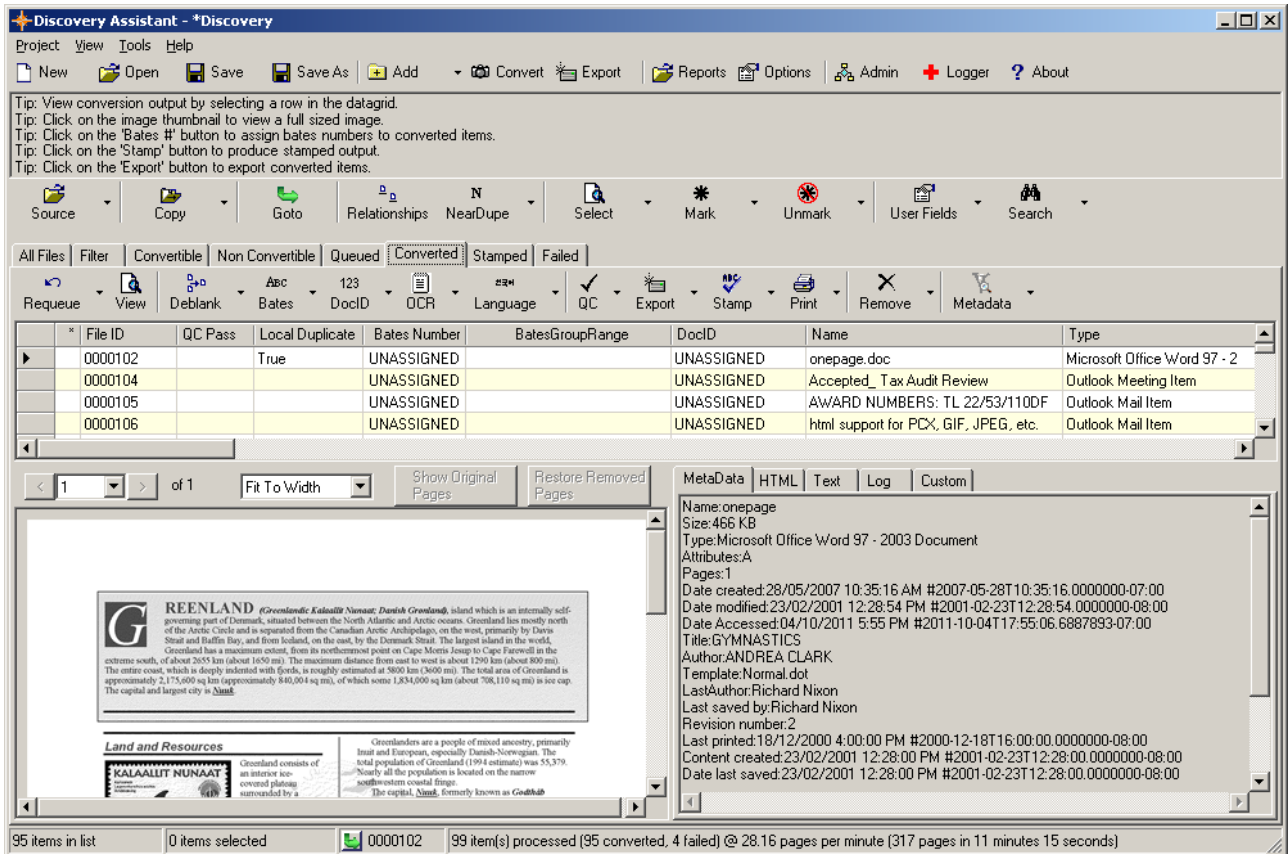# Introduction

On behalf of the entire Discovery Assistant team, Welcome!

If you are new to Discovery Assistant and would like to follow a step-by-step guide to its use, please see Getting Started.
For a more in-depth discussion of how to use Discovery Assistant, see Using Discovery Assistant.
For a brief overview of what Discovery Assistant does, see Overview.

Features

ImageMAKER Discovery Assistant is a powerful automation program designed to manage the conversion of electronic files to TIFF and/or PDF. The program combines a database front end capable of managing millions of documents with a automated document conversion back end, capable of converting tens of thousands of documents a day. Documents are automatically converted using the native application (Word, Excel, Acrobat, etc.) to TIFF or PDF.

The ImageMAKER Discovery Assistant application supports the following features:

Batch mode conversion to TIFF or PDF The ImageMAKER Electronic Discovery Conversion Server automatically converts documents in batch mode to TIFF Group 4 or Group 3 Black and White compressed file formats, as well as several colour TIFF formats such 24-bit JPEG, and PDF files. Document types that can be converted include: Microsoft email messages and attachments, Zip files, RTF files (Rich Text Formatted), Microsoft Word, Excel, PowerPoint, Visio and, HTML, Text, Adobe PDF, Word Perfect, Corel Draw, Cad Cam, as well as a host of other common file types.

Advanced Process Monitoring Discovery Assistant includes 'auto-close' and 'auto-kill' intelligence. This ensures that the conversion process runs un-attended for hundreds of thousands of conversions. If an application gets stuck for any reason while converting the document, a daemon application that monitors the process can be trained to look for and close dialogs and/or terminate the problem process. Conversion status is logged, and users can at any time review the status of converted documents.

Upgrade path to support multiple servers For large batch jobs that are too big for a single machine, we provide an upgrade path that allows multiple servers to attach to the same conversion database to help process jobs and improve throughput.

Electronic Mail (PST, MSG and ZIP files) support Electronic mail. E-mail attachments and Zip files are pre-processed when imported into the conversion database. The contents of the document are listed as separate documents, each of which can be converted separate from the other jobs. PST files - Microsoft Message Archival / Personal Folder File MSG Files - Microsoft Outlook message file Zip files – compressed document folders (pkzip).

Multiple output file formats Standard output file formats include TIFF Class F Group 4, TIFF Group 3, and PDF scanned image formats. Output resolutions include 204x198 dpi (faxable format), 200x200 dpi, and 300x300 dpi (scanner resolution). Output resolutions can be customized as needed. Additional document types include 24 bit color formats (TIFF , JPEG, and PDF), DCX, BMP, and PNG.

Extracted Text contents A formatted ASCII text stream is created for each converted file, separate from the image file. This text file can be indexed and searched separate from the image file.

Unique Identifier Assigned to each Document page (Bates Stamping) Automated sequential stamping of scanned or electronically converted TIFF images. Discovery Assistant places a sequential number onto every TIFF page generated. User can control placement of Bates stamps to any of the 4 corners of the page. Support for different fonts and font sizes. Customizable interface. Batch or single user mode.

Document Storage ImageMAKER products can be integrated with over 60 different electronic document management and content management systems, such as: DocuLex, Summation,

Concordance, iCONECT, Hummingbird and other prominent legal content management systems.

Custom Development ImageMAKER Custom development services are designed to insure rapid and seamless integration of our components and meet our customers' specific requirements. Our custom development services offering complements our customers' development team with integration advice and development expertise gained from over hundreds of integration projects in application areas including Litigation Automation, Document Management, Document Delivery & Storage Systems, E-Mail, CTI (Computer Telephony Integration), MFD (Multi-Functional Device), Fax, Unified Messaging, Embedded and Mobile applications.

Additional Features:

- Compatible with the leading email formats.
- Complete metadata and full-text extraction from virtually any file formats.
- Hundreds of file formats recognized.
- User selectable print control options for spreadsheets.
- Hard drive scan features to identify business documents.
- Quick check.
- Document tracking to monitor the conversion of millions of documents.
- De-Duplication.
- NearDuplicate detection

Benefits
With Discovery Assistant, electronic discovery and litigation software, clients gain unmatched benefits. With the capacity to manage millions of pages, Discovery Assistant offers organizations the ability to easily access, review and manage electronic documents.

- Review documents faster.
- High client satisfaction.
- Lower litigation risks.
- Improve productivity.
- Build attorney-client relationship through increased collaboration.
- Reduce cost and save time associated with electronic discovery.
- Preserve a digital chain of evidence.
- Control your electronic documents.

Overview

ImageMAKER Discovery Assistant is a powerful automation program designed to manage the conversion of electronic files to TIFF and/or PDF. The program combines a database front end capable of managing millions of documents with an automated scaleable document conversion engine back-end. Documents are automatically converted using the native application (Word, Excel, Acrobat, etc.) to TIFF or PDF.

The process of preparing documents for Electronic Discovery requires that large collections of documents, in a variety of different formats, be converted to a common image format for inclusion within large legal databases. Source document types include: e-mail message files and their attachments; Microsoft Office documents such as Word, Excel, HTML, PowerPoint; and other standard document types including Acrobat PDF, Corel Word Perfect, Corel Draw, and AutoCad.

Once documents have been converted to TIFF, they are then Bates Stamped, indexed, imported into large legal databases for additional processing, then delivered into discovery. Common legal database programs include Summation and Concordance.

Because most documents are created and stored electronically, it is more cost effective to convert documents in-place, using document conversion software, than it is to print the documents to paper, then scan the images back in. Image quality is better, speed is better, there are fewer errors, and the overall cost can be significantly less.

Automated Conversion of Electronic Documents to TIFF and PDF

Source Documents:
DOC, PDF, PST, HTML, MSG,
WPG, DGF, PPT, TXT, XLS, JPEG,
TIFF etc

Output Documents:
Tiff or PDF Image
MetaData
Extracted Searchable Text

Ready for Importation into Summation, Concordance or other industry standard Document Storage Systems.

Electronic Discovery Process:

Electronic Discovery is the process of gathering, reviewing, and managing documents originating from computers. Electronic Discovery is important in the management of large evidentiary collections.

The process of Electronic Discovery can be defined as follows:

1. The collection of documents from Electronic Document Repositories (electronically stored documents like word processing documents, e-mails and electronic spread sheets), i.e., the collection of data from computers, including network servers, desktop PCs, laptops, backup tapes, handheld devices, and any other storage media.

