be found using the E-print Network and Information Bridge. Journal literature is also covered in databases searched by EnergyFiles and Science.gov.

Conference Proceedings and Papers

An excellent way for researchers to share their ongoing or completed projects is by delivering papers or making presentations at scholarly conferences. Thus, conference proceedings are excellent sources of STI. They are also the bane of librarians everywhere. Identifying libraries owning specific volumes of proceedings can be exceedingly difficult. This is due to limited dissemination, poor citations, conference title variations, and changes in publishers, locations and funding organizations. In the past, the problems with locating conference papers often dissuaded people from pursuing the information. Today, conference proceedings and individual conference papers are often available on the web, greatly improving access. Still there are many of challenges to providing bibliographic control of this material.

OSTI addresses the problems with finding and accessing conference proceedings and papers by providing Science Conference Proceedings. This search tool seeks out the conference proceedings and papers on websites and in databases hosted by scholarly societies, institutes and national labs with a strong interest in the physical sciences. Full text coverage varies, depending on the publishing organization. Conference papers are also found as e-prints using the E-print Network. Conference

papers or presentations prepared by DOE-funded researchers may also be accessible in the Information Bridge.

E-prints

E-prints are scientific or technical documents that are circulated electronically, usually by the author. This may include pre-publication versions of various document types, such as technical reports, conference presentations, journal article preprints, and other electronic documents that have not been formally issued. Preprints are an important category of e-prints. Making the text of a journal article available electronically prior to its eventual publication in a journal allows the research results to be available much sooner. This electronic dissemination is done to facilitate peer exchange and scientific advancement.

Many researchers are placing copies of all their publications and reports on their personal web pages, mixing published articles with unpublished reports and presentations. Government agencies are also collecting the publications and reports from their employees and contractors and making them accessible through databases or web pages. Private companies and research organizations are doing the same with their own employees, although they do not always make the documents publically accessible. These efforts have resulted in an enormous amount of full-text STI being available via the web that is extremely difficult to bring under bibliographic control.

The standard web search engines work best for e-prints when searching for a specific title. If the document is on the web and is not in a database or buried too many levels down on a web page, you should have some success finding it. Searching for a specific author is much more challenging due to variation in how names are presented. When seeking scientific or technical e-prints on specific topics, web search engines will usually overwhelm you with non-relevant results.

OSTI has created a very specialized search engine for e-prints that searches websites and databases. The E-print Network searches over 5 million documents from e-print websites. It also searches over 3,100 professional societies in the sciences and engineering. These societies may be browsed by discipline and/or by one of eleven languages. Over fifty institutional repositories and multidisciplinary collections are searched using technology specifically developed for OSTI that reaches into the deep web.

Patents

Patents allow the spread of information about technological inventions while protecting the property rights of the inventor. A patent issued by the U.S. Patent and Trademark Office excludes others from making, using, offering for sale, or selling the invention throughout the U.S. or importing the invention into the U.S. for a limited time in exchange for public disclosure of the invention when the patent is granted. This public disclosure is extremely important in furthering scientific research. Technology moves on, but information remains useful forever.

Thomas Jefferson, an inventor himself and appointed by George Washington to the first Patent Board, was, essentially, the first patent examiner. He found that “the issue of patents for new discoveries has given a spring to invention beyond my conception.”

DOE and its predecessor agencies, ERDA and AEC, are responsible for creating a tremendous amount of new technology. Much of this research resulted in patents. OSTI has created a central collection of DOE patent information from the 1940s to the present. DOepatents contains bibliographic citations with links to the full text where available. Patent topics range from flexible horseshoes to nuclear power stations. Patents and patent applications can also be found in Information Bridge, Energy Citations Database and EnergyFiles.

The E-print Network will include patents from other countries. DOE R&D Accomplishments will highlight some of the more significant DOE patents.

Project Summaries

Some government agencies try to keep the public and the research community informed of their R&D activities by providing summaries of ongoing or recently completed research projects. The summaries generally provide the project title and abstract, the sponsoring agency, the principal investigator and the start and end dates of the project. Some of these projects will eventually generate a technical report, journal article or other literature, but for many of the projects, the summary is all the public will see. These summaries are very helpful in making connections between researchers working on similar projects.

