

Gerry McKiernan

EnergyFiles: the virtual library of energy science and technology

The targeted audience for EnergyFiles is end users and organizations having a need for energy-related scientific and technical information. Anticipated primary users include researchers, scientists, project managers, program managers, academia (educators and students), and associated information professionals (United States Department of Energy, Office of Scientific and Technical Information, 1997, 1999).

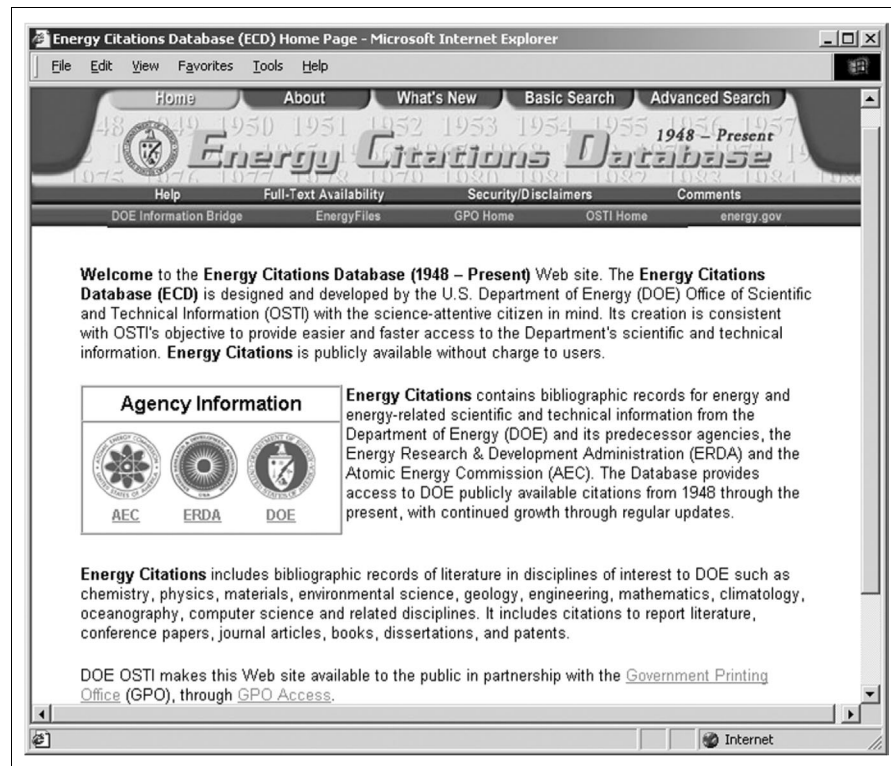
The Truth is Out There (*The X-Files*).

Energy Citations Database

The Energy Citations Database (ECD) (www.osti.gov/energycitations/) is a bibliographic database developed by the Office of Scientific and Technical Information (OSTI) of the US Department of Energy (DOE) to improve access to energy and energy-related scientific and technical information produced or funded by DOE and its predecessor agencies, the Atomic Energy Commission (AEC) and the Energy Research and Development Administration (ERDA) (see Figure 1). ECD is made available by OSTI in partnership with the US Government Printing Office (GPO) through GPO Access (www.access.gpo.gov/su_docs/index.html). It was launched in October 2001 and presently contains approximately 2 million records and is currently updated weekly.

The ECD provides full bibliographic records to the literature of disciplines of interest to the US Department of Energy, notably chemistry, physics, materials, environmental science, geology, engineering, mathematics, climatology, oceanography, computer science, and related disciplines. ECD provides citations to books, conference proceedings and papers, dissertations

Figure 1
Screen print of the Energy Citations Database homepage



and theses, government documents, journal articles, patents and patent applications, software manuals, as well as the report literature. Access to the full text of selected publications is also available (McKiernan, 2002).

Although specific in coverage and content, the Energy Citations Database in many respects represents the general nature and scope of the electronic resources made available through EnergyFiles: the Virtual Library of Energy Science and Technology (www.osti.gov/EnergyFiles/), the portal of which it is a component. Sponsored, developed, and maintained by OSTI, EnergyFiles “combines information, tools, and technologies to facilitate the use of scientific resources at the desktop” ... [that can] “... maximize access to and usefulness of the vast body of knowledge that exists in the

worldwide scientific community ...”. EnergyFiles is intended to offer “a fully searchable dynamic information resource that provides organized access to a comprehensive resource of scientific and technical information in physics and other disciplines of concern to the [US] Department of Energy.”

