

Page-Turning Solution Propels Books into the Digital Age

Synopsis: Kirtas Technologies, a provider of book digitization solutions and services, uses ABBYY Software's optical character recognition engine to capture text and images of published works. When Kirtas put several OCR engines through their paces, ABBYY FineReader out-performed them all. And, as the relationship has evolved and expanded, ABBYY continues to provide outstanding support to Kirtas and its growing customer base.

Customer: Kirtas Technologies, Inc.

Headquarters: Victor, N.Y.

URL: www.kirtas.com

Founded: June 2001

Founder and CEO: Lotfi Belkhir

Industry: Digitization of out-of-copyright bound books and other publications

Clients: Small, midsize and enterprise businesses; individuals; libraries; government agencies; not-for-profit groups

Customer List: Includes Microsoft; LexisNexis; Johns Hopkins University, Idaho National Laboratories, The United Nations, U.K. Climbers' Club

Problem: Address the growing digitization market by creating a complete hardware and software solution that would digitize books and other publications rapidly, accurately and gently, and result in an easily searchable format

Technology Tools: ABBYY FineReader Engine, Kirtas' automated page-turner solutions

Systems: Kirtas' APT 1200; APT 1212; APT 1600; APT 2400; APT 2400RA; BookScan Editor; BookScan Editor Pro 2.0, integrated with ABBYY FineReader Engine

Results: Kirtas developed a family of digitization solutions that allow the rapid digitization of content that is locked into bound documents, such as books and manuals. Using BookScan Editor Pro 2.0 and ABBYY FineReader Engine, Kirtas' solution provides extremely accurate OCR conversion in more than 177 languages and multiple output formats, such as online, CD, hard drives, e-books or printed-on-demand (POD)

ABBYY FineReader Engine Complements Kirtas' Sophisticated Digitization Systems

From pages yellowed by time and touch, to carefully typed reams of documentation, Kirtas Technologies specializes in digitizing the vast array of knowledge currently housed on library shelves, in government archives and corporate storerooms.

The Victor, N.Y.-based company developed an automated solution that gently turns even the most fragile pages. Kirtas' software – including BookScan and BookScan Editor Pro, coupled with ABBYY's FineReader Engine – transforms words, pictures and charts into digitized images that can easily be scanned and manipulated.

"The average book takes less than eight minutes," said Lotfi Belkhir, founder and chief executive. "Kirtas developed the first and fastest automated book-scanner. The whole mission of our company is focused on the fact that digital is better than analog, and our belief that future generations will want to access knowledge electronically."

Digitizing Tools

From its inception, the Kirtas team knew it would need powerful, accurate tools to keep a book's text and graphics intact and exact.

"There was no software out there to crop images with the level of quality you needed to print," Belkhir said. "There were no software packages customized for books. We also needed an optical character recognition (OCR) engine that was accurate, reliable and flexible."

The company developed the Kirtas APT BookScan systems, which feature intuitive user interfaces to simplify the capture, organization and storage of books. It also created Kirtas BookScan Editor to ensure consistent page positioning and outstanding type integrity that maintains or enhances the quality of photographs and illustrations. BookScan Editor PRO features extremely accurate OCR conversion in more than 177 languages and multiple output formats, including PDF, XML, MS Word and text.

BookScan Editor PRO is powered by ABBYY FineReader Engine, an intelligent and accurate OCR engine. “We looked at many different OCR engines,” Belkhir said. “ABBYY FineReader was by far the most accurate with the best throughput. Although accuracy was good with some other programs, ABBYY also won the reliability tests hands-down. FineReader far exceeded all other OCR applications we tested.”

Reliability is critical to Kirtas’ success. Their systems can handle up to 2,400 pages per hour and the company digitizes up to 25,000 books a month through its in-house service bureau. In addition to providing this service, Kirtas also sells its hardware and software to organizations that want to keep the process in-house, Belkhir said.

Given the diverse formats and art elements of published materials, flexibility was also vitally important, he said. FineReader’s feature-rich Software Development Kit (SDK) allows Kirtas to quickly and easily customize the software as needed, Belkhir added.

“We have built automated language selection into our process, and ABBYY’s flexibility is the tool that allowed us to do that,” he said.

Strong Support

Kirtas has a full staff of professional image technicians who are proficient in digitization using FineReader and Kirtas’ array of products. According to Belkhir, whenever the internal team had a question, ABBYY’s technical support department never let it down.

“We did all the customization, and ABBYY’s outstanding technical support was a large enabler of that,” he said. “We know their developers by first name, and they always work with us through our problems until we are completely satisfied.”

This support is helping Kirtas expand. When the company wanted to support more languages, it relied heavily on ABBYY’s technical staff to achieve this goal. Kirtas plans to continue adding languages – and continuing to turn to ABBYY for design and implementation assistance, Belkhir said.

Off the Shelf

An ever-growing number of organizations are turning to Kirtas to transform their hard-bound documents into digitized publications. The seven-year-old firm’s clients include Microsoft, LexisNexis, Yale University, Cornell University, Johns Hopkins University, and the United Nations. But Kirtas also reaches out to smaller organizations, such as the Climbers’ Club in the United Kingdom, which tapped Kirtas to digitize its 121 journals, ranging in publication date from 1898 to 2005.

Duke University, a satisfied customer, introduced the Climbers’ Club and Kirtas, recalled Christina Sherman, marketing manager at Kirtas. What began as a way to capture and protect journals, some of which were more than a century old, actually has evolved into a revenue-generator for the Climber’s Club, she said.

“Anyone interested can search for the club’s digitized journals and, when copies are ordered, revenue is generated for the club,” said Sherman. “They chose Kirtas because of the outstanding image quality and the OCR accuracy produced by our process.”

Due to the documents’ age and condition, damage-prevention was also important to the club, said Rick Parrinello, production manager at Kirtas Service Bureau (KSB). “Our page-turning machines are gentle enough to work with a book of any age,” he said. “We get a lot of older books that, if you touched them incorrectly, would fall apart.”

Digitized images, of course, can withstand any amount of man-handling. By bringing long-lost printed treasures, copies of ancient manuscripts and different cultures’ writings to the global community, Kirtas and ABBYY FineReader Engine are delivering priceless views of the world from today and yesterday.