OSTI has created DOE R&D Project Summaries. The information contained in these project summaries is different than the bibliographic data contained in the Information Bridge or Energy Citations Database, and this is why a separate database is required. Since other federal agencies also have somewhat similar databases, OSTI is able to create a virtual collection of federal project summaries with Federal R&D Project Summaries using federated searching to link nine federal agency R&D summary databases.

Theses/Dissertations

Theses and dissertations written by students in pursuit of advanced degrees in accordance with requirement of their academic institutions can be extremely hard to get a firm grip on. There is a commercial publisher that produces a good index to this material, but submission of dissertations and theses is usually voluntary and, each year, a sizable number do not make it into the index. Many universities are now requiring electronic submission of theses and dissertations and maintain institutional repositories. However, many authors choose not to make their dissertation or thesis available via the web, fearing that it will hurt their chances of getting it published as a book. Networking of these digital dissertation repositories has vastly increased accessibility, but it still is not comprehensive.

Searching for theses and dissertations is possible in a number of OSTI resources. Full-text copies of dissertations and theses that received funding support from DOE are found in the Information Bridge. Citations to earlier DOE supported dissertations and theses are found in Energy Citations Database. Energy Citations Database also has extensive coverage of nuclear energy related dissertations and theses from U.S. and foreign universities between 1948 and 1973. Dissertations and theses may also be included in the results from EnergyFiles and the E-print Network, although full-text coverage will vary.

Scientific Research Data

In the course of their research, scientists can compile a great deal of scientific data. Collections of this data may include numeric files, simulations, data plots, figures and scientific images such as those captured during sky mapping or particle collisions. While the results of this research are published in text formats, the scientific data collections can also be of tremendous value to other researchers, students and the public. Where scientific research data collections reside, users often find specialized search interfaces or software toolkits developed by data owners. These allow deeper searches into the data files to help users understand, analyze, and use the data.

The major challenges to researchers using scientific data are discovering the data exists and getting access to the data. Although the Internet has made accessing the data easier, discovering these scientific data files and locating where they reside on the web is still a challenge. For data files created under the sponsorship of DOE, OSTI has removed much of this challenge with the DOE Data Explorer. OSTI has identified scientific data files that are freely accessible at national laboratories, data centers, universities, professional organizations, international organizations, and other websites. Researchers can search or browse the descriptive citations in the DOE Data Explorer and then follow links to the data.

Software

Another form of scientific and technical information that is centrally made accessible by OSTI on behalf of the Department is scientific and technical computer software resulting from DOE-funded R&D projects. Scientific research often includes the development and modification of software, which may be designed to run on personal computers, workstations, mainframes, and supercomputers. Many packages are designed to run in several environments, using different machines and operating systems.

Collection, licensing, and distribution of federally funded software that is developed by national laboratories and other DOE facilities/contractors is managed by the Energy Science and Technology Software Center within OSTI. The software is announced via an online searchable catalog.

Appendix 3

Statutory Authorities

Established in 1947, DOE's Office of Scientific and Technical Information *(OSTI)* fulfills the agency's responsibilities related to the collection, preservation, and dissemination of scientific and technical information emanating from DOE R&D activities. This responsibility has been codified in the organic, or enabling, legislation of DOE and its predecessor agencies and, more recently, was defined as a specific OSTI responsibility in the Energy Policy Act of 2005.

- Energy Policy Act of 2005 *(P.L. 109-58)*, Section 982, called out responsibility of OSTI: "The Secretary, through the Office of Scientific and Technical Information, shall maintain with the Department publicly available collections of scientific and technical information resulting from research, development, demonstration, and commercial applications activities supported by the Department."
- Department of Energy Organization Act of 1977 *(P.L. 95-91)* provided for maintaining a central source of information and disseminating information *(42 U.S.C. Sec. 5916, 7112)*.
- Energy Reorganization Act of 1974 *(P.L. 93-438)* defined responsibilities for developing, collecting, distributing, and making scientific and technical information available for distribution *(42 U.S.C. Sec. 5813, 5817)*.
- Atomic Energy Acts of 1946 *(P. L.79-585)* and 1954, as amended *(P.L. 83-703)* established a program for the dissemination of unclassified scientific and technical information and for the control of classified information *(42 U.S.C. Sec. 2013, 2051, and 2161)*.
- America COMPETES Act of 2007 *(P.L. 110-69)*, Section 1009, required that Federal agencies that conduct scientific research develop agency-specific policies and procedures regarding the public release of data and results of research.
- America COMPETES Act of 2007 *(P.L. 110-69)*, Chapter 5, Section 3191, "National Energy Education Development, *(a)* The Secretary [of Energy] . . . in consultation with the Director of the National Science Foundation shall establish a program to coordinate and make available to teachers and students web-based kindergarten through high school science, technology, engineering, and mathematics education resources relating to the science and energy mission of the Department, including existing instruction materials and protocols for classroom laboratory experiments."

