

Opus 3

PRODUCTION DIGITIZATION WORKFLOW SOFTWARE



One Product ... Six Exceptional Advancements

- Next Generation Workflow with Automatic Archive
- Superior Automatic Image Treatment
- Fully Integrated Hierarchical Metadata Capture
- Web-ready Output with Virtual 3-D Page Turning
- Open RAID Digital Archiving & Migration Management Software using Automated Data Migration Facility
- Enlightened Architecture

A New Era in Digitization

Get into the game with Opus! Long have the complexity and high startup and operating costs of digitization projects kept most libraries, museums and archives on the sidelines. Opus leverages today's high-performance, low-cost computers to obliterate those barriers with its paint-by-numbers digitization system that produces the highest quality digital collections possible at less than 1/3 historic startup and operating costs. In as little as one day, Digital Library Systems Group (DLSG) installation and training staff will leave you with a fully operational digitization system. Follow-on training sessions via the Web ensure that your staff is producing multiple beautiful digital books per day in no time, complete with hierarchical metadata.

With the new low prices and high-performance of the latest book scanning hardware, complete digitization labs are surprisingly affordable.

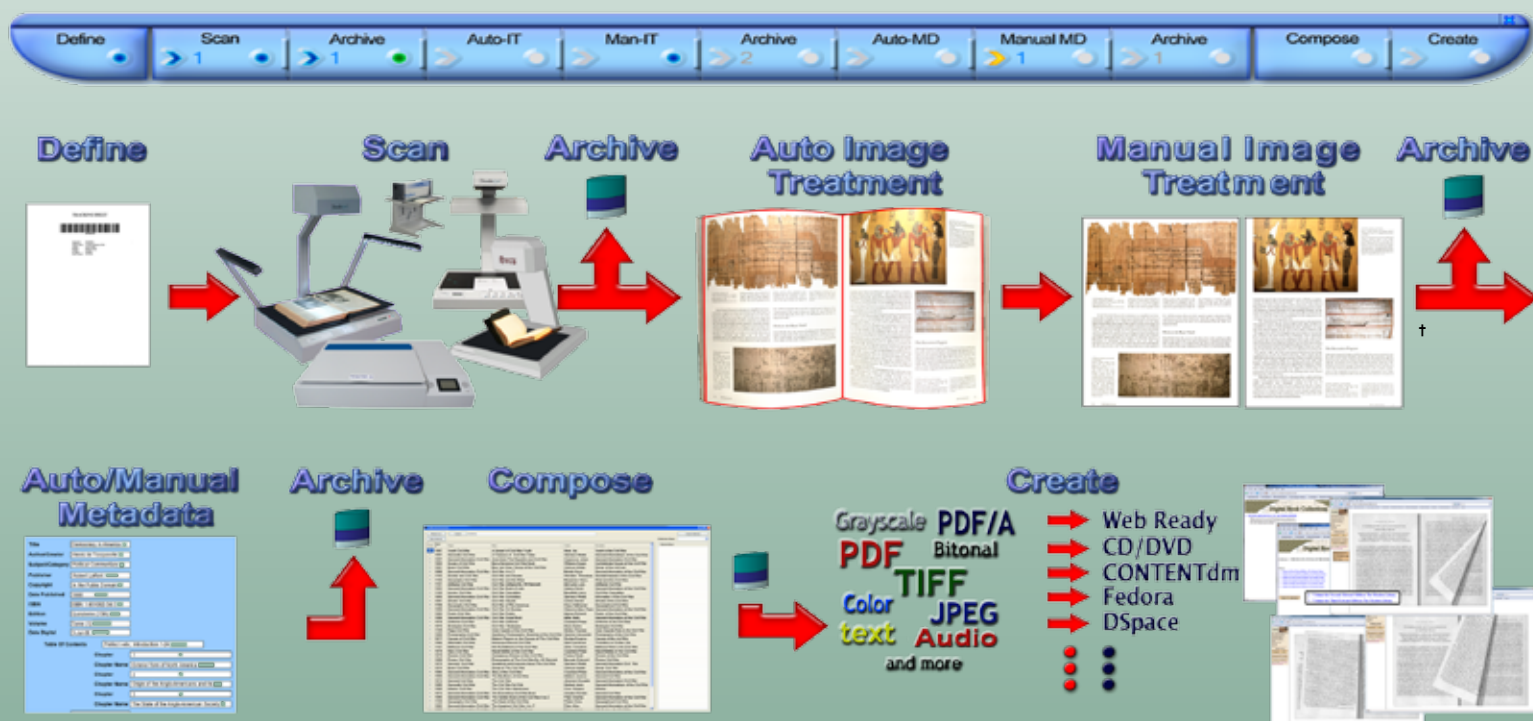
The digital archives that Opus creates can be accessed without proprietary software. Using its Compose Collection and Create Collection facilities, Opus can create an unlimited number of digital collections in nearly unlimited formats (derivatives) any time in the future. Simply select the desired volumes and data formats and let Opus create your collection(s) overnight. In addition, Opus' automated data migration facility makes it easy to transfer an entire digital archive to new media while simultaneously converting the image and metadata files to the latest format standards.