2. Importation of the documents into a managed database, ready for conversion. Documents can be sorted by file type, size, date, whether or not they can be converted, and point of origin. Email files and Zip files are expanded in-place to allow full enumeration of available documents.
3. The conversion of electronic documents to a single industry standard file format (TIFF or PDF). Output files include the image file, a meta-data file, and extracted TXT contents (for searchable indexing).
4. Creation of a unique identity (Bates Stamp) for each converted document. The txt files can then be fully indexed for search purposes.
5. Managing the exportation of the list of converted documents for inclusion back into legal databases.

Electronic Discovery and Email:
Every day, more than 10 billion new email messages zip through the Internet and across corporate networks. Attached to many of these email messages are electronic documents. These attachments are created with a variety of software programs in different file formats. The result is a staggering volume and variety of incompatible electronic documents, all of which are potentially discoverable.

Electronic Discovery and You:
Attorneys use Electronic Discovery services in litigation and antitrust work, as well as in other situations when electronic documents must be reviewed and produced for use by another party. Typically, the cost of providing the Electronic Discovery service is an up-front fee, and a per-copy fee.

Electronic Discovery Services use a variety of tools to manage the conversion of documents from a wide variety of sources to TIFF or PDF. Service organizations will either buy an off the shelf product (such as the ImageMAKER Discovery Assistant), or build their own tools from library components (print drivers, conversion server utilities, database development environments).

Attorneys are increasingly looking to manage the process of electronic discovery in-house - either as a way to control costs, or as a way of generating extra income. The ImageMAKER Discovery Assistant product is one way to bring the required conversion expertise in-house.

Getting Started

This section will provide you with a step-by-step guide to using DiscoveryAssistant. You can follow these steps in order or jump to whichever section suits your needs.

Features:

- [Creating a New Project](#)
- [Adding Files to A Project](#)
- [Verifying Files](#)
- [Queuing Files](#)
- [Converting Files](#)
- [Viewing Converted Files](#)
- [Viewing Failed Files](#)
- [Stamping Converted Files](#)
- [Viewing Stamped Files](#)
- [Exporting Results](#)

Adding Files to a Project

There are four ways to add files to a Discovery Assistant project:

- Select Files
- Scan a Folder
- File List
- Drag and Drop from Windows Explorer

The content of all folders, subfolders, zip files, MSG files, and PST files are added to a Discovery Assistant project automatically when one of these objects is scanned into or dropped onto Discovery Assistant. For ways to limit the files that are added during a scan or Drag and Drop operation, see Project Options.

Select Files

To add files from an Open File Dialog, click the Add button in the All Files Toolbar and then click the Files... button from the drop-down menu.
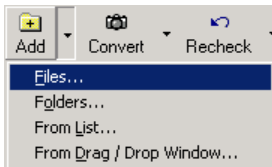
An open file dialog will appear. Select the files you want to add - you can use the Shift and Ctrl keys to select multiple files - and then click the Open button to add the files to Discovery Asistant.

Scan a Folder

To add files by scanning a folder, simply click the Add button in the All Files Toolbar and then click the Folders... button from the drop-down menu.
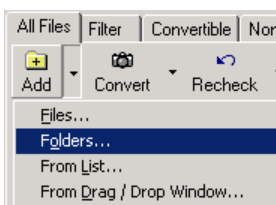
You will then be presented with a dialog to choose the folder to scan.
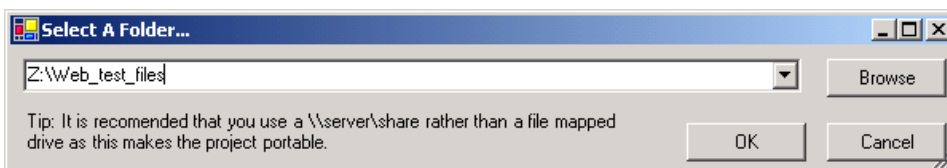
All files in the chosen folder and any subfolders will be added to the current project.

File List

To add files from a file list, simply click the Add button in the All Files Toolbar and then click the Folders... button from the drop-down menu.
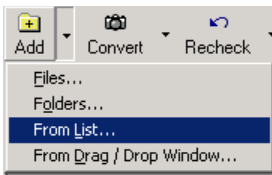
An Open File Dialog will appear. Select the file list and Discovery Assistant will add the files in the list. File lists should be text files with each file path on its own line.

Drag and Drop from Windows Explorer

To use this method to add files to a Discovery Assistant project, simply drag any number of files or folders from any Explorer window and drop them on the All Files Data Grid of Discovery Assistant.
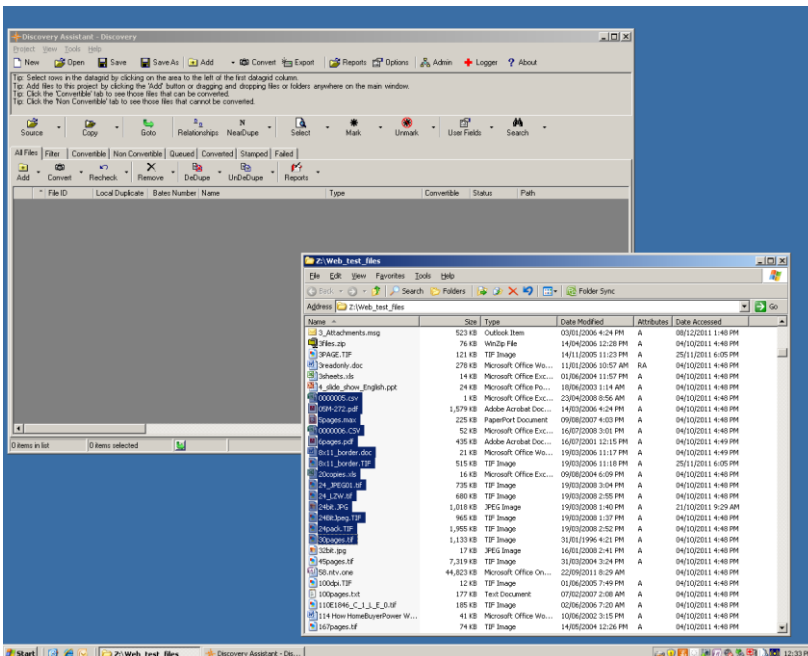
Clicking the Add button in the All Files Toolbar and then the From Drag / Drop Window... button from drop-down menu will open an Explorer window for the user to drag and drop files.
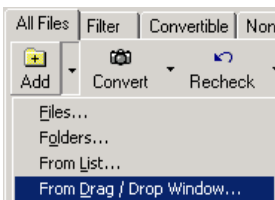
For ways to limit the files that are added, see Project Options.

The next step is to Verify the files that were added to the project.

Creating a New Project

When you first start Discovery Assistant, you will be presented with the Open Project dialog. If there are no recently opened projects, you will start in the "New Project" tab. Type a name in the Project Name field and click Open.
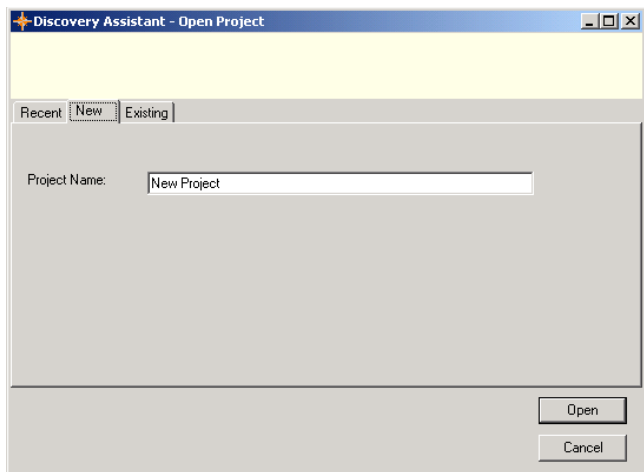
You will then be prompted to save the project with the Save As dialog.
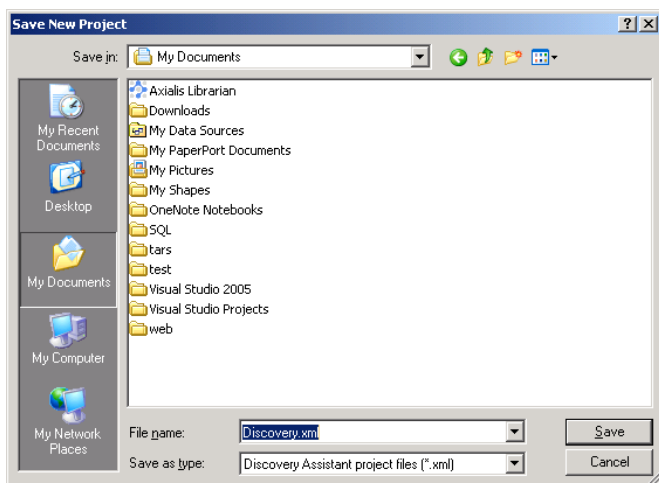
Choose the folder that you want the project to be saved in and press the Save button. Discovery Assistant will create several subfolders in addition to the XML project file. These subfolders are <projectName>.tmp, <projectName>.cnvt, <projectName>.stmp, <projectName>.rpt, <projectName>.cust, <projectName>.ND and <projectName>.sav. They contain temporary files, converted files, stamped files, report files, custom metadata files, near duplicate related files, and saved files (e.g., from before deblanking), respectively. Typically, you do not need to be concerned with the contents of these folders as all viewing and exporting of converted files is handled for you through the Discovery Assistant interface.

After the project has been saved, you will be presented with the main Discovery Assistant interface.

The next step is to Add Files to the project.

