

Other UK websites

[UK Subscriptions](#) [Overseas Subscriptions](#)

[Blog](#) | [News](#) | [Jobs](#)

[Science](#) | [Business & Market](#) | [Legal](#) | [Medical](#) | [Academic & Humanities](#) | [News & Reference](#) | [Information Management](#) | [Library](#)



[RSS](#) [All articles](#) [Site map >](#)

[Information World Review](#) > [Features](#) > [Library issues](#)



Don't let the grey fade away

Online efforts underway to open the door to a wealth of quality information that would otherwise languish on shelves or disappear forever

By **Kim Thomas** 07 May 2008

Despite its distinctly unglamorous name, "grey" literature – that mass of booklets, research reports and PDF documents produced in-house by organisations rather than through a recognised publisher – is a hot topic in the library world.

While the question of how to find grey literature and make it available has always been a challenge for librarians, the advent of web-based tools is starting to make the job easier. At the same time, the web has enabled the production and publication of a much greater body of grey literature. But unless this literature is archived, it could disappear from public view altogether.

Jennie Grimshaw, lead content specialist for social policy and official publications at the British Library, says: "Grey literature is enormously valuable to users because it is often more informative than articles in peer-reviewed journals.

"You'll get the full methodology, which is often omitted from peer-reviewed journals, because authors are often asked to cut an article down to 10 pages.

"The research report will give you more information about the data set, and more information about the research that was unsuccessful, because there's a tendency for people to publish formally only the interventions that were successful, not the ones that didn't work."

Say it again, and again...

These reports provide an opportunity to see the results of a research project long before they appear in published form, adds Grimshaw. "There's often a long delay in something coming out in a peer-reviewed journal, whereas a research report is available much more quickly," she says. "And academics are under enormous pressure to publish, so they have a horrible tendency to salami-slice, and they'll produce endless articles saying more or less the same thing."

And it's not just research reports. Government departments, quangos and thinktanks all publish booklets and reports that are in danger of getting lost. Then there are the postgraduate students producing theses and dissertations that, although indexed, spend much of their lives gathering dust.

But that has begun to change, thanks mainly to the internet. "Before the internet, you had to rely on a smattering of incomplete sources, hit-and-miss journal and monograph sources and dissertation indices," says John Hagen, manager of institutional repository programs and co-ordinator of the electronic thesis and dissertation programme at the University of West Virginia. "So much of the body of grey literature remained hidden, languishing on library shelves."

Co-ordinated efforts are being made to preserve grey literature in electronic form and make it accessible. One of the earliest UK projects to address the issue was run by Cranfield University. The project, which concluded in 2002, looked at grey literature in the engineering sector, and found a body of largely uncatalogued documentation, including technical reports from the Aeronautics Research Council (ARC) that still had immense value.

News centre

- [Science](#)
- [Business & Market](#)
- [Legal](#)
- [Medical](#)
- [Academic & Humanities](#)
- [News & Reference](#)
- [Information Management Technology](#)
- [News & Reference](#)
- [Information Management](#)
- [Library](#)
- [XML](#) RSS feeds

KnowledgeBANK

- [SilkCentral Issue Manager - Defect Tracking Tool](#)
- [EnCase® eDiscovery Suite](#)
- [bf.documents -- Streamline Collaborative Document Creation and Management](#)

More from Information World Review

- [NHS trials remote monitoring of patients](#)
- [Information Commissioner gets stronger powers](#)
- [MySpace offers opt-in data sharing](#)

ADVERTISEMENT



Paul Needham, who worked on the project and is now a research and innovation specialist at Cranfield University, says: "Some materials would be regarded as seminal: a lot of the reports were to do with basic aerodynamic experiments. People still wanted access to these reports as part of their current research rather than having to do the investigations and experiments."

The project resulted in a prototype national reports catalogue. "We created a large database that was populated from library catalogues and harvesting," says Needham. "We demonstrated how it was possible to link together reports, the corporate sources of those reports, descriptions of series of reports, and collections where those reports were held."

Needham believes the project was ahead of its time. Now there are numerous initiatives to create collections of grey literature and publish them on the web.

e-theses

One of the biggest success stories has been electronic theses and dissertations. The pioneer in the field is Virginia Tech University, which began making theses and dissertations available electronically in 1995.

Many higher education institutions around the world have followed suit by making the electronic deposit of PhD theses mandatory.

While students have, on the whole, been happy to deposit their theses in an electronic repository, says Hagen, there have been some issues.

"We've realised the political difficulties, with the tradition of publishers being reticent to publish material that had already been distributed on the web," he says, "so it's taken some time for us to work with faculty, graduate students and publishers to find common ground where everyone can be comfortable in providing open access to their research."

One of the key factors in persuading people, Hagen says, has been Stevan Harnad's research, which showed that papers made available under an open access model are cited more frequently (sometimes as much as five times more frequently) than those that are not.

To make it easier to search across institutional repositories, the Networked Digital Library of Theses and Dissertations (NDLTD) has initiated a project to harvest metadata from university electronic theses and dissertations.

Hagen believes such initiatives have revolutionised scholarship. "Universities can now immediately share their intellectual property output with the world," he says. "Electronic theses and dissertations also assist developing regions by providing unfettered access to cutting-edge research, and they allow for low-cost distribution and access to their research as well. In many ways they are the great equaliser."