OVERVIEW

Opus software processes raw images directly from book scanners, flatbed scanners and digital camera systems. Opus groups images into objects as it captures them from the scanner or camera system, then manages those objects through a complete image treatment, metadata capture and archiving system, ensuring that all required processes are performed on each page of each object. It then archives the objects in non-proprietary formats of your choosing. Finally, Opus can output repeatedly to the Web, to CD, to viewing or presentation software, etc. Opus is also compatible with Open RAID digital archiving and migration management software.

Opus Digitization Workflow System



Opus Highlights

- Controls scanners and groups images into objects
- Automatically archives raw scanned images
- Automatically separates two-up images, removes bookfold distortion, fan, gutter, skew and other undesirable artifacts of digitization
- Allows manual review and adjustment of automatic image treatment results
- Supports third-party image editing software (e.g. Photoshop)
- Automatically archives treated images
- Automatically captures hierarchical metadata
- Allows manual review, adjustment of automatically captured hierarchical metadata and manual keying of additional hierarchical metadata
- Automatically archives hierarchical metadata
- Unlimited numbers of archive volume storage units
- Automates data migration
- Data Migration to new formats as they evolve
- Allows an unlimited number of collections to be composed from an archive
- Outputs an unlimited variety of image and metadata derivatives in a wide variety of formats
- Creates complete, ready-to-go digital library Web sites with virtual 3-D page turning
- Displays real-time workflow status changes
- Provides easy, real-time tracking of objects as they are processed through the workflow
- Supports multiple predefined as well as custom image treatment templates
- Supports multiple predefined as well as custom metadata capture templates
- Supports multiple workflow workstations
- Allows for multiple metadata capture sets to be collected for each object
- Allows for the reprocessing of objects under new or different image treatment algorithms without requiring the object to be rescanned

Elements Of Opus

Workflow Manager Opus Workflow Manager completely manages all objects throughout the entire digitization process while also allowing third party software to be used where desired; provides tools for monitoring all objects as they move through the digitization workflow system to make managing projects and scheduling workers a breeze; and minimizes the cost of rework, whether it be skipped or improperly performed work, or even a change in the fundamental rules governing a digitization project.

Automatic and Manual Image Treatment Opus completely automates the Image Treatment processes by separating two-up images, removing bookfold distortion, fan, gutter, and skew; all in a single operation. Opus also provides a streamlined graphical interface to manually perform the specialized image treatment functions required for both digital preservation and the various types of access.

Integrated Hierarchical Metadata Capture Opus' Metadata Capture stages are not part of any digital collection viewing software, but instead are built seamlessly into the Opus digitization workflow system. The Opus Metadata Capture stages allow users to select from a list of standard hierarchical metadata templates, create new ones and restrict users to a subset of templates for a given project. The templates include metadata characteristics like requiring certain metadata to be captured, allowing some types of metadata to be repeated and requiring a single value for other types.

State-of-the-Art Customizable Output Derivatives and Formats Opus' Create Collection stage provides output that can be imported into open source and proprietary content management and digital library systems. Since storage and internet bandwidth requirements rise as image quality rises, Opus lets the user select the ideal compromise for each access system that collections are deployed to. If requirements change, simply output the collection again with different clarity parameters.