Verifying Files

Now that you've added files to your Discovery Assistant project, you will want to remove duplicate files. The DeDupe button in the All Files tab will flag duplicate files. These flagged files will not be furthered processed by Discovery Assistant. To undo this deduplication, press the UnDeDupe button in the All Files tab.
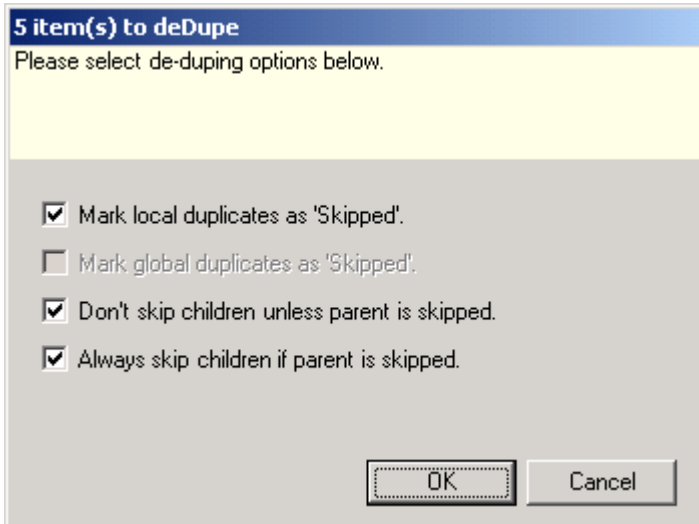
The All Files Data Grid shows all the items that you have added. However, the Convertible column of the Data Grid shows if a given item is convertible. Some items such as Zip files and PST files are considered Containers and are not themselves convertible although their contents are.
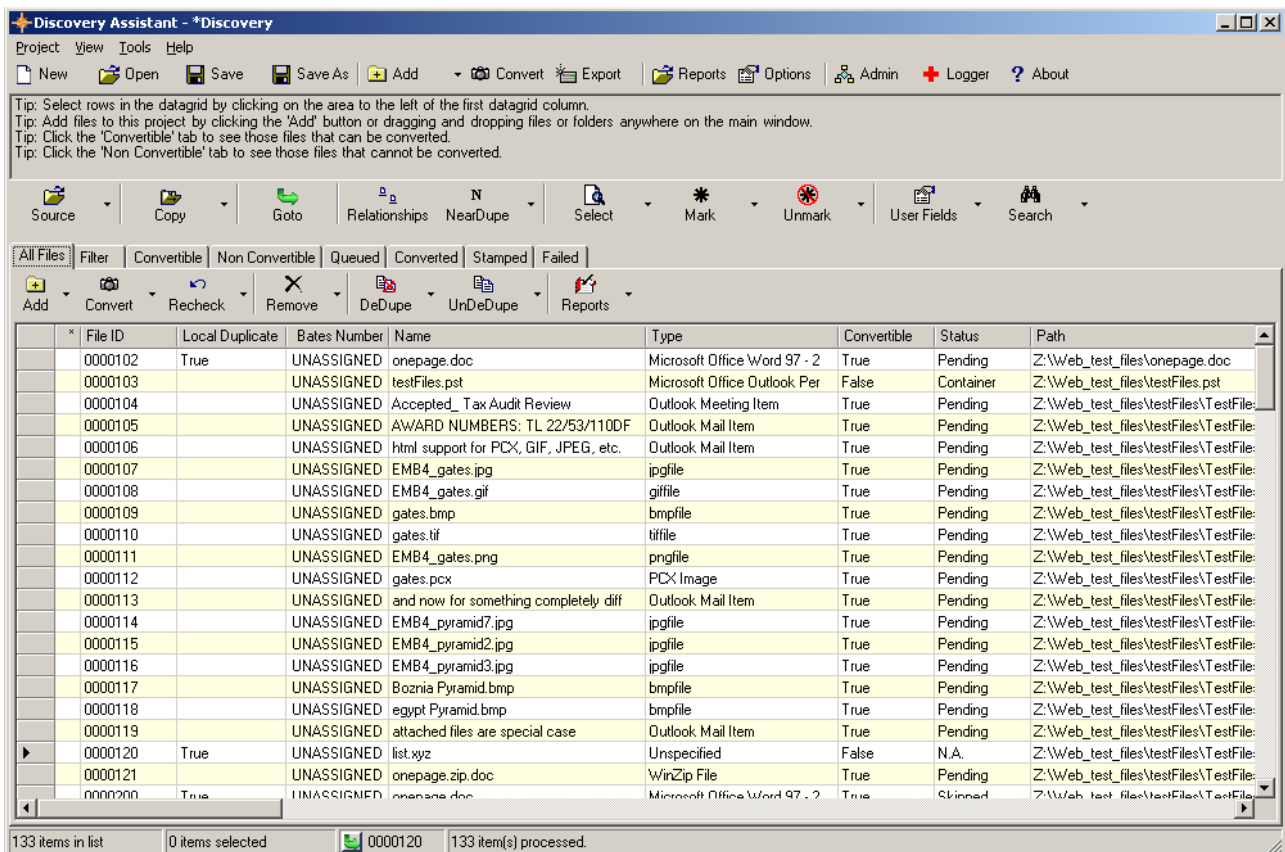
For more information on what you can do in the All Files tab see All Files Tab.

Convertible Files Tab

To view only those items that are considered to be convertible, click on the Convertible tab.
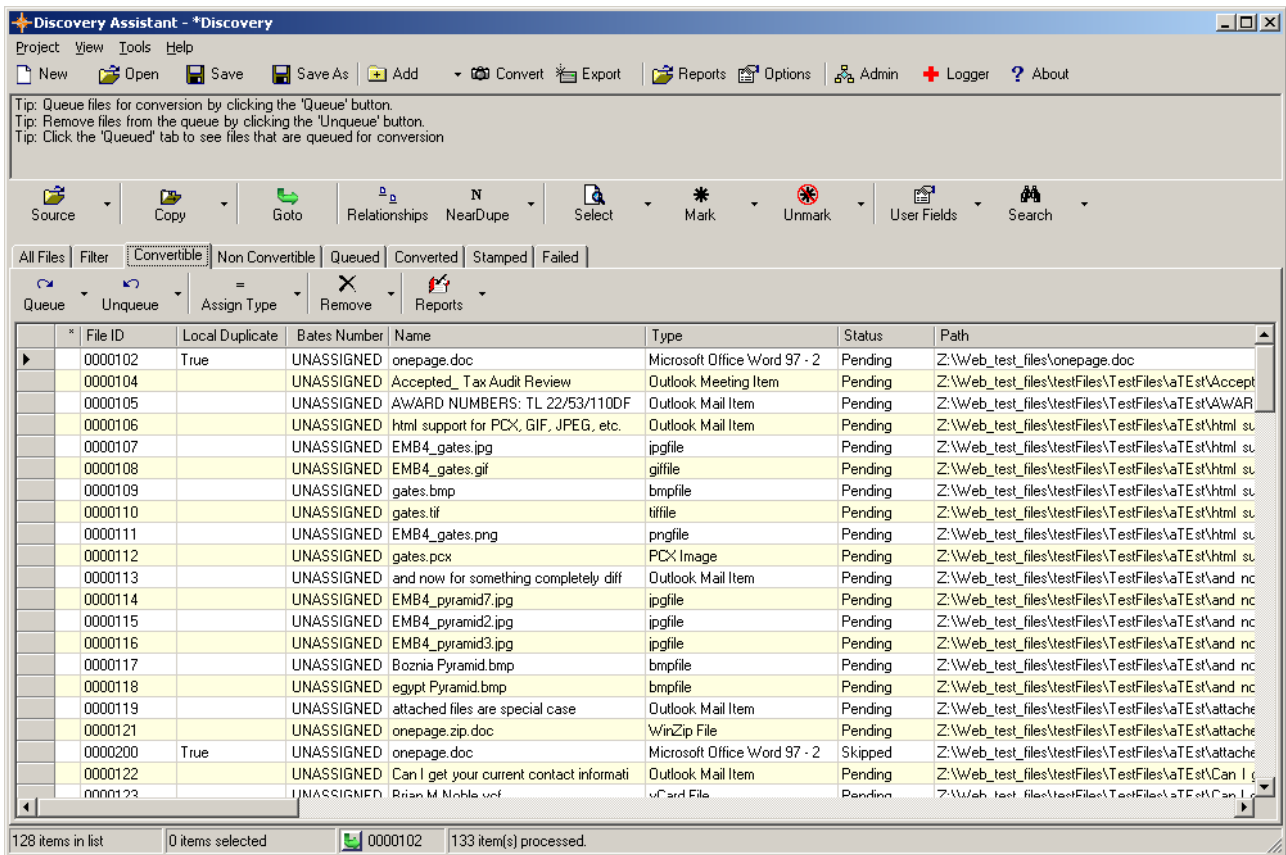
For more information, see the Convertible Tab.

