
XML Marks the Spot: XML Helps Move Knowledge from Books to Bytes

Lotfi Belkhir

Abstract

The discussion will share advancements in the areas of digital capture, storage, management, access and output. It will review the significant benefits and cultural implications with the digitization of information, focusing on software and storage solutions creating easy access and search capability for scanned information. A demonstration and review of the automatic bookscanning process relating to the use of XML will share how modifications can be made to a pre-existing XML file.

With multiple file outputs, scanning technology allows for optional hard copy or digital output, giving users options for sharing and using information. Post-processing allows for image enhancement and clean-up, eliminating marks on older texts and creating a completely clean and searchable digitized version.

The digitization of information will have a notable impact on digital culture and preservation:

- o Cultural heritage and digital culture now can work in tandem, preserving and sharing both ancient learnings and the latest findings in a seamless digital stream.
- o The impact on productivity and society is significant. With such an unprecedented amount of information sharing, organizations can share intelligence more rapidly and in higher volumes.
- o The ramifications for cultural institutions, government, religious, finance, law, and accounting are also significant with digitization.

Specifically, BookScan automatic book scanners feature the "APT Manager," and "BookScan Editor" to facilitate the sharing and searching of scanned information. The APT Manager offers a number of post-processing software solutions, offering a significant amount of control to the end user operating the automatic book scanner. Technical metadata indicates camera and imaging setup, and other software capabilities allow control of the scanner page width, clamp pressure, vacuum and fluffer speeds, machine settings, and machine status. New or modified settings can be applied to the machine easily. Such tight control of scanned settings and information from the APT Manager allows for greater efficiency and productivity in the scanning process.

The BookScan Editor features three levels of fully-customizable metadata. It uses an automatic batch mode processing operation that performs extensive image post-processing functions. These include:

- * Image Cleanup: rotation, de-skew, cropping, clamp removal
- * Image Enhancement: brightness, contrast, TRC, luminance, sharpen adjustments, bitonal (DLT algorithm) and grayscale conversion, curvature correction, image segmentation, text centering, page padding, checksum, page number contact sheet.
- * Mutiple output options: 300-600 DPI and TIFF, JPEG and compressed PDF (2x to 100x).
- * Integrated interface for structural metadata

Focusing on the structural metadata, using MSWord or Notepad or any other text editor, Kirtas Technologies can input and track scanned information as metadata in a customer's scanned document directory. Up to 16 elements of metadata can be submitted for inclusion in the scanned document's directory. Each scanned book has its own XML-formatted metadata file within the book's directory.

Table of Contents

1. Product Presentation Paper	4
-------------------------------------	---

1. Product Presentation Paper

The author did not prepare a summary of this product presentation for the proceedings.

Biography

Lotfi **Belkhir**

CEO

[Kirtas Technologies Inc](http://www.kirtas-tech.com) [<http://www.kirtas-tech.com>]

Victor

New York

United States of America

Lotfi Belkhir brings more than two decades of experience in research, development, and management to his role as CEO of Kirtas Technologies, Inc. In 2001, Lotfi founded Kirtas with the vision of digitizing information from bound documents, achieving an unprecedented level of speed, accuracy, and quality. At Kirtas, Lotfi oversees the management and implementation of the company's revolutionary automatic bookscanning technology, which is redefining digitization and information sharing in the marketplace. In his career, Lotfi has held key positions at Xerox Corporation facilities in New York, Connecticut, and California. While at Xerox PARC, Lotfi negotiated the exclusive license to key book scanning technology. Lotfi holds a Ph.D. in Condensed Matter Physics from the State University of New York at Stony Brook, and an MBA in Management of Technology from National Technology University. For more information on Kirtas, please visit www.kirtas-tech.com.