Collections and access

The EnergyFiles project was initiated five years ago in Spring 1997 with the appointment of a development team charged with defining “a virtual library within the context of the scientific and technical information (STI) created and maintained by the Department of Energy and building a prototype to demonstrate the concept.” The prototype was unveiled in May 1997 at InForum, the annual

information management and technology conference sponsored by OSTI; the current revision was implemented in May 1998.

EnergyFiles offers access to three major types of publications (see Figure 2):

- journal literature;
- gray literature; and
- electronic preprints (e-prints).

Journal literature

From within EnergyFiles access is provided to three bibliographic databases:

- (1) Infotrieve® Article Finder;
- (2) Scirus;
- (3) PubSCIENCE.

Infotrieve® Article Finder (www4.infotrieve.com/search/databases/newsearch.asp) provides access to a database of 20 million bibliographic citations with 10 million abstracts from more than 30,000 scientific, technical, medical journals, while Scirus (www.scirus.com/) provides access to Web resources, as well as free and

subscriber resources such as those available from or provided by ScienceDirect, IDEAL, MEDLINE on BioMedNet, Beilstein Abstracts on ChemWeb, Neuroscion, BioMed Central, United States Patent and Trademark Office, E-Print ArXiv, Chemistry Preprint Server, Mathematics Preprint Server, CogPrints and the National Aeronautics and Space Administration (NASA). Collectively, Scirus provides access or links to more than 90 million Web pages, and millions of bibliographic citations, abstracts, and full-text publications or e-prints (Elsevier Science, 2002).

PubSCIENCE (pubsci.osti.gov) is a free gateway search engine that provides access to the citations, and, in select cases, to the full text of journal articles provided by cooperating publishers. PubSCIENCE citations are derived from two sources: participating publishers and information intermediaries maintaining citation collections based on agreements negotiated with OSTI, and the DOE Energy Science and Technology Database journal citations maintained by OSTI. To date, more than 1,000

separate journal titles have been integrated into PubSCIENCE.

Recently, the Department of Energy has proposed that PubSCIENCE be discontinued, due in part to the overlap in coverage of the resources made available by Infotrieve® Article Finder and Scirus.

Gray literature

Gray literature may be defined as “literature which is not readily available through normal book-selling channels, and therefore difficult to identify and obtain.” “Examples of grey literature include reports, technical notes and specifications, conference proceedings and preprints, supplementary publications and data compilations, trade literature and so on” (Auger, 1998, p. 3). Among the more significant types of gray literature is the technical report, a specialized publication generally issued as a formal report of a completed organizational, institutional, or research project (Auger, 1998, pp. 8-15).

Access to the technical reports produced by the national laboratories of the Department of Energy, and its grantees, is provided from within EnergyFiles via the DOE Information Bridge. The DOE Information Bridge offers access to the full text of more than 60,000 searchable reports issued since 1995.

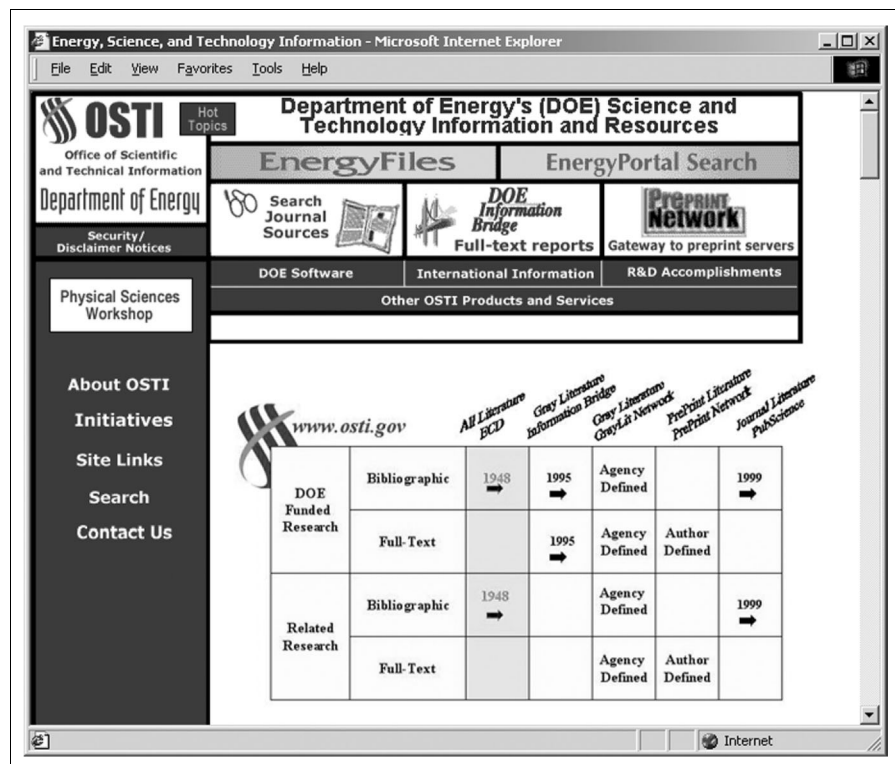
In addition, EnergyFiles provides access to the GrayLIT Network (www.osti.gov/graylit/), an OSTI information service that offers access to nearly 120,000 full-text technical reports from a wide range of Federal agencies.