Automated Archival Opus adheres to the Open RAID standard. During the digitization process, there are three points at which archival is necessary: 1) archival of images immediately after scanning and before image treatment is performed; 2) archival

of images after image treatment has been performed; and 3) archival of metadata after metadata has been captured. Opus creates archives that are open architecture in formats that are accessible by today's communities of users, and with the proper data migration plan, will be accessible by its future communities of users.

Data Migration Facility Opus provides an Archive Management and Data Migration facility using the Open RAID standard, which is intended to be run periodically, before the possibility of any media-related data loss becomes significant and before the media or data formats become obsolete. Opus' data migration facility runs unattended. While moving the data to new media, it can convert images and/or metadata from its existing standardized formats to new standardized formats, ensuring that data is never lost and is always in a format that is accessible by its community of users.

Digital Collection Web Site Creator Opus provides Web-ready output that allows viewers to select a volume, open it and navigate from section to section and from page to page. A lifelike 3-D virtual page turning experience is provided as well. Opus creates the image files, downloadable PDFs, HTML, etc. including an entry point file called default.html and places them in a folder. To deploy a Web-ready digital collection, simply place the contents of the folder created by Opus on a Web site server and link the default.html file to an existing Web page. A thousand-book collection can be deployed in little more than the time it takes to copy the data.

Enlightened Architecture The Opus architecture originates from a clear understanding of the core purposes of each element of digitization and access systems. In the book publishing business, rarely are books manufactured in the stores that sell them. Not only is manufacturing separate from retail, but distributors exist to attend to the specific needs of each of their stores while providing a simple, consolidated relationship with book manufacturers. In some cases, distributors provide warehousing services and in other cases, product warehousing is performed by the manufacturer. Opus serves as the manufacturer, warehouse and distributor.

Opus Digital Workflow VS. Digital Collection Viewing Software

Opus digitization workflow software has very little in common with digital collection viewing software. Opus creates digital assets and the latter provides day-to-day access. Opus users are predominantly archivists and digitization workers whose work is mostly done once their digital collections have been created. In contrast, the digital collection viewing systems are used by a much larger community with far more varied needs and interests.

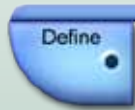
In fact, Opus creates output for presentation and content managers as well as output that can be accessed directly by Web browsers. In the future, when new content management systems are developed to meet new and different needs, Opus will output part or all of the digital collections it has created to the new systems in new formats.

Open RAID Archive and Migration Management Software

Opus supports the Open RAID storage method for archiving digital masters. Open RAID was designed for static data that requires extremely safe, but ideally low cost, long-term storage. Since digital data does not degrade over time, it is a perfect method of archiving important information, provided the risk of data loss can be eliminated. The Open RAID storage method alone is inadequate to protect data for extended

periods. However, since Open RAID is very low cost and its data can be transferred with little human effort, entire archives can affordably be transferred to newer media every three to five years. During the data transfer, formats of the data can also be updated to conform to the latest standards.

OPUS WORKFLOW MANAGEMENT



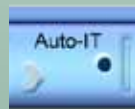
This stage is used to print a tracking sheet for each book, magazine, newspaper, map or other object to be scanned. To identify an object, the scan operator scans the tracking sheet prior to scanning the object and Opus reads the object identifying information from the sheet. A metadata template can be specified in this stage for each object to be digitized.



The Scan stage supports scanning of ultra-large images, including multiple gigabyte raw TIFF images produced by the world's largest scanners. Image objects which are managed by Opus and processed with the digitization workflow stages are created in this stage.



The first Opus Archive stage automatically archives the raw images prior to any image treatment so that the subsequent image treatment stages can be rerun with different parameters, a month or even years later, to meet future needs. For example, an object's images can be passed through the image treatment stages to produce the best possible representation of the 'aged' original. Another pass through image treatment can produce a representation of the 'unaged' original, as it was when it was initially printed.



The Automatic Image Treatment stage automatically determines book skew, locates image content for creating consistent margins, determines bookfold distortion borders and detects the most representative background color to allow for image smoothing. Image treatment functions are not limited by image file size.



