



Reference Parsing for Journals

A leading content aggregator

The client is a leading provider of learning and research solutions for the academic, professional, and library markets worldwide.

Project / Service Category

Publishing - Reference Parsing

Services Provided

XML conversion

Background

The client wanted to convert the references available in journals into XML format.

Challenges

- Various input formats like PDF, HTML, XML, ASCII
- Varying reference patterns
- Handling West European languages
- Turnaround time – 24 to 36 Hours

Solution

This project involved extracting metadata and text from the input files and performing semi-automated tagging for the various elements. Our semi-automated solution minimized manual intervention by 80% and increased productivity. We analyzed unique reference patterns and created templates ahead of production to handle varying reference patterns. Auto Translation helped us to handle West European languages.

Benefits

- HTC's output consistently exceeded the client's quality expectations
- Semi-automated approach helped to reduce the production time and enabled the client to realize cost benefit



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