From within GrayLIT Network, users can individually or simultaneously search any or all of the following databases, notably:

- Defense Technical Information Center (DTIC) Report Collection;
- DOE Information Bride Report Collection;
- NASA Jet Propulsion Laboratory (JPL) Technical Reports;
- NASA Langley Technical Reports;
- EPA National Environmental Publications Internet Site (NEPIS).

The Defense Technical Information Center (DTIC) Report Collection

Figure 2
Screen print matrix of selected EnergyFiles collection gateways



(stinet.dtic.mil) is a full-text collection of more than 61,500 unclassified documents dating from 1974. The DOE Information Bridge Report Collection (www.osti.gov/bridge/) provides access to the full text of more than 65,000 reports with a total of more than 4.3 million pages. More than 12,500 technical reports are available from the NASA Jet Propulsion Laboratory (JPL) Technical Reports collection (jpltrs.jpl.nasa.gov) in Portable Document Format (PDF). The NASA Langley Technical Reports collection (techreports.larc.nasa.gov/ltrs/ltrs.html) provides access to nearly 3,000 publicly available NASA documents covering aeronautical and space science in PDF, HTML, or compressed PostScript format, with searchable citations and abstracts.

National Environmental Publications Internet Site (NEPIS) (www.epa.gov/cgi-bin/claritgw?op=NewSearch&template=epa) provides access to the full text of more than 7,000 EPA documents in the areas of water quality, wastewater, pesticides, ecological issues, wetlands, and other environment-related topics.