Non Convertible Tab

To view those items that are not convertible, click on the Non Convertible tab.
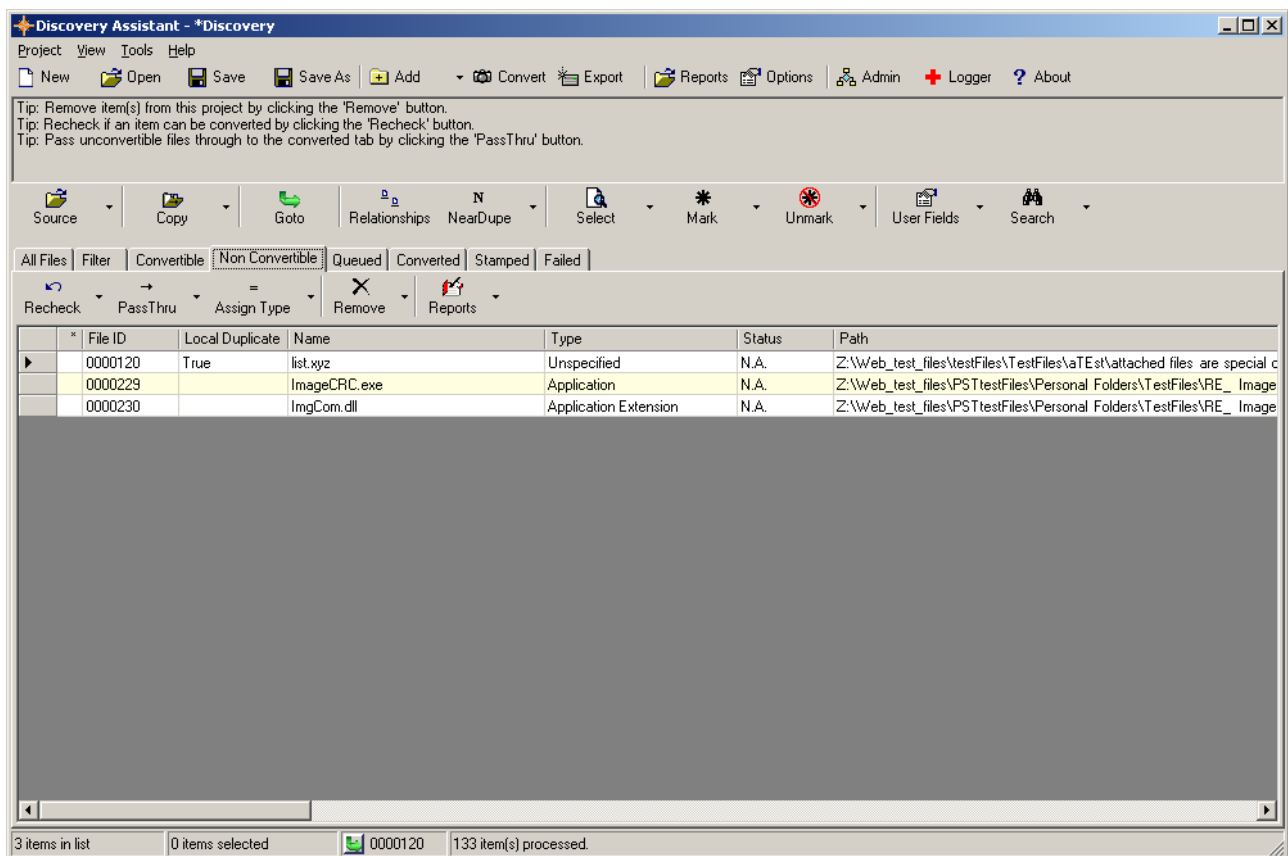
Note that although Zip and PST files are considered non-convertible, they do not appear in the Non Convertible tab because their content is convertible.

If there are items that appear as non-convertible that you think should be convertible, such as Word documents or Excel spreadsheets, then it is probable that you do not have the required application installed to print these files (in this case Word or Excel). Install the required programs and then click the Recheck button.
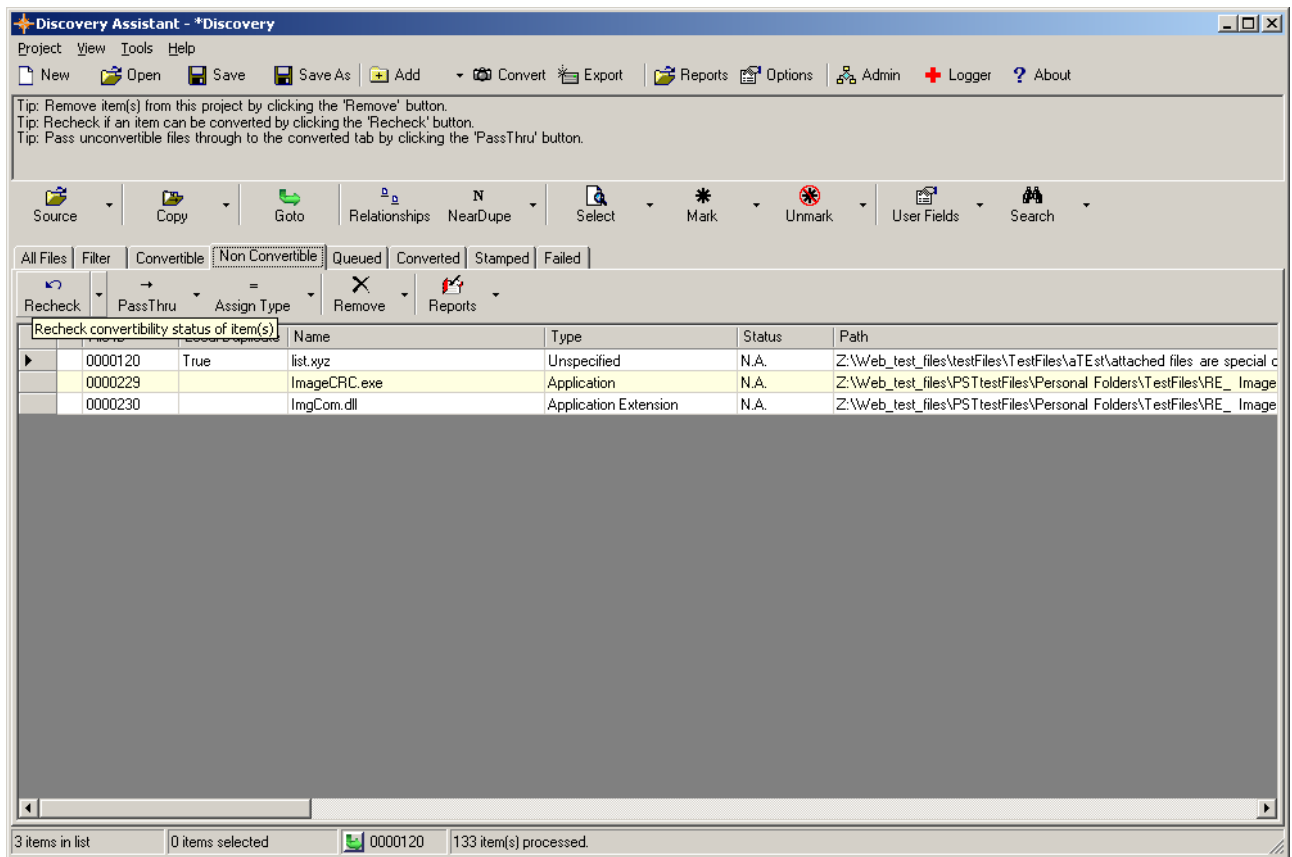
For more information, see the Non Convertible Tab.

The next step is to Queue files for conversion.

Queueing Files

To queue files for conversion, simply click on the Queue button in the Convertible Tab.
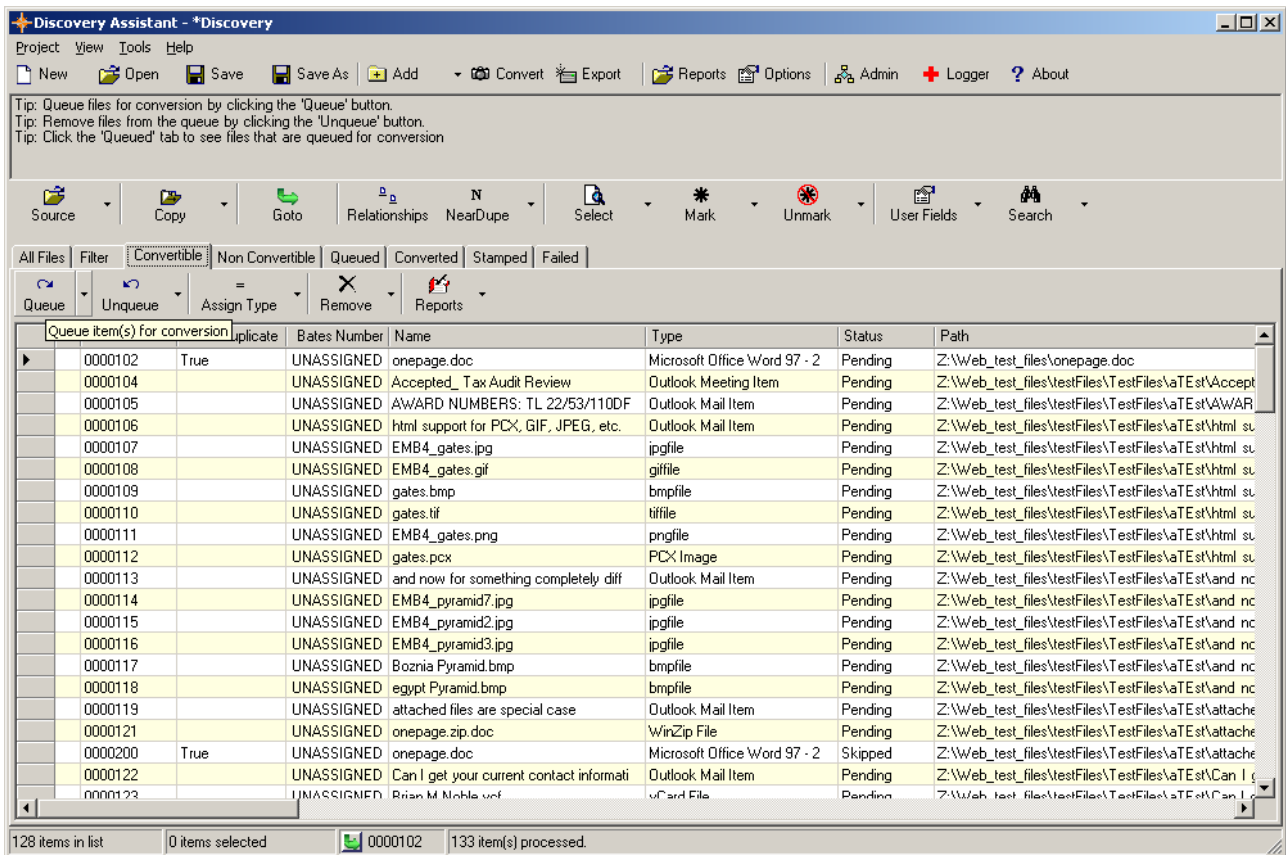
You can select specific items to queue by clicking to the left of the first column in the Data Grid before clicking the Queue button.

