

# ELSEVIER DIGITISES SCIENTIFIC HERITAGE

*Karen Hunter*

*The project to digitise all back issues of Elsevier Science journals and make these electronic backfiles available to the community is described. More than 40 million pages of text have been digitised; the print issues involved occupied 3.8 km of shelf space and filled six containers.*

*Karen Hunter, Senior Vice-President, Elsevier Science, 655 Avenue of the Americas, New York, NY 10010, USA  
Tel: +1 212 633-3787  
Fax: +1 212 633-3764  
E-mail: k.hunter@elsevier.com*

## **Introduction**

With researchers now using ScienceDirect, the online full-text database from Elsevier Science, many are requesting comprehensive retrospective access to a whole range of scientific journals. In response, Elsevier Science recently decided to expand electronic access on ScienceDirect to include all back issues for its entire publication list. This unique resource enables researchers to access past issues of Elsevier Science journals electronically. In many cases, first issues will be accessible – for example the first issue of the *Lancet* which was published in 1823. The backfiles complement the current volumes on ScienceDirect, which start with the mid-1990s.

## **Scanning six containers**

The backfiles project involves the conversion of more than 40 million pages of journals into electronic format. Elsevier Science is investing \$40 million in scanning and text tagging to make this vast body of articles available electronically. Within Elsevier Science alone, at least ten people are already working full-time on the backfiles project. The project, which began in December 2000, is expected to be complete by the end of 2003. Responsibility for producing the electronic backfiles lies with the Philippine company SPI Technologies Inc., an IT services company. Content Sciences Inc. (CSI), a subsidiary of SPI, currently operates with 1053 skilled Filipino IT workers, who are fully trained for, and dedicated to, the backfiles project. Elsevier Science ultimately expects to have 1200-1300 people from Content Sciences working for them. Twelve months ago, six containers full of journals from the out-of-print archive were shipped to the Philippines. The

journals – on a mixture of paper and microfiche – weighed more than 60 tonnes.

### Special treatment for fragile issues

The process of obtaining all volumes of the journals is a complex and fascinating one. Some date back to the nineteenth century. Elsevier Science has begun an extensive search for missing volumes and journals, to ensure that as much of the content as possible is made available. The extensive search goes from libraries' cellars, basements and offices to editors. Elsevier Science also has a project with the National Library of the Netherlands to help source and scan the volumes and issues that are not found this way. Very old and fragile publications, such as the first issue of the *Lancet* of 1823, would not withstand transportation to the Philippines and require special treatment to ensure they are not destroyed in the process.

### 3.8 kilometres in digital format

ScienceDirect customers already accessing the collections are finding that their high level of usage of the articles more than justifies the investment. The unit cost of supplying the articles electronically drops to a level much lower than an average inter-library loan or document delivery charge. Libraries could of course scan the material themselves, but at a significantly higher cost. ScienceDirect also has the benefit of providing excellent functionality. There are also cost saving benefits in terms of shelf space – the whole collection fills more than 3 kilometres of shelving!

### User convenience

The journal backfiles are maintained by and accessed through ScienceDirect, enabling users to utilise ScienceDirect's advanced search capabilities. All the backfiles' metadata, as well as abstracts, are fully searchable by every ScienceDirect user. Full-text access is available as a separate one-time purchase for current ScienceDirect and ScienceDirect OnSite subscribers. The text is viewable in PDF format, with all metadata (authors, titles, etc.), abstracts and references available in HTML. The body of

the text is fully searchable via an ASCII file in the background. All the articles in the backfiles collections will be registered with CrossRef, so that researchers can link to the abstract or full text of the referenced article, depending on what is electronically available and the researchers' entitlements (determined by the journals' publishers). This makes it easy to navigate between current research and historical perspectives. Electronic access to the full text of each backfiles subject is available under a single purchase payment – a one-off fee for unlimited usage. This payment allows customers to benefit from future version upgrades, as long as they remain ScienceDirect subscribers. Customers also have the option of choosing a three-year payment plan.

Customers purchasing backfiles collections have full archival rights to those collections, as defined in the ScienceDirect licence. So long as the library is a ScienceDirect customer, they have access via the online platform. If they are not a ScienceDirect subscriber, purchase (at the cost of reproduction) of a copy of the files is a possibility.

### Enthusiastic feedback

Full-text access will be made available in stages, each covering a particular subject area. Elsevier Science began by concentrating first on those disciplines for which archived literature remains most relevant to contemporary research and analysis. Subject areas already available are organic chemistry (which includes eight organic chemistry journals); inorganic chemistry; and chemical engineering. Coverage starts with volume 1, number 1, for titles in this field, with the longest-published journal (*Tetrahedron*) dating back to 1957. All journals on the subjects of chemistry, mathematics and economics are already available. The completed sections are already in use and Elsevier Science has received enthusiastic feedback from the research community. Other subjects soon to be released in 2002 include:

- business management and accounting
- psychology
- social sciences
- physics and astronomy
- pharmacology, toxicology and pharmaceuticals

- immunology and microbiology
- biochemistry, genetics and molecular biology

The following subjects are also scheduled to be worked on:

- engineering and technology
- neuroscience
- earth and planetary science
- medicine
- materials science
- decision science
- computer science
- environmental sciences
- energy and power
- agricultural and biological sciences

### Copyright issues

For the last 25 years Elsevier Science has obtained a written transfer of copyright for the articles published. While there may be an occasional article that is published under licence rather than copyright transfer, that licence also provides the rights needed to do the backfiles project. Prior to this time the transfer was more informal, but it was stated in instructions to authors that by submitting an article and having it accepted and published, the author was agreeing to transfer copyright to the publisher. While it is possible that an author could raise a question on older material, Elsevier Science believes that researchers will be pleased to have this more efficient electronic access to their older articles.

### Electronic archiving

As the world's largest publisher of scientific, technical and medical journals, Elsevier Science understands that the permanent archiving of

electronic information, and reliable access to these archives, are of critical importance to its customers and the scientific community at large. As a forerunner in the digital archiving of electronic journals, Elsevier Science was among the first to make a commitment to the preservation of digital files and involved librarians and authors to ensure that their policies and procedures would meet with approval. They have been in intensive discussions and intend to continue to take a lead in this area. This has resulted in an archiving policy and commitment in ScienceDirect licences that, for example, states that it is the intention of Elsevier Science to maintain the electronically archived copies of its journals permanently. The archives will be modified as the technology for document storage, display and access evolves.

Elsevier Science is also in active discussions with a number of university libraries, national libraries and preservation organisations to explore other alternatives for permanent archiving and is currently developing archival partnerships. In August, 2002, Elsevier Science and the Koninklijke Bibliotheek (KB), the National Library of the Netherlands, have signed a groundbreaking agreement: the KB will become the first official digital archive for Elsevier Science journals. This means the library will receive digital copies of all Elsevier journals made available on ScienceDirect, approximately 1500 journals covering all areas of science, technology and medicine, and exceeding 7 TB of data. This is an enormous project, just as backfiles, with a huge impact for everybody involved in research and the communication of research results, and again a proof of the fact that Elsevier is showing responsibility towards the scientific community in making an effort to keep not only print, but also digital archives available in perpetuity.