

Bulk Scanning and OCR

Scanning books page by page on a single sheet scanner is a slow process, so what are the alternatives?

Use a multi-function scanner/printer with a sheet feeder

Many modern networked printer/photocopiers can also scan documents and save them as digital files (usually image PDF or TIFFs). One option is to cut the spine off a book and feed the pages into the printer/photocopier. The machine will then scan each page and save it or email it to you. You still have to OCR the file to convert it to text, and add structure, but it saves you scanning it yourself.

Use a document scanning agency

CALL has developed a working relationship with DDSR (Digital Document Scanning and Retrieval), which is a local company specialising in document scanning. We have tested different scanning and OCR procedures and feel we can recommend the service that DDSR provide.

You send your books to DDSR and they send you back a CD or DVD with a PDF of the book, along with the book itself, if you want (minus the spine, of course).

We suggest you ask DDSR to:

- scan the books at 300 dpi and in colour (unless it's a black and white book);
- send you both an 'image' PDF and an OCRd PDF. The OCRd files are about as accurate and complete as they would be if you did it yourself using Acrobat Pro, i.e. textbooks with lots of images and textboxes may not be OCRd accurately, so you may need to check and correct them.

DDSR will also insert a copyright page if you wish and can add bookmarks to the PDF. It's a fast and cost-efficient service if you don't have your own multi-sheet scanner: DDSR charge about 6p per page for more than 1,000 pages and turnaround times are a few days. Contact:

Richard Wong
Managing Director
DDSR LIMITED
242 Glasgow Road
Wishaw
North Lanarkshire ML2 7QH
www.DDSR.com
01698 37 32 37