When items have been queued for conversion, their status changes from Pending to Queued.
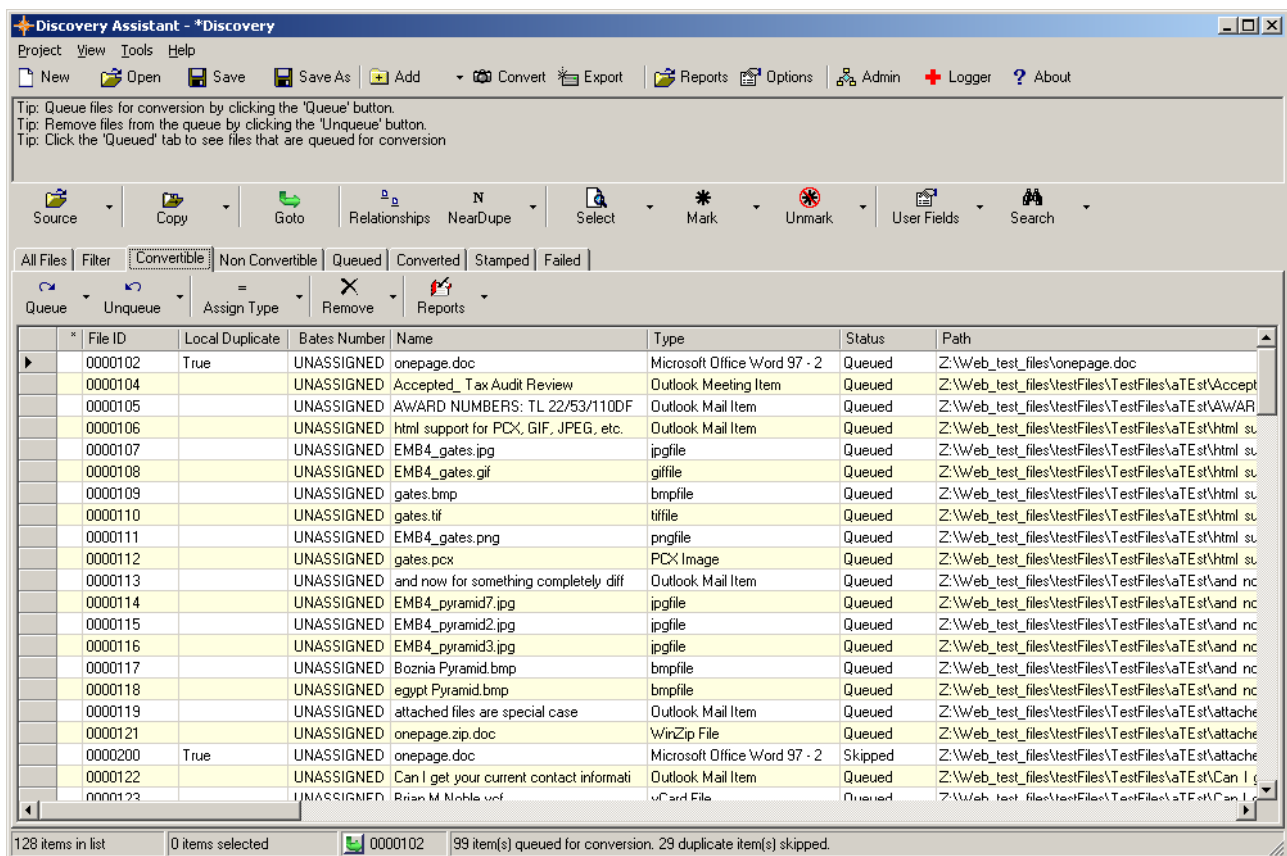
The next step is to [Convert](#) these queued items.

Converting Files

To Convert queued files, simply press the Convert button in the Queued Tab.

You can select specific items to convert by clicking to the left of the first column in the Data Grid before clicking the Convert button.

When the conversion process starts a progress bar will appear indicating which item is being converted and the various stages of the conversion process.

Discovery Assistant - *Discovery

Project View Tools Help

New | Open | Save | Save As | Add | Convert | Export | Reports | Options | Admin | Logger | About

[0000163] esize.xls   Init Job

Pause

Source | Copy | Goto | Relationships NearDupe | Select | Mark | Unmark | User Fields | Search

All Files | Filter | Convertible | Non Convertible | Queued | Converted | Stamped | Failed

Convert | PassThru | Remove

*	File ID	Local Duplicate	Name	Type	Path
►	0000163	True	esize.xls	Microsoft Office Excel 97-20	Z:\Web_test_files\testFiles\TestFiles\aTEst\Test Files XLS, TIF, DOC, PDF\esize.x
	0000164	True	font.pdf	Adobe Acrobat Document	Z:\Web_test_files\testFiles\TestFiles\aTEst\Test Files XLS, TIF, DOC, PDF\font.pd
	0000165	True	genoa.ppt	Microsoft Office PowerPoint	Z:\Web_test_files\testFiles\TestFiles\aTEst\Test Files XLS, TIF, DOC, PDF\genoa.
	0000166	True	letter_legal_a4.doc	Microsoft Office Word 97 - 2	Z:\Web_test_files\testFiles\TestFiles\aTEst\Test Files XLS, TIF, DOC, PDF\letter_
	0000167	True	letter_legal.TIF	tiffile	Z:\Web_test_files\testFiles\TestFiles\aTEst\Test Files XLS, TIF, DOC, PDF\letter_
	0000168	True	letter_legal.pdf	Adobe Acrobat Document	Z:\Web_test_files\testFiles\TestFiles\aTEst\Test Files XLS, TIF, DOC, PDF\letter_
	0000169		FW: 6Sigma Story	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\FW_ 6Sigma Story.msg
	0000170		inv540241	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\inv540241.msg
	0000171		inv54024.tif	tiffile	Z:\Web_test_files\testFiles\TestFiles\aTEst\inv540241\inv54024.tif
	0000172		Untitled	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\Untitled.msg
	0000185		FW: method example	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\FW_ method example.msg
	0000188		RE: method example	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\RE_ method example.msg
	0000190		Re: can you send me the sample MSG	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\Re_ can you send me the sample MSG
	0000191		FW Mid Year Operations Meeting.msg	Outlook Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\Re_ can you send me the sample MSG
	0000201		EMB4_RonDeal.png	pngfile	Z:\Web_test_files\testFiles\TestFiles\aTEst\Re_ can you send me the sample MSG
	0000202		EMB4_image002.jpg	jpgfile	Z:\Web_test_files\testFiles\TestFiles\aTEst\Re_ can you send me the sample MSG
	0000192		test with XLS in the middle.	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\test with XLS in the middle..msg
	0000193		3sheets.xls	Microsoft Office Excel 97-20	Z:\Web_test_files\testFiles\TestFiles\aTEst\test with XLS in the middle.\3sheets.xls
	0000194		FW: Preuss Pharmacy-CPM	Outlook Mail Item	Z:\Web_test_files\testFiles\TestFiles\aTEst\FW_ Preuss Pharmacy-CPM.msg
	0000195		Preuss Pharmacy COG table.xls	Microsoft Office Excel 97-20	Z:\Web_test_files\testFiles\TestFiles\aTEst\FW_ Preuss Pharmacy-CPM\Preuss F
	0000196		Preuss Pharmacy 5years.xls	Microsoft Office Excel 97-20	Z:\Web_test_files\testFiles\TestFiles\aTEst\FW_ Preuss Pharmacy-CPM\Preuss F

42 items in list | 0 items selected | 0000163 | 57 item(s) processed (57 converted, 0 failed) @ 27.95 pages per minute (130 pages in 4 minutes 39 seconds)

As conversion is completed for each item in the Queued Data Grid, its status changes from Queued to either Converted or Failed and it is removed from the Queued Data Grid. Finally, when the entire conversion process is complete, the Queued Data Grid will be empty (unless you selected only specific items to be converted) and the progress bar will be removed.

Discovery Assistant - *Discovery

Project View Tools Help

New | Open | Save | Save As | Add | Convert | Export | Reports | Options | Admin | Logger | About

Tip: Convert queued files by clicking the 'Convert' button.
Tip: Click the 'Converted' tab to see files that converted successfully or the 'Failed' tab to see files that failed in conversion.

Source | Copy | Goto | Relationships NearDupe | Select | Mark | Unmark | User Fields | Search

All Files | Filter | Convertible | Non Convertible | Queued | Converted | Stamped | Failed

Convert | PassThru | Remove

| *	File ID	Local Duplicate	Name	Type	Path

0 items in list | 0 items selected | | 99 item(s) processed (95 converted, 4 failed) @ 28.16 pages per minute (317 pages in 11 minutes 15 seconds)

17

Also note that the status bar at the bottom of the Discovery Assistant window will give you a running total of the items converted, how many succeeded, how many failed, and the average Page Per Minute rate of conversion.

Click here for more information on the Queued Tab.

The next step is to view Converted and Failed items.

Viewing Converted Files

To view converted items, click on the Converted Tab.

Click to the left of the first column in the Data Grid to view a specific item.

Click here for more information on the Converted Tab.

Your next step will either be to Stamp the converted items or, if you aren't applying stamps, then directly to Export the results. You can also send your converted files to the printer of your choice by clicking on the Print button in the Converted Toolbar.