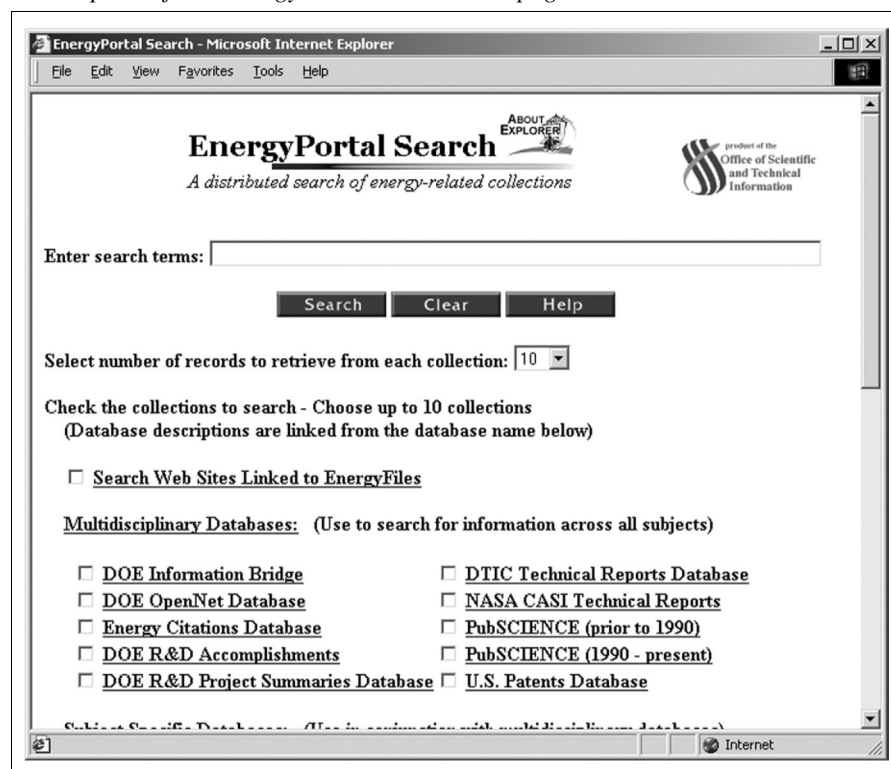
E-prints

In general, “preprints are manuscripts that have not yet been published, but may have been reviewed and accepted; submitted for publication; or intended for publication and being circulated for comment.” Electronic versions of preprints are commonly referred to as “e-prints.” Worldwide, it is estimated that there are more than 7,600 preprint servers. Among the largest and well-known e-print servers is arXiv.org, the service originally based and developed by Paul Ginsparg at the Los Alamos National Laboratory in New Mexico (McKiernan, 2000).

The PrePRINT Network is a searchable gateway to preprint servers that cover the scientific and technical disciplines of interest to DOE (see Figure 3). Such disciplines include physics, materials, and chemistry, as well as biology, the environmental sciences and nuclear medicine. Through its search feature, the PrePRINT Network allows users to launch parallel searches on any or all of the incorporated preprint servers. PrePRINT Alerts allows users to

Figure 3

Screen print of the EnergyPortal Search homepage



subscribe preprint servers to receive weekly e-mail updates. The PrePRINT Network can be searched directly, or concurrently with other networks, databases, and resources incorporated within EnergyFiles.

Other components

EnergyFiles also includes several other complementary databases and DOE compilations, notably:

- Energy Citations Database;
- DOE R&D Accomplishments;
- DOE R&D Project Summaries;
- Federal Research and Development Project Summaries;
- Federal Research and Development Project Summaries.

As previously noted, the Energy Citations Database (www.osti.gov/energycitations/) provides bibliographic records, as well as abstracts, to a full range of energy and energy-related scientific and technical information produced by the Department of Energy (DOE) and its predecessor agencies. DOE R&D Accomplishments (www.osti.gov/accomplishments/) is a database that offers access to

descriptions of previous DOE research and development projects that “have had significant economic impact, have improved people’s lives, or have been widely recognized as a remarkable advance in science.” Current DOE research and development projects are available from the DOE R&D Project Summaries database (www.osti.gov/rnd/), while those of all Federal agencies are available from Federal Research and Development Project Summaries (www.osti.gov/fedrnd/).

Subject Pathways

Within EnergyFiles, select collections are organized into subject-specific categories called Subject Pathways, which offer “one-stop access to and use of widely distributed information” and provide “the foundation for an electronic national library of energy-related science and technology.” The categories shown in Table I are included.

For each collection a general description is provided along with a listing and access to associated electronic resources incorporated within EnergyFiles, for example.

Table I*Categories included in the Subject Pathways*

Biology and medicine	Geosciences
Chemistry	Materials science
Energy storage, conversion, and utilization	Mathematics, computing, and information science
Engineering	Physics
Environmental sciences, safety and health	Plasma physics and fusion
Fission and nuclear technologies	Power transmission, distribution and plants
Fossil fuels	Renewable energy

EnergyFiles: materials science

The materials science subject area provides information to support materials research directly associated with energy production, conversion, or utilization or with studying radiation effects on materials. Energy-related applications of metals, alloys, ceramics, cermets, refractories, composites, polymers, plastics and other materials are (also) covered:

- (1) EnergyPortal Search.
- (2) Resource Index:

- STI collections and databases;
- Electronic journals, preprints, and conference proceedings;
- Related resources;
- Virtual workspace.

For listed items in the Subject Pathways, a detailed list with annotations and links is provided for each; for example:

- (1) *STI collections and databases:*

- DOE Energy Citations Database (DOE ECD), DOE and DOE Contractor Web site – Bibliographic records for energy and energy-related scientific and technical information from the Department of Energy (DOE) and its predecessor agencies, the Atomic Energy Commission (AEC) and the Energy Research & Development Administration (ERDA), including records (domestic and international) from the Nuclear Science Abstracts (NSA) and from the

Energy Science and Technology Database (EDB). DOE ECD provides access to bibliographic records from 1948 forward.

