

At the Royal Library in The Hague more than 8 million newspaper pages will be digitised by 2011

Zeutschel makes headlines in Europe's biggest project for digitisation of newspapers

To historians and others, the newspapers we read and then toss aside so casually have a significance that endures long after the paper itself has turned to dust. For newspapers tell the story of our times. But how to preserve these historical records when paper, particularly newsprint, has a tendency to crumble with age? And how to make the history documented in such pages permanently and easily available to the reading public? Scholars, libraries and archives might once have raised such questions. But in recent times, digitisation is offering a new and wonderfully effective solution. And one of the first pioneers to take advantage of digitisation on a grand scale, at least in Europe, is The Royal Library of the Netherlands in The Hague.

“Europe’s largest digitisation project for newspapers sets new quality standards”

Project Manager Hans van Dormolen of the Royal Library in The Hague



How to preserve newspapers

Newspapers have a long tradition in the Netherlands. The first Dutch weekly paper was published in Amsterdam as early as 1618. This paper from a remote past is one of many hundreds to be permanently preserved in Europe’s biggest and most impressive project for digitisation of newspapers.

The project of the Royal Library is remarkable in its scope. The newspapers date back to the early 17th century, many printed on paper that threatens to fall apart. But it is the sheer volume that is staggering. More than 8 million pages of Dutch newspapers spanning 400 years, from the Library’s historical collection, must be digitised by 2011 and made accessible to the public via a web portal.

Such an ambitious project requires not only technology that meets the highest standards of excellence, but also people with the knowledge, expertise and experience needed to manage the many challenges it offers. Zeutschel OS 14000 book scanners ensure that each of the 8 million reproductions is of the highest quality, and that the very delicate, very old newspapers are protected in the process.

Millions of challenges

Digitising newspapers is not quite the same as digitising books or other forms of printed text. Digitising 8 million pages of newspapers that are hundreds of years old further complicates matters. Explains Hans van Dormolen, the man responsible for the project at the Royal Library: ‘Newspapers through the centuries have con-

tinuously changed the way they look. Not only has the format of the papers changed, but also the quality of paper used. Then there are the various typefaces and font sizes, which differ from paper to paper. Not to mention the many different kinds of binding.’ What was evident was that it would require fresh new guidelines and new standards to be set.

The project demanded high-resolution, uncompressed scans in a JPEG2000 format, in both grey scale and color, including the respective metadata containing structural information about page layouts. Images had to be segmented down to article level so that advertisements and journalistic articles, for example, could be displayed separately. This in turn would necessitate data preparation of scanned documents by means of automatic text recognition or optical character recognition. Ultimately, targeted searches would have to be possible through keywords using the library’s internal search engine module. Although the Royal Library has its own technical department, which has expertise in all aspects of digitisation, the sheer size of the project made the Library decide to outsource the task of data preparation to Hamburg-based digitisation specialist CCS GmbH, and the responsibility for scanning to Dutch specialist M&R Microimaging BV. Says Van Dormolen, ‘For this project we saw ourselves more in the role of a customer: specifying guidelines and monitoring their compliance.’ The Library decided, however, that conceptualising and setting up of the web portal, for public access to the digitised documents, would be done by its own experts.

Royal Library of the Netherlands in The Hague

The Royal Library (Koninklijke Bibliotheek) of the Netherlands was established in 1798. Today it is committed to preserve all texts – whether printed or handwritten – that make up the cultural heritage of the country, and makes this knowledge accessible to all. The collection of works at the Royal Library runs to approximately 3.5 million, of which 2.5 million are books. In addition, there are 15,000

magazines. The Humanities play a central role in the Royal Library’s holdings, the focus being on Dutch history, language and culture. The Library has been an independent administrative body since 1993. It is financed by the Dutch Ministry of Education, Culture and Science.

<http://www.kb.nl/>



The Royal Library (KB) of the Netherlands preserves the cultural heritage of the country

The one in a million scanners

The scanning of newspapers is a critical part of the digitisation project. Because of the wide variations in paper quality and also the fragile condition of much of the old newsprint, very finely nuanced machines are needed. Zeutschel OS 14000 work with precision and exactitude to consistently offer the high standards of quality that the project demands. For instance, high resolutions of up to 6.3 lp/mm (A0-Model) and 8 lp/mm (A1-Scanner) guarantee reproductions almost identical to the original. A patented LED illumination system gives low light exposure minus harmful UV/IR radiation so that both users and fragile old newsprint are protected. This special LED system also ensures high-speed scanning, an important feature given that millions of pages must be scanned.

