



# Digitization of Substantial Collection of Eighteenth Century Historical Works

## A Large Publishing Company

A leading educational content, and service provider for higher education, professional, and library markets worldwide.

## Project/Service Category

Enterprise Content Management

## Technology

- JDK, Eclipse, Maven, JEE, XML, XSLT, Spring
- Sybase
- TomCat

## Customer Speak

The quality of images and XML data that you have delivered has helped us create a highly sophisticated digital library. This has revolutionized the way research is done using old English books.”

- Director – Technology

## Background

The client wanted to digitize a huge collection of 18th century books available as microfilms and convert the text images to XML data for the world's largest library of historical works (from 15th to 19th century) with more than 150,000 printed multilingual books and publish their digital editions for enabling easy search, retrieval, display, and print.

HTC was contracted to design, develop, and implement the application for Eighteenth Century Collections Online (ECCO) and digitize the Eighteenth Century Collections.

## Solution

This project involved publishing a digital edition of one of the world's largest library of historical works with more than 150,000 printed books having greater than 46 million pages in a time span of one year for web availability. To achieve the desired objective of creating XML format data, HTC leveraged its Imaging and Document Management framework to deliver over 3.5 million pages per month. The framework enabled HTC to perform image capture, Optical Character Recognition (OCR), the capture of optional items (index fields, table of contents etc.), and generation of XML output of text according to client specifications.

## Benefits

- Easy search and retrieval of digitized collection with the implemented next generation ECCO application
- Potential to provide access to the digital books for a large number of users from anywhere at any time with ease
- Ability to reuse existing resources limited by its format such as the materials stored on microfilm



- Over 25 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