- DOE Energy Citations Database (ECD), Public Web site – Bibliographic records for energy and energy-related scientific and technical information from the Department of Energy (DOE) and its predecessor agencies, the Energy Research & Development Administration (ERDA) and the Atomic Energy Commission (AEC). Energy Citations Database provides access to DOE publicly available citations from 1948 forward.
- DOE Information Bridge, Public Web site – (no registration or password required) Searchable and downloadable bibliographic records and full text of DOE research report literature from 1995 forward.
- DOE Information Bridge, DOE and DOE Contractor Web site – (Registration and password required) Searchable and downloadable bibliographic records of worldwide energy research and full text of DOE sponsored or acquired scientific and technical information from 1995 forward.
- DOE R&D Accomplishments Database – Searchable and downloadable bibliographic records and full-text reports of past DOE R&D accomplishments.
- DOE Research and Development (R&D) Project Summa-

ries – Subset of the research projects funded by the Department of Energy in fiscal years 1995 through 1999.

- Making Matter: The Atomic Structure of Materials – The Institut Laue Langevin provides this site, which illustrates the 3D structure of inorganic materials including perovskites, zeolites, and gems and minerals.
- NIST Property Data Summaries for Advanced Materials – Property Data Summaries are collections of property values derived from surveys of published data including thermal, mechanical, structural, and chemical properties.
- Preprint Network – Searchable gateway to preprint servers that deal with scientific and technical disciplines of concern to DOE such as physics, materials, and chemistry, as well as portions of biology, environmental sciences, and nuclear medicine.
- WebHTS – The NIST WWW Version of the High Temperature Superconducting Materials Database.
- PubSCIENCE – Provided by DOE's Office of Scientific and Technical Information and the US Government Printing Office, PubSCIENCE offers users the capability to search across a large compendium of peer-reviewed journal literature with a focus on the physical sciences.

- (2) *Other materials science resources.* Electronic Journals, Preprints and Conference Proceedings | Related Resources.

Resource List

The Resource List provides “an enhanced alphabetical listing of all resources available through EnergyFiles . . .” which can be browsed or searched; access to a categorized collection is available through the Subject Pathways gateway.

A full range of information resource types is incorporated in the Resource List and includes publisher Web sites,

conference Web sites, Web resource collections, newsletters, e-journals, directories, federal agencies' Web sites, information services, catalogs, research centers' Web sites, databases, and technical report collections, among many.

EnergyPortal Search

Within EnergyFiles, users may search all incorporated resources using the EnergyPortal Search. The EnergyPortal Search provides access to a variety of multidisciplinary databases, notably the DOE Information Bridge, DOE OpenNet Database, Energy Citations Database, DTIC Technical Reports Database, NASA CASI Technical Reports, and PubSCIENCE among others (see Figure 3).

The DOE OpenNet Database is a component of the DOE "Openness Initiative" for "public awareness, public education, and public input" and "designed to provide easy, timely access to recently declassified information, including information declassified in response to Freedom of Information Act requests." The NASA CASI Technical Reports database provides access to the NASA Center of AeroSpace Information RECONplus database that "contains worldwide information on aeronautics, astronautics, chemistry and materials, engineering, geosciences, life sciences, mathematical and computer sciences, physics, social sciences, and space sciences published from 1915 to the present."

In addition, access is provided to a range of subject-specific collections in biology and medicine, engineering, environmental sciences, physics, renewable energy. Of particular note are the databases accessible from within the environmental and renewal resources databases, notably:

- The Atmospheric Radiation Measurement Program (ARM) (www.arm.gov) site which focuses on global climate change research with coverage of "sunlight, radiant energy, and clouds on temperatures, climate, and weather."
- The Carbon Dioxide Information Analysis Center (CDIAC) (cdiac.ornl.gov/about/intro.html), which provides access to information

about the greenhouse effect and global climate change.