Appendix 4

Federated Search and Directed Harvesting

OSTI combines two unconventional technologies to make the web work for science: (1) federated search and (2) directed harvesting.

The first unconventional technology, federated search, is the simultaneous search of multiple web portals in real time via a single search query. Further value is added by applying sophisticated ranking technology to return highly relevant results. OSTI pioneered federated search of government research to make available the richest science content found in collections that were mostly inaccessible through conventional search engines such as Google, MSN Search, and Yahoo *(see p. 15)*.

The second unconventional technology, directed harvesting, uses human expertise or machine guidance to direct the web crawler to search and to "mine" information from hand-picked sets of scholarly websites. The process includes downloading documents found at these sites, identifying which are scholarly, and cataloging them for search. OSTI harvests documents to increase the number of scholarly documents it makes available to researchers.

OSTI deploys federated search and harvesting as complementary technologies. Federated search works best where there are large silos *(databases or portals)* of scientific and technical information *(STI)* that cannot be easily accessed by conventional search engines. Federated search permits researchers and the public to search multiple databases all at once. Harvesting permits the gathering of documents that do not lend themselves to being searched online in real time. Additionally, and more importantly, harvesting performs document analysis as part of the process of making the documents searchable. Harvesting ensures that the most relevant and highest quality information is made searchable while non-relevant information is weeded out.

Essentially, when OSTI is seeking hard-to-find information from portals and databases, federated search tools are developed and deployed. When OSTI is seeking scholarly documents that must be hand-picked from credible websites around the world, harvesting tools are developed and deployed. An example of a seamless integration of federated search and harvesting into a simple and intuitive search application is E-print Network.

The E-print Network provides searchable access to over 5 million e-prints. E-prints are scholarly and professional works electronically produced and shared by researchers with the intent of communicating research findings to colleagues. They may include preprints, reprints, technical reports, conference publications or other means of electronic communication. Some of these documents are stored in databases and portals, others on websites around the world. By implementing a blend of federated search and web harvesting, OSTI works to make all the documents searchable and retrievable.

This gives E-print Network some key advantages when compared with conventional search engines. Federated search does not place any requirements or burdens on owners of the individual data sources, other than handling increased traffic. And because they are searched in real time, federated searches are inherently as current as the individual data sources. Web harvesting-based searches focus exclusively on quality e-print sources.

Appendix 5

Diffusion Research

OSTI is conducting research to find ways to accelerate the diffusion, or sharing, of scientific and technical knowledge among researchers.

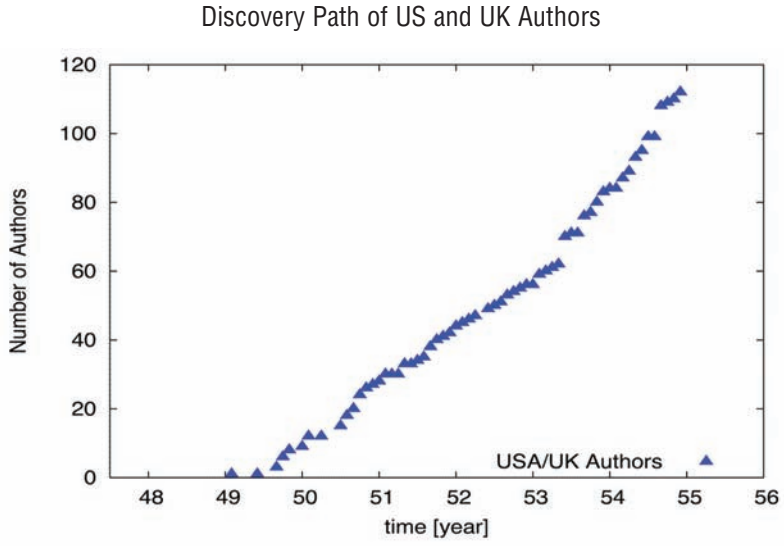
OSTI has used epidemiological research and concepts related to the "spread of infectious disease" to examine whether the spread *(or diffusion)* of information can be measured and modeled and if the "rate of knowledge infection" can increase scientific productivity. OSTI's research suggests that the diffusion of knowledge can be measured and modeled, and can be accelerated by increasing the "transfer rate" among researchers *(see diffusion model charts below)*.