Viewing Failed Files

If any files in your Discovery Assistant project fail conversion, you can check the reason by going to the Failed Tab.
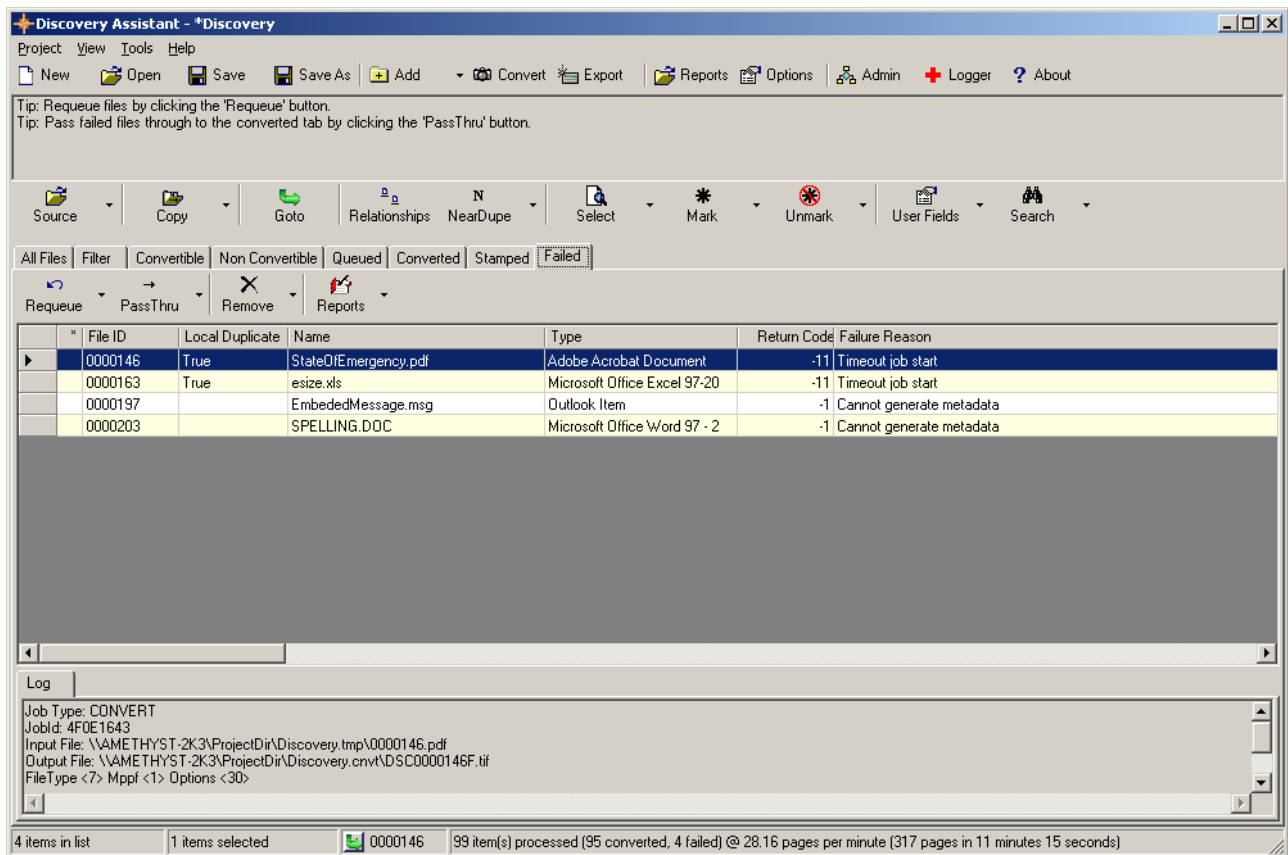
Each item that failed conversion is listed in the Failed Data Grid along with a return code and failure reason. You can choose to requeue the failed item(s) and try again (perhaps after rectifying the problem) or you can remove them from your project.

Click here for more information on the Failed Tab.

Stamping Converted Files

You can apply stamps such as Bates numbers to your converted files easily with Discovery Assistant. Simply click on the Stamp button in the Converted Toolbar.

Enter the requested details and click Ok. A progress bar will show as the files are stamped.

[0000164] font.pdf

Pause

All Files | Filter | Convertible | Non Convertible | Queued | Converted | Stamped | Failed

	*	File ID	QC Pass	Local Duplicate	Bates Number	BatesGroupRange	DocID	Name	Type
►		0000102		True	DSC00001		DOC_0000001	onepage.doc	Microsoft Office Word 97 - 2
		0000104			DSC00002		DOC_0000002	Accepted_ Tax Audit Review	Outlook Meeting Item
		0000105			DSC00003		DOC_0000003	AWARD NUMBERS: TL 22/53/110DF	Outlook Mail Item
		0000106			DSC00005	DSC00005-DSC00014	DOC_0000004	html support for PCX, GIF, JPEG, etc.	Outlook Mail Item

MetaData | HTML | Text | Log | Custom

Name:onepage
Size:466 KB
Type:Microsoft Office Word 97 - 2003 Document
Attributes:A
Pages:1
Date created:28/05/2007 10:35:16 AM #2007-05-28T10:35:16.0000000-07:00
Date modified:23/02/2001 12:28:54 PM #2001-02-23T12:28:54.0000000-08:00
Date Accessed:04/10/2011 5:55 PM #2011-10-04T17:55:06.6887893-07:00
Title:GYMNASTICS
Author:ANDREA CLARK
Template:Normal.dot
LastAuthor:Richard Nixon
Last saved by:Richard Nixon
Revision number:2
Last printed:18/12/2000 4:00:00 PM #2000-12-18T16:00:00.0000000-08:00
Content created:23/02/2001 12:28:00 PM #2001-02-23T12:28:00.0000000-08:00
Date last saved:23/02/2001 12:28:00 PM #2001-02-23T12:28:00.0000000-08:00

95 items in list   0 items selected   0000102   57 of 95 items processed.

For more information on defining the position and content of stamps, see Project Options.

The next step is the view the Stamped items.

Viewing Stamped Files

To view stamped items, click on the Stamped Tab.

Click here for more information on the Stamped Tab.

Your next step may be to Export the results. You can also send your stamped files to the printer of your choice by clicking on the Print button in the Stamped Toolbar.
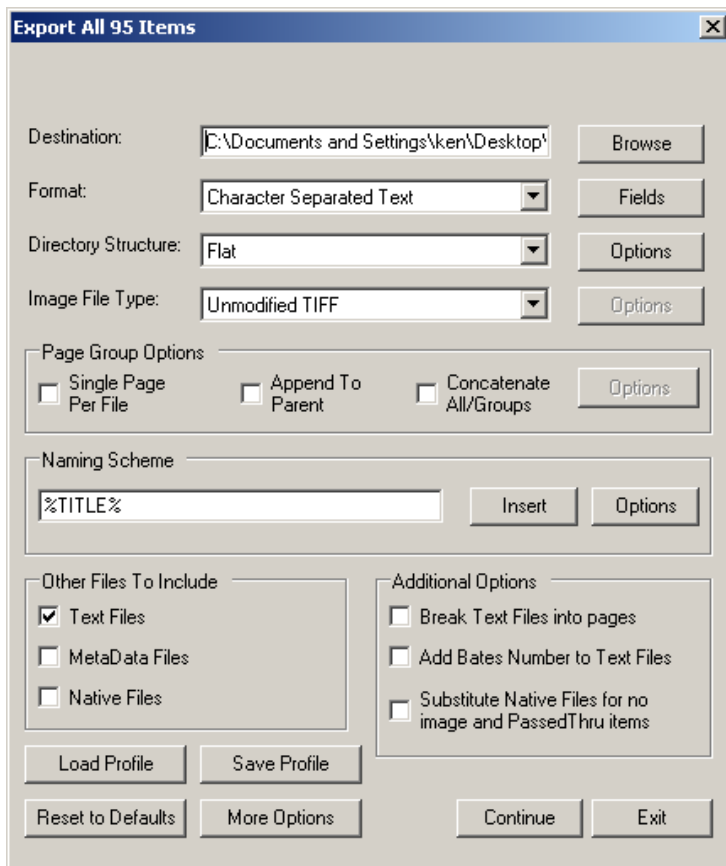
Exporting Results

To export converted or stamped items, simply press the Export button in the main Toolbar.

A dialog will appear allowing you to select Export options.

Click on the Browse button to select the folder to output results to. Click here for more information on Export options.

Click OK and the Export process will begin. A progress bar will appear showing which item is currently being exported. When the progress bar disappears, the export process is complete.

Using Discovery Assistant

This section provides an in-depth discussion of the many features of Discovery Assistant and how to use them.

- [The Tabbed Interface](#)
- [Project Options](#)
- [Tips and Tricks](#)
- [Using Admin](#)

The Tabbed Interface
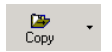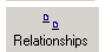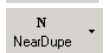
Discovery Assistant uses a tabbed interface to organize and guide you through the steps of E-Discovery. Each tab has a toolbar that makes the commands applicable to that stage of the process easy to access.

- All Files Tab
- Convertible Files Tab
- Non Convertible Files Tab
- Queued Files Tab
- Converted Files Tab
- Stamped Files Tab
- Failed Files Tab

Universal Toolbar

When you have any of the tabs open in Disscover Assistant, there is a toolbar above the tab selection. This toolbar has many universal options.

Opens the selected item with the application associated with it.

Allows you to copy selected source files to another folder on disk.

???

Brings up a dialog which displays duplicates, children and ancestors of the selected file.

Brings up a dialog for finding near duplicates and e-mail threads.

Provides multiple selection optioss, including selecting multiple rows based on a predefined file with the desired File ID's, DocID's or starting Bates Numbers separated by new lines. Some examples for the select file are (for File ID):
001
004
040
And (for DocID):
DOCID_01237
DOCID_01240
DOCID_01251

Marks the selected files.

Unmarks the selected files.

Brings up a dialog to manage the User Fileds.

Indexes and searches files in the All Files or current tab.