- Environmental Management Science Research Projects (EMSP) (emsp.em.doe.gov/portfolio/Multi-Search.asp), a searchable database "of research projects funded by the US Department of Energy, Office of Environmental Management Science Program, [with] ... emphasis in such research areas as chemistry, engineering, geochemistry and geophysics, health science, materials, microbial science, plant science, and related topics."
- Alternative Fuels Data Center (www.afdc.nrel.gov), which provides access to information sources on alternative fuels provided by the National Renewable Energy Laboratory.
- Bioenergy Information Network (bioenergy.ornl.gov), a bibliographic database of DOE-funded projects focusing on production, recovery, and conversion of energy crops to fuels and power generation, are notable resources in the Renewable Energy category.

Other options

Within the EnergyPortal, users can specify that the HML and PDF collections linked to EnergyFiles are also searched ("Search Web Sites Linked to EnergyFiles") (see Figure 3). In addition, the number of records to be retrieved from selected collections can be designated from a pull-down menu (i.e. 10, 25, 100, 150, 200, 250); a maximum of ten collections may be chosen.

Collection Development Criteria

To guide the development of EnergyFiles, a document that defines its purpose and scope has been prepared. This Collection Development Criteria document "states the principles and guidelines that address the selection and inclusion of information collections, databases, materials, software tools and other resources for the Department of Energy's EnergyFiles Virtual Library Environment." As such, it serves as "a planning document for future development and growth of

EnergyFiles," with the criteria to be revised and updated as necessary.

Key sections of the document include: Mission/objectives of EnergyFiles; Definition of scientific and technical information (STI); Audience; and Selection guidelines. The following are essential components of its selection guidelines:

- All resources, including reference materials and software tools, must support research in one of the energy-related disciplinary areas defined in this collection criteria document and must demonstrate value to the primary user community.
- Full text data, or complete numeric or other data where applicable, is preferred where available. If the information is not available in full text, then alternate forms will be considered.
- Resources and software tools must be maintained on a regular basis by the source or originating site and sites must show a demonstrated commitment to making the information or tools available long term.
- A willingness on the part of the information owners to participate in distributed searching and other collaborative efforts, where applicable, is desired.
- Resources and tools in English are preferred; information in other languages will be considered if a demonstrated need is evident.
- Emphasis is on current and recent information; however, legacy resources will be added as needed to meet the needs of the user community.
- Accuracy of the information as well as the quality of the writing and presentation format will be considered.

Within the collection development policy, specific subject areas relevant to the scope of EnergyFiles are delineated, and include the following:

- biomedical sciences;
- chemistry;
- energy storage, conversion, utilization and policy;
- engineering;

- environmental sciences and safety and health;
- fission and nuclear technologies;
- fossil fuels;
- geosciences;
- materials science;
- mathematics, computing and information science;
- physics;
- plasma physics and fusion;
- power transmission, distribution and plants;
- renewable energy.

In addition, each subject area is annotated with details of its specific coverage, including, for example:

- *Energy storage, conversion, utilization and policy.* This area contains information resources to support research on equipment and methods to reduce energy consumption, increase energy efficiency, and enable the substitution of more plentiful or environmentally favorable energy sources. Examples of research supported include mechanical, chemical, electromagnetic and thermal methods for storing energy; methods of converting heat into direct electrical energy; design and development of advanced propulsion systems for land, sea and air vehicles; and energy conservation within buildings, transportation, industry and agriculture, and within municipalities and communities.
- *Fission and nuclear technologies.* The Fission and Nuclear Technologies section is designed to provide information useful to researchers in the areas of nuclear fuels, nuclear power and reactors, and isotope and radiation source technology. Within the area of nuclear fuels, information to support research on mining, uranium enrichment, spent fuels, waste management, and safeguards and accountability is included. Research in the isotope and radiation source technologies includes isotope separation, radiation sources and isotopic power supplies. Various aspects of nuclear power reactors are covered, including the specific types of reactors, and associated technologies such as fuel elements, control systems and reactor safety.

- *Power transmission, distribution and plants.* Information supporting research and development in utility-size fossil-fuel power plants and the transmission and distribution of power is provided in this section. Research into new designs, developments and technologies for both the generation of power at the plants as well as power transmission and delivery are included.