OSTI attempts to increase the "transfer rate" and accelerate diffusion by:

1. Identifying high-quality databases
2. Combining these high-quality databases to act as an integrated whole
3. Compiling search results from the databases and organizing them by relevancy rank

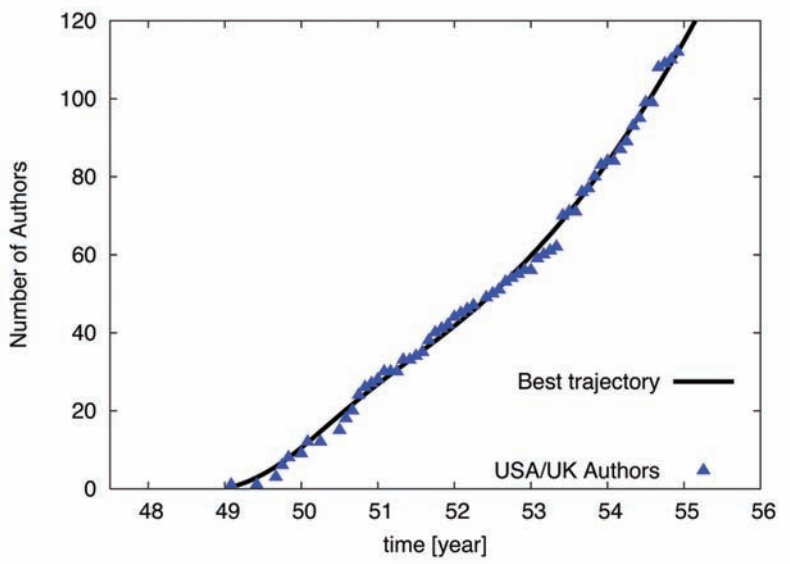
Making STI easily searchable and relevant increases the "transfer rate" of knowledge among researchers; the results of this is expected to increase the pace of discovery.

The Spread of Knowledge about Feynman Diagrams



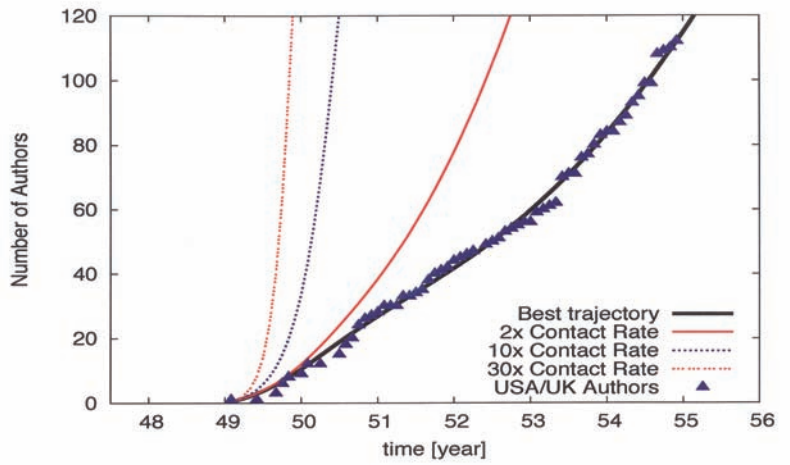
From The Power of a Good Idea: Quantitative Modeling of the Spread of Ideas from Epidemiological Models, Luis M. A. Bettencourt, Ariel Cintron-Arias, Carlos Castillo-Chavez, David Kaiser, May 2005

Path of Best Trajectory



From Report for the Office of Scientific and Technical Information: Population Modeling of the Emergence and Development of Scientific Fields, Luis M. A. Bettencourt, Carlos Castillo-Chavez, David Kaiser, David E. Wojick, October 2006

Paths of Acceleration



From Report for the Office of Scientific and Technical Information: Population Modeling of the Emergence and Development of Scientific Fields, Luis M. A. Bettencourt, Carlos Castillo-Chavez, David Kaiser, David E. Wojick, October 2006

Appendix 6

Science.gov: A Case Study in the "Art of the Possible"

OSTI combines passion, vision, and innovation to get things done, producing tangible and useful results more efficiently and effectively. This is OSTI exceptionalism. This is the art of the possible.