All Files Tab

The All Files tab lists all of the files in the open project.
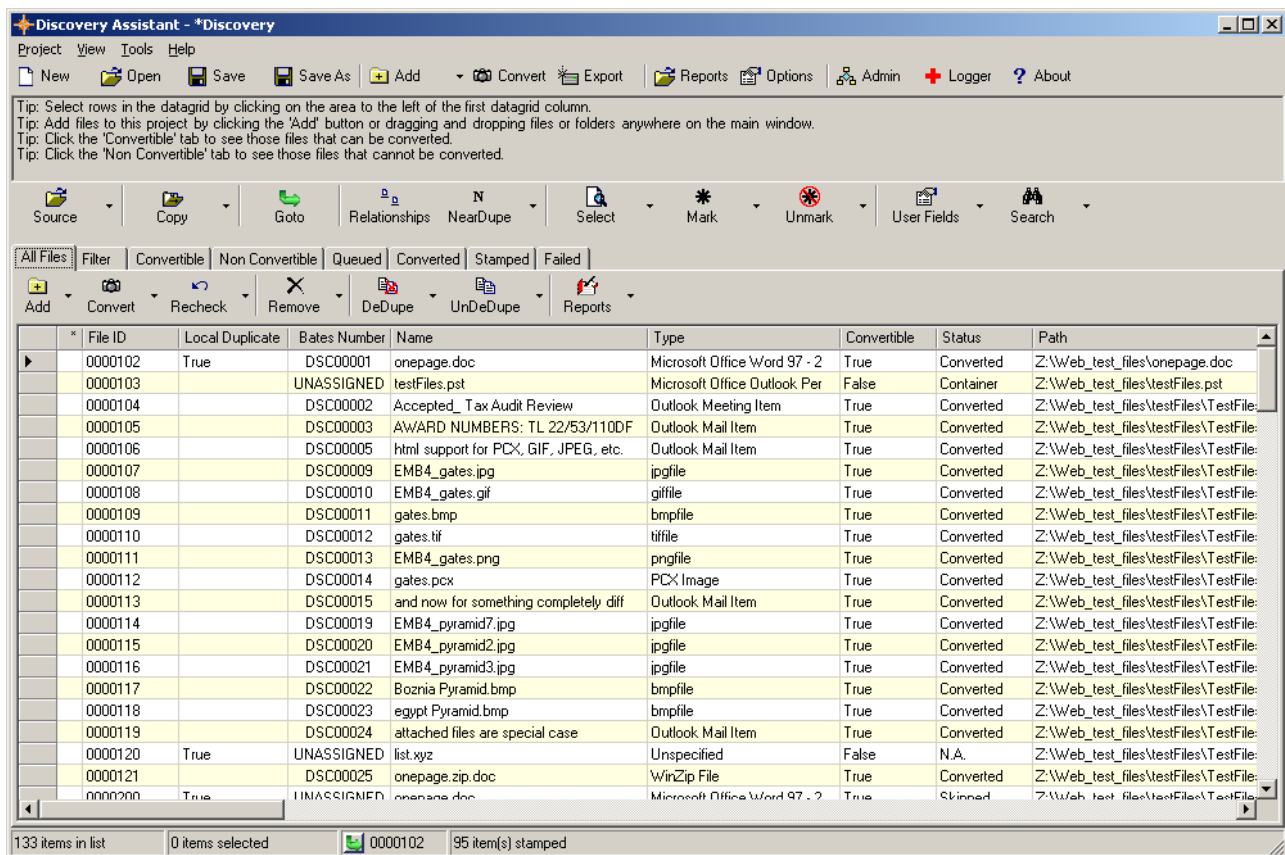
Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will contain the word "True".
- Bates Number - If the item has been Bates stamped, this column will contain the starting Bates number.
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Convertible - This column indicates whether the item is deemed convertible. For an item to be convertible in must have a "Print" or "PrintTo" association defined. Container items such as PST or ZIP files are marked as non-convertible even though they may contain convertible items.
- Status
  - N.A. - indicates that the item is not convertible and is not a container of other items.
  - Container - indicates that the item is a PST file. It is not, itself, convertible but may contain items that are.
  - Skipped - item is a duplicate of another convertible item that does not have a status of Skipped
  - Pending - item is convertible but has not been queued or converted yet.
  - PassedThru - item has been passed through the converted tab.

- o Queued - item has been queued for conversion.
- o Converted - item has been converted successfully.
- o Failed - item failed conversion or failed to be extracted from its parent file.
- o ERROR - an error occurred extracting files from this file (e.g., PST, ZIP, etc.)
- Path - The full path to the item. For items contained in PST, ZIP, MSG, EML, NSF, RAR, PDF and Office files, this path will indicate the parent container.
- Hash Code - The MD5 hash code for the item.
- Modify Date - The last modified date of the item.
- Size - Size, in bytes, of the item on the disk.
- Media Label - The media label of the disk containing the item.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

Adds files to the project. (See Adding Files to a Project).

Automatically queues and converts the selected items (or all items if no items are selected).

Causes the selected items (or all items if no items are selected) to be rechecked for convertibility and their hash codes to be regenerated. This is useful if new software has been installed that defines new Print or PrintTo associations.

Removes the selected items (or all items if no items are selected) from the project. It does not delete them from disk.

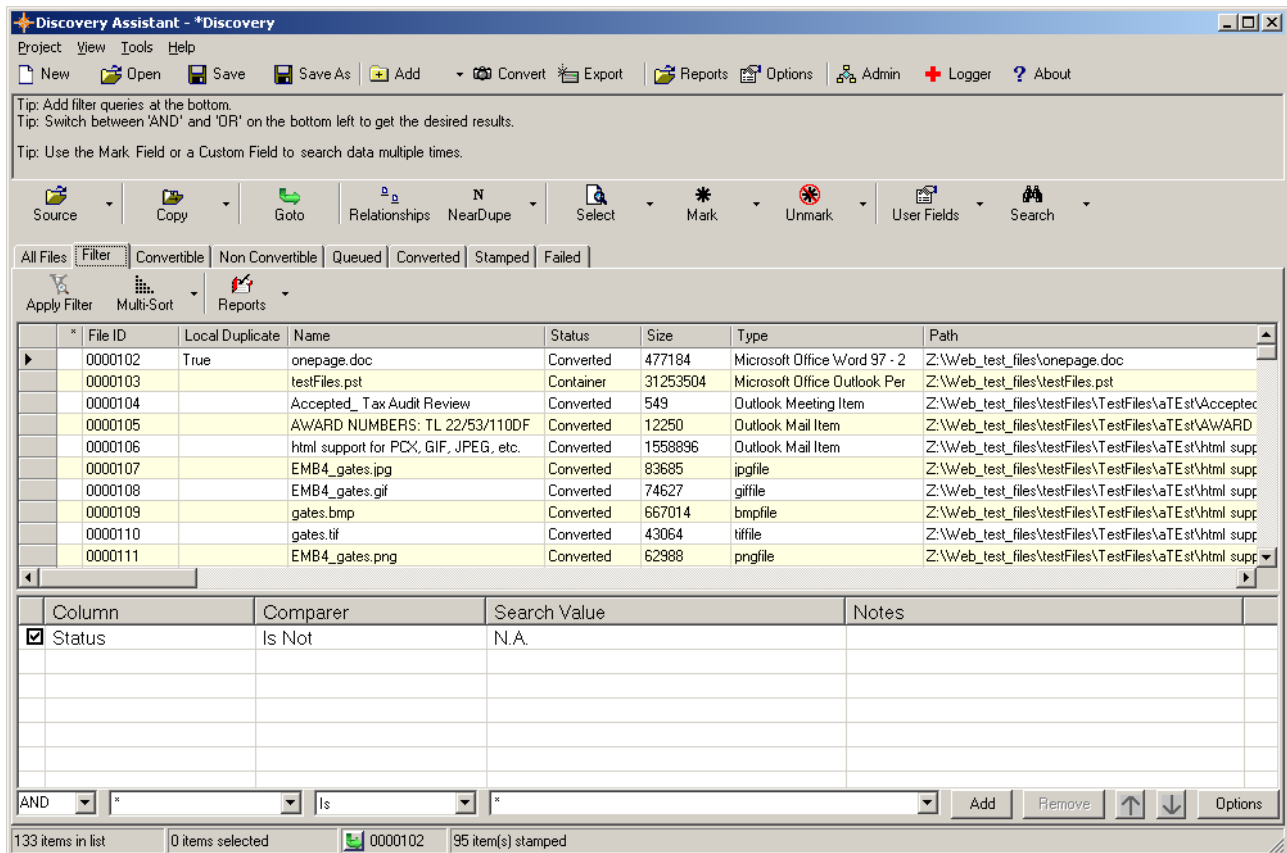Finds and flags duplicate files among the selected items (or all items if no items are selected) in the project.

Undoes the Depulication process among the selected items (or all items if no items are selected).

Creates a Detail or Summary Report in a CSV file.

Filter Tab

The Filter Files Tab can be used to custom filter DA project items in your project.

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will say "True", else it will be blank.
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Status
  - N.A. - indicates that the item is not convertible and is not a container of other items.
  - Container - indicates that the item is a PST file. It is not, itself, convertible but may contain items that are.
  - Skipped - item is a duplicate of another convertible item that does not have a status of Skipped
  - Pending - item is convertible but has not been queued or converted yet.
  - PassedThru - item has been passed through the converted tab.
  - Queued - item has been queued for conversion.
  - Converted - item has been converted successfully.
  - Failed - item failed conversion or failed to be extracted from its parent file.
  - ERROR - an error occurred extracting files from this file (e.g., PST, ZIP, etc.)
- Size - Size, in bytes, of the item on the disk.
- Type - This column contains the file type description as provided by the operating system.

