

CASE STUDY- Information and Media - Publishing

Collection of Construction Project Information in North America (BIDS)



About the Client

A Fortune 500 company that provides industry news, market analysis, trends and forecasts, online industry intelligence, project information for the Construction Industry

The Business Challenges

- The information is in HTML, word format, images, PDFs, feeds, across websites at the federal, state and local county, municipality level. This information is time sensitive and must be gathered within stipulated window periods
- Track over 20,000 sources of information daily
- Bid information from websites to be delivered within 24 hrs of posting
- De-duping records that come from multiple sources
- Pool the output as a XML feed

Results

- Crawlers were developed to handle PDFs, Feeds, And other challenging formats and the output is pooled to send out a single XML feed to the client
- Time zone specific – we schedule sites based on the time zones. We instruct the crawlers to run at specific time during the day

Benefits

- Distributed information – when a page contains disturbed information in various columns/rows/headers/footers we can scrape all this information more efficiently than a human can
- A single XML feed containing all the day's construction information is of immense utility to the client who uses for their analysis and intelligence products

For more information, Contact info@mobius-bizinfo.com

"Pantheon Plaza"

485-486, Pantheon Road, Egmore,
Chennai – 600 00, Tamil Nadu, India

Phone: +91-44-4298 6400

www.mobiuservices.com