Many UK universities now have institutional repositories that hold electronic theses and dissertations. But it can still be difficult to find theses on a particular topic if it means carrying out individual searches on each institutional repository.

The Electronic Theses Online Service (ETHOS) project, a partnership between the British Library and institutions of higher education, is creating a central service that, from September, will enable users to access electronic theses and dissertations held by UK institutions. Institutions can deposit them at the central portal run by the British Library; if they choose not to, users searching for a particular thesis will be redirected from the portal to the relevant institutional repository.

Neil Jacobs, programme manager at the Joint Information Systems Committee (JISC), which is part-funding the project, says: "Theses are digitised on demand, so if someone comes to the service and asks for a thesis from 1861, the service will go back to the relevant institutions to ask for the thesis so it can be digitised."

Other European countries are also making theses and dissertations available electronically. The French national catalogue, Sudoc, lets users search academic repositories throughout the country. Other countries, such as Sweden, Germany and Holland, have similar catalogues, says Christiane Stock, head of monographs and grey literature at France's Institute for Scientific and Technical Information. Most e-repositories use the open access initiative protocol for metadata harvesting (OAIPMH), which allows search engines and catalogues to pick up their metadata.

Keeping secrets

Some authors are concerned about plagiarism, although Hagen says widely available plagiarism detection software can solve the problem. Another concern, says Stock, is that theses may include confidential information. In these cases, it is usual to remove the confidential data.

Copyright issues can also arise from the use of illustrations, Stock points out. "Students often scan large parts of works, and university libraries fear that they might have legal problems if those scanned illustrations are in the thesis, so they might not make it publicly available."

Theses and dissertations are only a small part of the grey literature world. How are users to find technical and research reports that may contain valuable information but are hidden on scientific society websites or in institutional repositories?

Some help is available through specialist search engines, one of the most well-known of which is Elsevier's **Scirus** tool. Scirus indexes a mass of web-based scientific information: journal content, scientists' homepages, courseware, pre-print server material, patents and institutional repository and website information.

No equivalent search engines in other disciplines can match Scirus's scope – a claimed 450 million science-specific web pages – although **Google Scholar** searches scholarly articles in all disciplines, including those held in institutional repositories or by scholarly organisations, as well as those published in journals.

The excellent **OpenDoar** provides a comprehensive directory of academic open access repositories, and has launched a trial search service for the full text of material held in open access repositories listed in the directory.

OAIster is another very useful service that catalogues thousands of digital resources, including e-books and articles, audio files and data sets, from more than 900 contributors.

But a very wide body of grey literature produced by organisations remains difficult for librarians to find and users to access. Grimshaw says many organisations are unaware of their legal obligations to deposit published reports with the British Library, and she spends much of her time writing to university departments, research institutes and government agencies asking for copies of publications. And while the internet has made it easier to make such reports available to the public, it has also created extra difficulties, says Grimshaw, because there is no legal requirement to deposit electronic publications with the British Library.

An increasing number of organisations, including the Institute for Public Policy Research (IPPR) publish online only, so there is a danger of some publications being lost forever.

Regulations to mandate the deposit of material published electronically are in hand but are unlikely to be implemented before autumn 2009. Currently the Legal Deposit Advisory Panel is looking at the issue of legal deposit as it relates to a variety of media, such as scholarly journals, handheld electronic media (CDs, for example) and material on websites.

Once in place, Grimshaw hopes the regulations will enable the British Library to harvest websites for reports and documents. Although the library already has a web archiving programme, it has to ask permission from the rights holders to archive documents published on those websites.

One major producer of grey literature is government itself.

"Government commissions vast quantities of social science research at great public expense on the whole range of social problems and science and technology," says Grimshaw. Government departments, she adds, increasingly publish their documents only on the web, and then remove those documents when they update their websites. Much valuable material is lost this way.

Grey literature is not, of course, confined to reports and academic dissertations. Two innovative projects, both part-funded by JISC, take a very broad view of what constitutes grey literature.

Talking heads archive

At Manchester University a team is creating a digital repository of Access Grid videoconferencing events, including seminars, conferences and workshops.

And at King's College London, a project is under way to create an electronic

repository of documents relating to committee meetings, such as minutes of those meetings and papers presented at meetings. The job of finding, cataloguing and indexing grey literature is herculean. Before the internet, it was, as Hagen says, "mission impossible". Even now it is a job that can never be completed.

But thanks to the development of web-based resources, information professionals can begin to provide their users with access to a treasure trove of grey literature that would previously have been out of reach.

Survival story

The problem of grey literature preservation is being addressed by the National Archives, which has launched a web continuity project.

From November this year, the National Archives will harvest all government websites three times a year and create a digital store of government documents. If a user finds a document through Google that is no longer on the website, they will be redirected automatically to the archive.

"It's a desire to ensure that all publications made available through government websites will survive for researchers in the future," says David Thomas, chief information officer at the National Archives, who points out that, before 1996, websites were not archived, and almost none survive from that period.

All Library issues

Like this story? Spread the news by clicking below:

 [del.icio.us](#)  [Digg this](#)  [reddit!](#)

Permalink for this story

Useful links: [About us](#) . [Privacy policy](#) . [Terms & conditions](#) . [Advertise](#)

[UK Subscriptions](#) | [Overseas Subscriptions](#)

© Incisive Media Ltd. 2008

Incisive Media Limited, Haymarket House, 28-29 Haymarket, London SW1Y 4RX, is a company registered in the United Kingdom with company registration number 04038503