The Manual Image Treatment stage provides a fast graphical interface that is optimized for the image treatment functions that are common to digital preservation, archive and access. In seconds, an image's skew angle can be corrected. In less than a minute, all functions can be performed on an image. If a user requires a special image treatment process that is not supported by Opus' production facilities, a popular photo editing program can be launched with the current image automatically loaded. Upon completion, the modified image is automatically reintegrated into the Opus system.



The second Opus Archive stage automatically archives the images processed by the image treatment stages. Multiple versions of an object can easily be maintained by the archive system.



Opus automatically captures much of the metadata specified in the metadata template that was previously chosen in the Define stage. The metadata is stored in a non-proprietary XML format during the digitization workflow process. When a digital collection is created, the metadata is transformed to the chosen output format(s).



This stage provides an intuitive hierarchical graphical interface for entering data. Data is keyed into a hierarchical entry screen which matches the template selected for the current object. The same capture interface is used to create new metadata templates or modify existing templates.



The third and final Opus Archive stage automatically archives the captured metadata in a non-proprietary XML format. Multiple derivatives can be created, such as a lower resolution (small file size) JPEG version for the Web created from the 'aged' originals, and a high resolution print-ready PDF version created from the 'unaged' original.



The Compose stage allows unlimited numbers of digital collections to be exported repeatedly with different image and metadata file formats. Each digital collection is comprised of a list of objects along with the desired output formats and clarity/size parameters. Output formats include many image formats (e.g. TIFF, PDF, JPEG, PNG, etc.) with file name control and many metadata format such as flat file, METS, MODS, MIX, Sublin Core, 3-D virtual pageturning, and website ready.



This stage provides unattended output of entire digital collections comprising many volumes. Processing can be performed overnight, unattended.

OPUS IMAGE TREATMENT

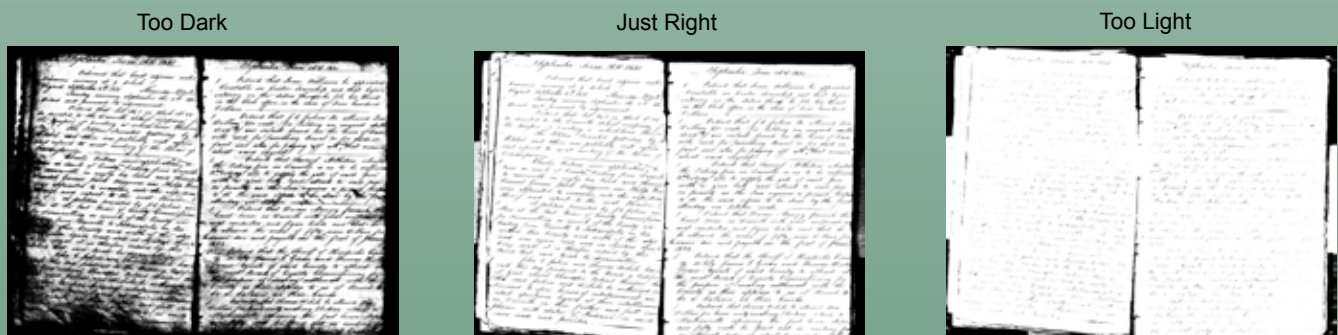
Opus takes the tedium out of image treatment by providing users with an interface that is wholly dedicated to processing images for archive, preservation and access. Even problem images can be fully processed in seconds. Prior to manually reviewing and adjusting images, Opus performs image treatment functions automatically. If any automatic image treatment function requires additional adjustment, Opus allows the user to graphically 'assist' Opus and make adjustments in seconds.



Automatic image treatment removes skew, adjusts the border, and eliminates fan and gutter. The treated image may be further refined by manual treatment.



Book curvature produced from scanning tightly bound books may be realigned using Opus' sophisticated book curve correction tool. Text and photos are returned to their original alignment and orientation.



Opus allows for the output of a large variety of derivatives. Controlling threshold optimizes the clarity for each type of derivative which, in turn, produces the highest quality image for that output request.