Science.gov, a search portal to more than 200 million pages of science information from 18 U.S. government science organizations within 14 federal agencies, was launched as the first-ever search capability across major U.S. science agencies. The path of Science.gov illustrates OSTI exceptionalism, in partnership with the interagency alliance. Science.gov:

- Grew from the April 2001 interagency workshop where the voluntary Science.gov Alliance was formed
- Was built from the ground up, at low cost, through innovation and collaboration across government
- Was launched as major new web portal in December 2002, hosted by OSTI
- Has grown steadily in content
- Represents 97 percent of the federal research and development *(R&D)* budget
- Improved search technology, with milestones in each new version
- Improved usability features
- Added handbook information *(Wikipedia science articles)* and EurekAlert! science news relevant to searches
- Made government resources act as if they were an integrated whole.

Version 1.0, December 2002: Provided for the first time wide public access and a unified search of the government's vast stores of scientific and technical information *(STI)*.

Version 2.0, May 2004: Introduced real-time relevancy ranking of government science information retrieval, helping citizens sort through the government's reservoirs of research and return results most likely to meet individual needs. An advanced search capability and other enhancements were added.

Version 3.0, November 2005: Took relevancy-ranked search to a higher level of precision, providing more refined search queries of federal science databases. Greatly enhanced fielded searching and the extensive Boolean capabilities offered new search options for Science.gov users. Alert Services were added.

Version 4.0, February 2007: Allowed patrons to refine search, view search results by title, author or date, as well as by relevancy rank or source, as in earlier versions, and email results. Relevancy ranking algorithms became more sophisticated.

Version 5.0, September 2008: Quadrupled content by adding seven new deep web sources and provided significant new features designed to improve users' search experience, including:

- ***Clustering***—Search results are automatically organized into subject and date categories allowing fine tuning of searches based on subtopics.

- ***Source totals***—Source summary information is provided for each search, letting the user know the sources that have provided search results and how many citations originated from each source.
- ***EurekAlert!***—Links to recent news articles and press releases that are relevant to a user's query terms are provided by the EurekAlert!, an online, global news service operated by the American Association for the Advancement of Science.
- ***Wikipedia articles***—Links to Wikipedia articles that are relevant to user search terms
- ***Limit results to a source***—Users can limit search results to those from a single source.
- ***Download citations***—Users can download citation information about selected documents to a file for importing into citation management software.

Science.gov will continue to evolve and provide an enriched search experience for anyone seeking reliable science information free of "web noise." OSTI continues to host the Science.gov website and work with the Science.gov Alliance to ensure access to reliable information. The voluntary interagency partnership includes the Departments of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services, the Interior, and Transportation, the Environmental Protection Agency, the National Aeronautics and Space Administration, the National Science Foundation, the Government Printing Office, the Library of Congress, and the National Archives and Records Administration.

Appendix 7

Precision Search

Precision search equals precise retrieval of the most relevant information

Information consumers are seeking accurate, timely and relevant information in response to their search queries. OSTI delivers by providing easy-to-use products with sophisticated relevance ranking and precision-searching capabilities.

In information retrieval, precision refers to the quality of the results. At OSTI, advanced, computer-based ranking algorithms are combined with user-friendly search interfaces to make searching efficient and productive. OSTI products allow fielded searching, by title, author, publication date, and other fields, so researchers can more quickly find the information they need. Results can be fine-tuned even further with numerous refine-and-sort options.

OSTI products are innovative and varied, and include focused literature collections such as the DOE Information Bridge, providing free access to full-text report literature, and DOepatents, a collection of patent information dating from the 1940s. In addition, the OSTI federated search applications, such as EnergyFiles and Science.gov, simultaneously query multiple, selected databases in search of relevant documents.