Virtual workspace

In addition to its various gateways, resource list, and collections, EnergyFiles includes a Virtual Workspace, an “interactive work environment containing tools and technologies that facilitate collaboration and peer interaction, remote experimentation, [and] discovery . . .” While a prototype, the workspace is envisioned as an effort to “interactively link library with laboratory.” Among the tools and technologies currently available are:

- ACTS Toolkit (www.nersc.gov/ACTS/). The ACTS (Advanced Computational Testing and Simulation) Toolkit is a set of DOE-developed software tools that facilitate the preparation of computer programs for high performance scientific applications for parallel computers.
- Ask a Scientist Service (newton.dep.anl.gov/#AAS). The NEWTON BBS service, operated by the Division of Educational Programs (DEP) of Argonne National Laboratory, Illinois, provides K-12 science, math and computer teachers and their students with “a place to practice telecommunications, to retrieve useful information in a wide variety of subjects, to contact research scientists from all over the world and to open communications between classroom teachers.”
- Ask an Energy Expert (www.eren.doe.gov/askanenergyexpert/). The Ask an Energy Expert service, provided by the DOE Office of Energy Efficiency and Renewable Energy, offers access to information fact sheets, relevant Web sites, and an opportunity to ask a

question related to energy efficiency and renewable energy.

- Software for Environmental Awareness (www.epa.gov/glnpo/seahome/). This EPA site provides access to software and programs on such topics as water, solid waste, air pollution, and environmental assessment.
- Technical Tools and Models (www.epa.gov/epahome/data/tool.htm). An EPA site that offers access to specialized databases and computer-based tools to assist in modeling environmental conditions and assessing water quality, among other activities.

Privacy, security, and accessibility

EnergyFiles provides explicit details about its policies and practices relating to privacy, security, and accessibility.

Privacy

Within EnergyFiles, selected site-access information is automatically captured and stored, and includes the:

- Internet Protocol (IP) address of the domain from which EnergyFiles is accessed;
- type of browser and operating system used;
- pages accessed; and
- referring Web site (if accessed from another Web site).

“Site-access information is aggregated and used to assess the value of the Web site in accomplishing its stated information dissemination goals, to plan for enhancements to make this site more useful to visitors, to optimize the site’s technical design specifications, and/or to identify system performance or problem areas. This information is used only as a source of anonymous statistical data, and is shared only when required by law enforcement investigation.” It is not used to “track or record information about individuals and their visits.”

Personal information is not collected unless the user voluntarily provides it through e-mail queries or requests. “Except for authorized law enforcement investigations, these e-mails and addresses are shared only to

the extent necessary to assure that an accurate and timely response is provided.”

EnergyFiles site does not use persistent cookies. “Session cookies may be used to provide salient session-related information and to allow product-specific functions to be performed, but in no case are cookies to be used to capture information that might compromise or threaten personal privacy.”

Security

EnergyFiles is “part of a Federal computer system used to accomplish Federal functions. Computer software programs as well as other surveillance methods are used to monitor network traffic on this Web site for security purposes. By accessing this Web site, . . . [the user] is expressly consenting to these monitoring activities.”

Accessibility

EnergyFiles complies with Section 508 of the Rehabilitation Act (www.section508.gov/), which “requires that Federal agencies’ electronic and information technology be made accessible to people with disabilities, including Federal employees and

members of the public.” The Bobby utility for evaluating Web sites for accessibility (bobby.watchfire.com/bobby/html/en/index.jsp) is used to ensure compliance.

Disclaimers

EnergyFiles “is made available by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of originators expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.”

REFERENCES

Auger, C.P. (1998), *Information Sources in Grey Literature*, 4th ed., Bowker Saur, London.

Elsevier Science (2002), “About Scirus”, available at: www.scirus.com/about/ (accessed 27 August 2002).

McKiernan, G. (2000), “arXiv.org: the Los Alamos National laboratory e-print server”, *The International Journal on Grey Literature*, Vol. 1 No. 3, pp. 127-38.

McKiernan, G. (2002), “Energy Citations Database”, *The Charleston Advisor: Critical Reviews of Web Products for Information Professionals*, Vol. 4 No. 1, pp. 25-31, 34-5.

United States Department of Energy. Office of Scientific and Technical Information (1997), *EnergyFiles: Collection Development Criteria*, DRAFT, Washington, DC.

United States Department of Energy. Office of Scientific and Technical Information (1999), *EnergyFiles: Collection Development Criteria*, DRAFT, Washington, DC.

Gerry McKiernan (gerrymck@iastate.edu) is a Science and Technology Librarian and Bibliographer, Iowa State University Library, Ames, Iowa, USA.