Artifacts of all types may be removed from an image while retaining the underlying variations in shading. Fingers, blemishes, stains or even handwritten notes may all be removed to produce an image that closely resembles the original. Opus enables you to publish both the treated and untreated image.

Photostats and negatives are easily inverted revealing an otherwise unreadable document.

OPUS METADATA CAPTURE

Opus' hierarchical metadata capture is controlled by templates that can be created and modified to fit virtually any requirements. Input field characteristics include: required/optional, repeatable/one occurrence only and other characteristics that can be specified when templates are constructed or modified. Opus provides several generic templates to capture the data necessary to create virtual library output formats.

Opus provides a straightforward graphical interface for creating and modifying hierarchical templates. The templates carry the characteristics of XML elements and attributes. Data entry rules such as required/optional and single/multiple can be assigned to each field. During metadata capture and entry, Opus enforces the rules, thus simplifying metadata entry and reducing the possibility of errors.

In addition, the Opus hierarchical metadata input facility allows a user to scroll back and forth through the images of an object (e.g. pages of a book) while reviewing and entering metadata.

The screenshot shows two windows from the Opus Metadata Capture interface. The left window, titled 'Template Name: Journal Style 3', displays a list of metadata fields: Abstract, Access Condition, Appendix, Author/Creator, Back Matter, Body, Chapter, Chapter Name, Classification, Contributor, and Copyright. The right window is a data entry form for a book titled 'Democracy in America' by Alexis de Tocqueville. It includes fields for Subject/Category (Political Commentary), Publisher (Robert Laffont), Copyright (In the Public Domain), Date Published (1968), ISBN (ISBN 1-931082-54-5), Edition (Quinzieme (15th)), Volume (Tome (1)), Date Digital (3-Jul-06), and a Table of Contents section with entries for Chapter 1 (Exterior form of North America) and Chapter 2 (Origin of the Anglo-Americans, and importance).

OPUS COLECTION COMPOSER

The first nine stages of Opus create normalized digital masters. The Opus Compose Collection facility withdraws objects from the stored normalized digital masters and publishes them to their final destination(s).



Composing collections couldn't be easier. Select the digital volumes that comprise the desired digital collection in minutes with Opus Collection Composer. Simply name the collection and begin adding objects by browsing the digital archives managed by Opus. A simple click of the mouse adds a volume to the collection.

Digital archives are reusable assets, especially when managed by Opus. Once you have created your archive, you can use it anytime to compose and automatically create new and different collections, course curriculum materials, etc. Online digital content and print-on-demand add to the many digital formats.

OPUS IMAGE OUTPUT - DERIVATIVES AND METADATA

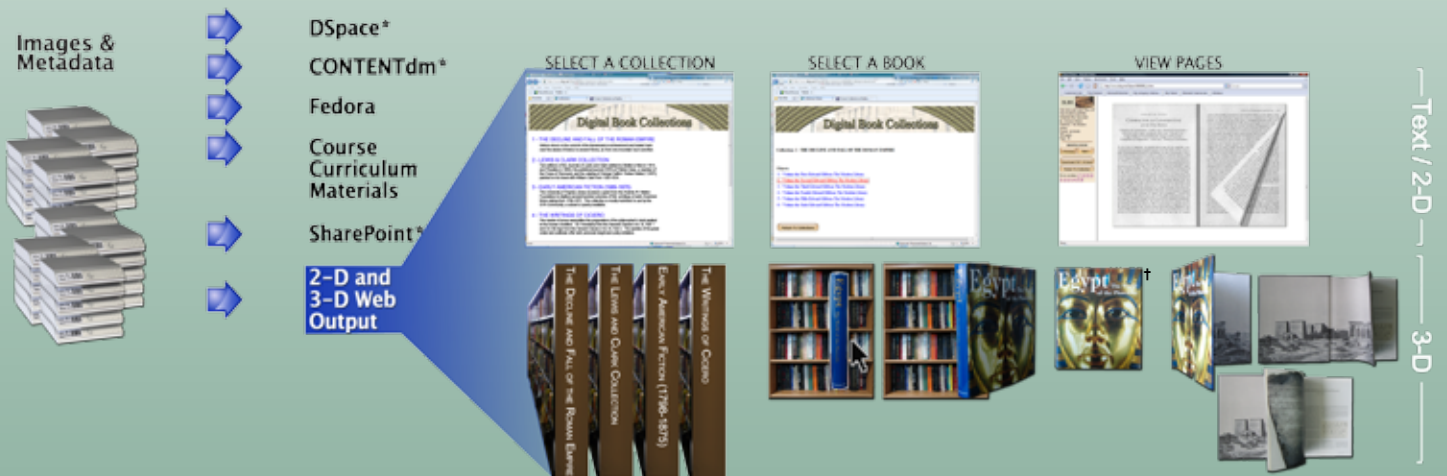
Opus has a library of script files for common XML, flat file and other output formats for the metadata and associated images of digital objects (e.g. METS, MODS, Dublin Core). This script library is expandable with an unlimited number of custom formats. In addition to selecting the desired output format(s) for the metadata, the clarity and output formats of the images can be specified (e.g. TIFF, JPEG, PDF, PNG,