Precision Search: Who will benefit?

Examples include:

- The student who can complete assignments more thoroughly and follow his/her natural curiosity
- The teacher who is creating lesson plans in areas where he/she is not expert
- The researcher entering a new field
- The industrial entrepreneur who needs a solution fast to get the product to market

Federated searching is accomplished with powerful technology that queries individual, often disparate databases, and produces one set of aggregated results. As with all OSTI products, searches are performed in real time rather than relying on stored indexes with outdated information. Fielded searching within federated search applications was once impossible due to the variance among sources, but OSTI pioneered this advancement, and now OSTI federated search applications use a number of databases that support fielded searching. Regardless of source, the results are relevancy-ranked according to the query provided by the user.

OSTI products are designed with the searcher in mind, offering powerful and intelligible search tools. Precision search is easily accomplished in record time with minimal user effort. Students and teachers with deadlines, researchers looking for breakthroughs or entering new fields, industrial entrepreneurs moving products to market, or just those curious about science, will be rewarded at OSTI with fast, relevant and precise information.

References

Belliston, J.C., Howland, J.L., Roberts, B.C. Undergraduate use of federated searching: A survey of preferences and perceptions of value-added functionality, https://contentdm.lib.byu.edu/cdm4/item_viewer.php?CISOROOT=/HBLLRep&CISOPTR=21&CISOBOX=1&REC=2

Bergman, M.K. September 2001. The deep web: Surfacing hidden value. BrightPlanet, http://brightplanet.com/pdf/deepwebwhitepaper.pdf

Bettencourt, L.M.A., Cintron-Arias, A., Kaiser, D.I., Castillo-Chavez, C. May 2005. Power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models, www.osti.gov/innovation/research/diffusion/epimodeling/powerofagoodidea.pdf

Bettencourt, L.M.A., Castillo-Chavez, C., Kaiser, D., Wojick, D.E. Editor's Copy. Report for the Office of Scientific and Technical Information: Population modeling of the emergence and development of scientific fields, www.osti.gov/innovation/research/diffusion/epicasediscussion_lb2.pdf

Cintron-Arias, A., Bettencourt, L.M.A., Kaiser, D.I., Castillo-Chavez. May 2005. On the transmission dynamics of knowledge, www.osti.gov/innovation/research/diffusion/epimodeling/transknowledge.pdf

Kaiser, D. June 2006. Notes toward a nanotech timeline, www.osti.gov/innovation/research/diffusion/nanotechdiscussionDK.pdf

Keeling, M. March 2001. The mathematics of diseases. http://plus.maths.org/issue14/features/diseases/

Needham, J.L. December 2007. Testimony to the Senate Homeland Security and Governmental Affairs Committee, Google public sector content partnership.

Roosevelt, President F.D. November 1944. In his charge to Vannevar Bush, then director of the Office of Scientific Research and Development.

Workshop Panel Participants. April 2007. *Workshop Panel Report on Accelerating the Spread of Knowledge about Science and Technology: An Examination of the Needs and Opportunities*, www.osti.gov/publications/2007/workshop.pdf

2000 Workshop Report on a Future Information Infrastructure for the Physical Sciences. May 2000, www.osti.gov/physicalsciences/

Developing the UK's e-Infrastructure for Science and Innovation—Report of the OSI e-Infrastructure Working Group, www.nesc.ac.uk/documents/OSI/report.pdf

DOE Science Accelerator. June 2006. www.osti.gov/innovation/scienceaccelerator.pdf

DOE Strategic Plan. 2006. www.cfo.doe.gov/strategicplan/docs/2006StrategicPlan.pdf

Global Discovery: Increasing the pace of knowledge diffusion to increase the pace of science. February 2006. Slides from the American Association for the Advancement of Science annual meeting, St. Louis, Missouri, www.osti.gov/speeches/fy2006/aaas/