Says Joerg Vogler, CEO of Zeutschel, 'Adherence to technological parameters is a key expectation of the Royal Library. In view of the huge number of documents, this makes great demands on the hardware being used. As far as Zeutschel is concerned, the calibration capabilities of our scanners are critical.'

The Zeutschel OS 14000 is designed with fine calibration in mind. There is an automatic function for white balance. Various important parameters can also be fine tuned: be it noise or sharpness, reproduction of tonal values or colour fidelity. In addition, a user can save different parameter combinations as permanent "set-ups" on the machine.

To prevent changes in image quality caused by mechanical, electronic or optical misalignments, test scans are conducted at regular intervals using calibration targets (comparable to test patterns on TV). Fine adjustments are then done manually.

A million times better in the future

But this whole process will soon be highly simplified with the development of a new universal test target (UTT) specifically designed for use in a library environment. The UTT will be the result of collaboration between Zeutschel, the Royal Library and German test laboratory Image Engineering, with Zeutschel acting in a leading capacity on behalf of the Association of Multimedia Information Processing (Fachverband für Multimediale Informationsverarbeitung e.V. – FMI e.V.).



400 years of Dutch newspaper history will be accessible on the Internet

Explains Volker Jansen, development manager at Zeutschel: 'With the aid of UTT's innovative test chart, all relevant parameters can be captured with only one scan. The image is then independently analysed using intelligent analysis software. Quality assurance is therefore completely automated. A calibration process which today might take two to three hours will be reduced to just a few minutes.' With this development, quality assurance will take place "in-line" during the production process rather than as a separate parallel process.

The specifications of the new universal test target (UTT) are accessible at <http://www.universaltesttarget.com/>. 'The objective of is to establish an open standard for quality assurance in digitisation, which will be supported by the community worldwide,' Jansen says.

Preserving eight million pages for posterity

Every week about 60,000 pages of newsprint are scanned at the M&R centre using 5 Zeutschel OS 14000 book scanners (4 x A1 and 1 x A0). So far, 1.2 million documents have been digitised. In order to comply with the strict quality specifications of the Royal Library, a quality control is carried out each day. 'Operators who've been specially trained conduct the test scans in the morning and then send the results to the Library, where the results are diligently checked. Only after we hear that our quality meets the specified standards do we start the daily scanning,' says Louis van Erven, Managing Director of M&R.

This leading Dutch scanning service provider has a state-of-the-art data center and years of experience in the areas of microfilming and digitisation. A part of the historical newspaper archive of the Royal Library was microfilmed years ago. In the second stage of the project, these documents will also be scanned and digitised so that they can be uploaded onto the web portal. For this purpose Zeutschel OM 1600 microfilm scanner is currently being tested.

www.zeutschel.com

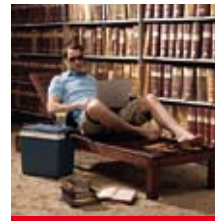
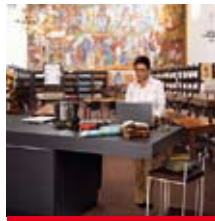
The Zeutschel Bookscanner series Omniscan 14000

Millions of books, maps and files are lost every year because they are mislaid or totally ruined. With the disappearance of these valuable cultural assets we normally also lose the information they contain – forever. What that means for the future of human society cannot be expressed simply in figures. Therefore libraries and archives are not giving up on their efforts to preserve this cultural heritage for future generations.

With our recently developed Omniscan 14000 high performance scanner, historic documents can now be scanned with even higher quality and speed. With a scanning speed starting at 6.5 seconds (A1-size) and an illumination system which excludes ultraviolet light, the most sensitive materials, measuring up to 880 mm x 640 mm, can be digitized. For the protection of cultural assets the new Omniscan 14000 represents a new benchmark in the market exceeding even highest requirements and combining efficiency with an ergonomic design.



Document access – like being there



Zeutschel – The Future of the Past

Zeutschel GmbH, based in Hirschau near Tübingen, is a well-known specialist company working in the area of the preservation of cultural assets and has been a supplier of equipment, solutions and systems for the management of documents and archives for more than 40 years.

Zeutschel develops, produces and distributes bookscanners, overhead scanners, capturing software, document delivery software, microfilm cameras, hybrid systems and microfilm readers. Zeutschel systems and solutions are used by libraries, archives, universities, land registries and commercial companies all over the world.



Zeutschel GmbH · Heerweg 2 · 72070 Tübingen
Tel.: +49 7071 9706-0 · Fax: +49 7071 9706-44
info@zeutschel.de · www.zeutschel.com