etc.). DLSG staff technicians can assist in modifying existing formats and creating new custom output formats.

The most exciting and immediately usable format that Opus creates is a Web-ready output that provides the experience of browsing the virtual stacks, selecting a digital book, opening it and turning the pages; as though the printed book were in front of you.

Digital Masters

Final Destinations



MAJOR ADVANCEMENTS IN DIGITIZATION HARDWARE

During the past few years, all makers of preservation quality scanners have released new products, and nearly all at lower prices than their predecessors. For example, the Bookeye 3, made by Image Access Europe, was released in early 2006 and represented a breakthrough in price/performance for very large format (24 x 36 inch) open scanners. It has a motorized book cradle and twice the bed size of the widely used Bookeye 2, yet it costs little more than the Bookeye 2. Then in early 2008, Image Access Europe released the Bookeye 3 A2 with true 600dpi resolution, an amazing breakthrough in planetary scanner image quality.

Since the introduction of digital storage media, costs have fallen nearly every year. In 1980, the cost to store the images of a single 300 page book on-line was well over \$10,000. Today, \$10,000 of on-line storage will hold the images of 250,000 books. In fact, this type of media has become so inexpensive that terabytes of self-contained, redundant disk storage devices are available in stackable, six inch cube packages. This type of media is becoming popular for digital archives because many terabytes can be connected quickly for automated migration to the next generation media. While CD storage is still the most affordable, its redundancy is not self-contained, and handling costs can outweigh the savings.

* CONTENTdm is a trademark of OCLC

SharePoint is a trademark of MicroSoft Corporation

DSpace is a trademark of Massachusetts Institute of Technology

† Published 1998 Könermann Verlagsgesellschaft GmbH Bonner Straße 126, D-50968 Köln | Edited by Regine Schulz and Matthias Seidel | Egypt the World of the Pharaohs

IMAGE ACCES AND THE DIGITAL LIBRARY SYSTEMS GROUP



Since 1993, Image Access has been providing digitization workflow solutions to commercial markets. Having developed its own OCR, barcode reading and image treatment algorithms, Image Access is capable of meeting virtually any and all specialized digitization needs.

Motivated by a great lineup of hybrid library products under development, Image Access created the Digital Library Systems Group (DLSG) in 2004; a division whose sole responsibility is the ongoing development, service and support of the best hybrid library digitization products possible at prices that virtually any academic or research library can afford.

Over the years, the Image Access factory in Europe has produced a variety of high performance production scanners, including 600dpi book scanners with cool white LED lighting and built-in motorized cradle, large flatbed scanners that can scan 18x25 inches in full color in 3 seconds, and feedthrough scanners capable of scanning large drawings and both sides of full broadsheet newspapers simultaneously. All scanners are equipped with the revolutionary Scan2Net® Ethernet/HTML interface. In addition, the Image Access factory provides its state-of-the-art scanner electronics to several major commercial scanner makers.

Image Access U.S. remains focused on enhancing its digital library software products while Image Access Europe continues to develop new and better scanners. Together, we're committed to providing ideal digitization solutions to libraries and other research institutions for the foreseeable future.