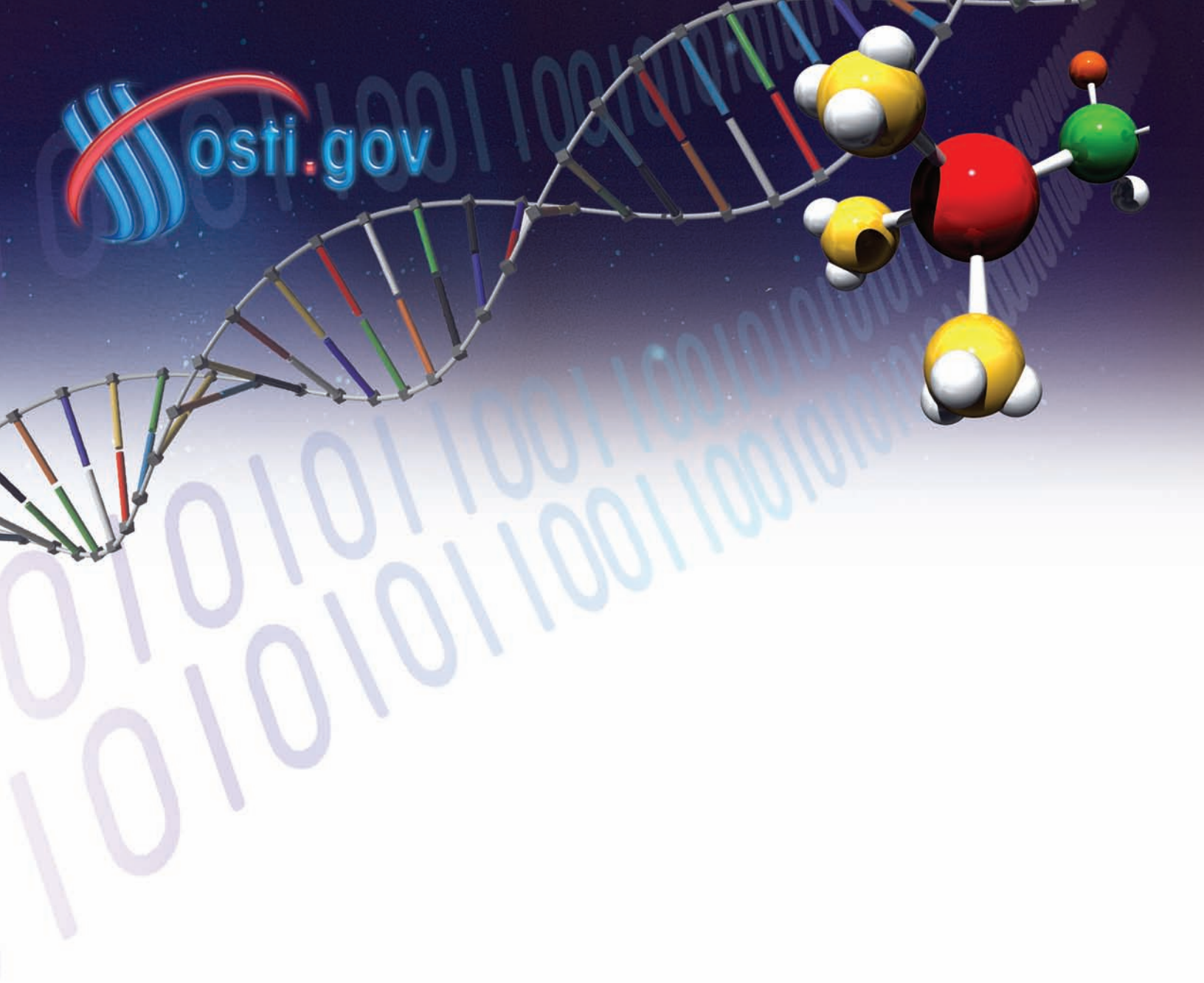
Milestones. OSTI's accomplishments for past decade. www.osti.gov/milestones

Office of Science Strategic Plan. February 2004. www.science.doe.gov/bes/archives/plans/SCSP_12FEB04.pdf

Perceptions of Library and Information Resources. 2005. Online Computer Library Center, www.oclc.org/reports/2005perceptions.htm

Small business innovation research grant opportunities: Discovery, search, and communication of textual knowledge resources in distributed systems. 2007. www.science.doe.gov/sbir/solicitations/FY%202007/41.OSTI.Search.htm

The State of Data Management in the DOE Research and Development Complex Report. November 2004. www.osti.gov/publications/2007/datameetingreport.pdf


osti.gov