



Digitizing 19th Century Newspaper Collections

A Large National Library

The client is the largest reference library in the world and is responsible for maintaining and developing a national collection of library materials

Project / Service Category

Archival Digitization

Services Provided

- Article Segmentation
- Article Categorization
- Indexing

Customer Speak

HTC delivered good quality work and were always forthcoming with questions to seek clarification with the team. This demonstrated their commitment to providing quality deliverables.

- Director – Digitization

Background

The library wanted to digitize over 4 million pages of 19th century newspaper collections as part of its mass digitization program to enable easy online access to researchers and the public. HTC was selected as a prime vendor for this digitization project.

Challenges

- High density of articles and varying content structure
- Varying image contrast levels

Solution

We deployed customized applications to automatically segment the articles, capture the metadata, and generate METS / ALTO xml files. We analyzed the newspaper contents well ahead of the production cycle and kept the solution ready for managing the digitization requirements for the varying content types. This helped us to deliver the project on time throughout the contract period.

Our customized imaging tools helped us to enhance the contrast level of the images to enable better OCR recognition. We delivered around 100,000 pages per week and met the tight delivery schedules in spite of the challenges.

Benefits

- Easy, anytime, anywhere online access for the contents
- Our consistent high quality deliverables minimized the workload for the client's Quality Control team
- The client registered a 25% increase in their online subscriptions



www.htcinc.com

- Over 27 years of IT and BPO experience in providing cost effective and innovative services across domains for a wide spectrum of global 2000 customers
- Quality and information security processes compliant to SEI CMM Level 5, ISO 9001, ISO 27001 and PCI DSS standards
- Global presence, large delivery centers, talented professionals, capabilities across multiple technologies, platforms and domains

Reimagine technology to accelerate your business

USA | UK | Germany | India | UAE | Australia | Malaysia | Singapore | Indonesia