- Path - The full path to the item. For items contained in PST, ZIP, MSG, EML, NSF, RAR, PDF and Office files, this path will indicate the parent container.
- Modify Date - The last modified date and time of the item.
- Output File - Full path to the output file.
- Pages - The number of pages in the output of the current item.
- Embedded Files - "True" if the item is an Office document or PDF file that has embedded files, else blank.
- Near Parent ID - If this is a child, the item ID of the immediate parent, else blank.
- Return Code - Return code from the conversion engine, with a negative value indicating that an error occurred.
- Failure Reason - Reason for conversion failure.
- Stamped File - Full path to the stamped output file.
- Output Type - Output file type.
- Hash Code - The MD5 hash code for the item.
- Convert Time - If the item has been successfully converted, the date/time of the conversion, else "Unavailable".
- Removed Pages - If deblanking was done on the file, the number of pages that were removed, else "unknown".
- DocID - The document ID assigned to the item, else "UNASSIGNED".
- Bates Number - If the item has been Bates stamped, this column will contain the starting Bates number, else "UNASSIGNED".
- Bates Group Range - If Bates numbers have been assigned, for items that are parents with one or more children items, this contains the range of Bates numbers from the first page of the top-level parent to the last child's page. Otherwise, this contains blanks.
- QC Pass - If the QC (Quality Control) has passed, it shows "True", else if QC has failed, it shows "False", else blank is shown if QC has not yet been on that item.
- Media Label - The media label of the disk containing the item.
- OCR'ed - "True" if the item has had OCR (Optical Character Recognition) applied to it, else blank.
- Language - If the Language option .
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

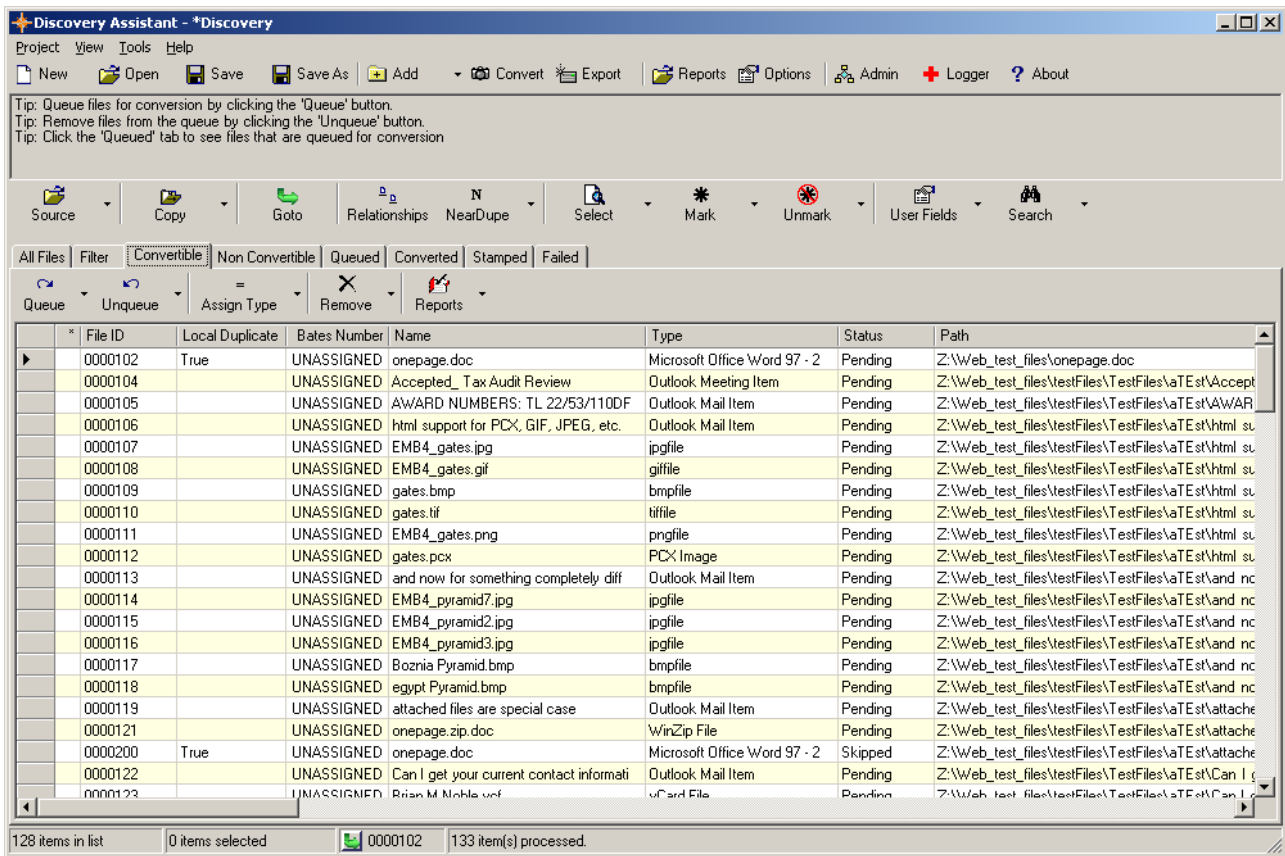 Applies or removes the specified filter of the project items.

   Sorts the items in this tab by multiple columns.

   Creates a Detail or Summary Report in a CSV file.

Convertable Tab

The Convertible Files Tab lists all files in the open project that are considered convertible.
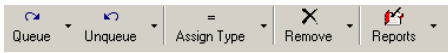

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will say "True", else it will be blank.
- Bates Number - If the item has been Bates stamped, this column will contain the starting Bates number, else "UNASSIGNED".
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Status
  - N.A. - indicates that the item is not convertible and is not a container of other items.
  - Container - indicates that the item is a PST file. It is not, itself, convertible but may contain items that are.
  - Skipped - item is a duplicate of another convertible item that does not have a status of Skipped
  - Pending - item is convertible but has not been queued or converted yet.
  - PassedThru - item has been passed through the converted tab.
  - Queued - item has been queued for conversion.
  - Converted - item has been converted successfully.
  - Failed - item failed conversion or failed to be extracted from its parent file.

34

- ○ ERROR - an error occurred extracting files from this file (e.g., PST, ZIP, etc.)
- Path - The full path to the item. For items contained in PST, ZIP, MSG, EML, NSF, RAR, PDF and Office files, this path will indicate the parent container.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

Queues the selected items (or all items if no items are selected) for conversion. If the drop-down menu option of Test is selected, then only the first instance of each file type is queued.

Unqueues the selected items (or all items if no items are selected).

Allows you to assign a file type to selected items (or all items if no items are selected).
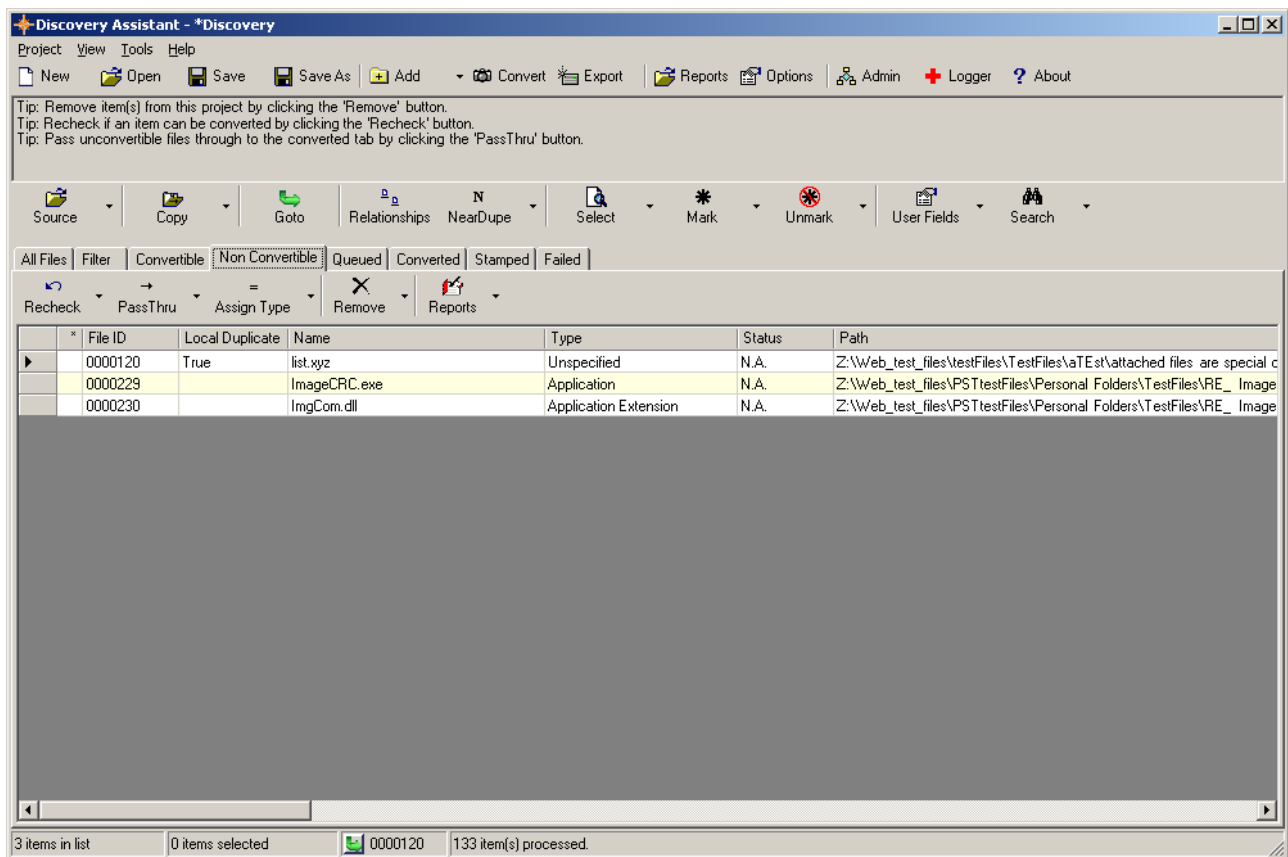
Removes the selected items (or all items if no items are selected) from the project. It does not delete them from disk.

Creates a Detail or Summary Report in a CSV file.

Non Convertible Tab

The Non-Convertible Files Tab lists all of the files in the open project that are considered non-convertible and are not containers (such as ZIP or PST files) for other items.

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will say "True", else it will be blank.
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Status
  - N.A. - indicates that the item is not convertible and is not a container of other items.
  - Container - indicates that the item is a PST file. It is not, itself, convertible but may contain items that are.
  - Skipped - item is a duplicate of another convertible item that does not have a status of Skipped
  - Pending - item is convertible but has not been queued or converted yet.
  - PassedThru - item has been passed through the converted tab.
  - Queued - item has been queued for conversion.
  - Converted - item has been converted successfully.
  - Failed - item failed conversion or failed to be extracted from its parent file.
  - ERROR - an error occurred extracting files from this file (e.g., PST, ZIP, etc.)

- Path - The full path to the item. For items contained in PST, ZIP, MSG, EML, NSF, RAR, PDF and Office files, this path will indicate the parent container.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

 Causes the selected items (or all items if no items are selected) to be rechecked for convertibility and their hash codes to be regenerated. This is useful if new software has been installed that defines new Print or PrintTo associations.

 Passes through the selected files (or all items if no items are selected) the normal comversion process by converting the metadata to image instead of the file contents.

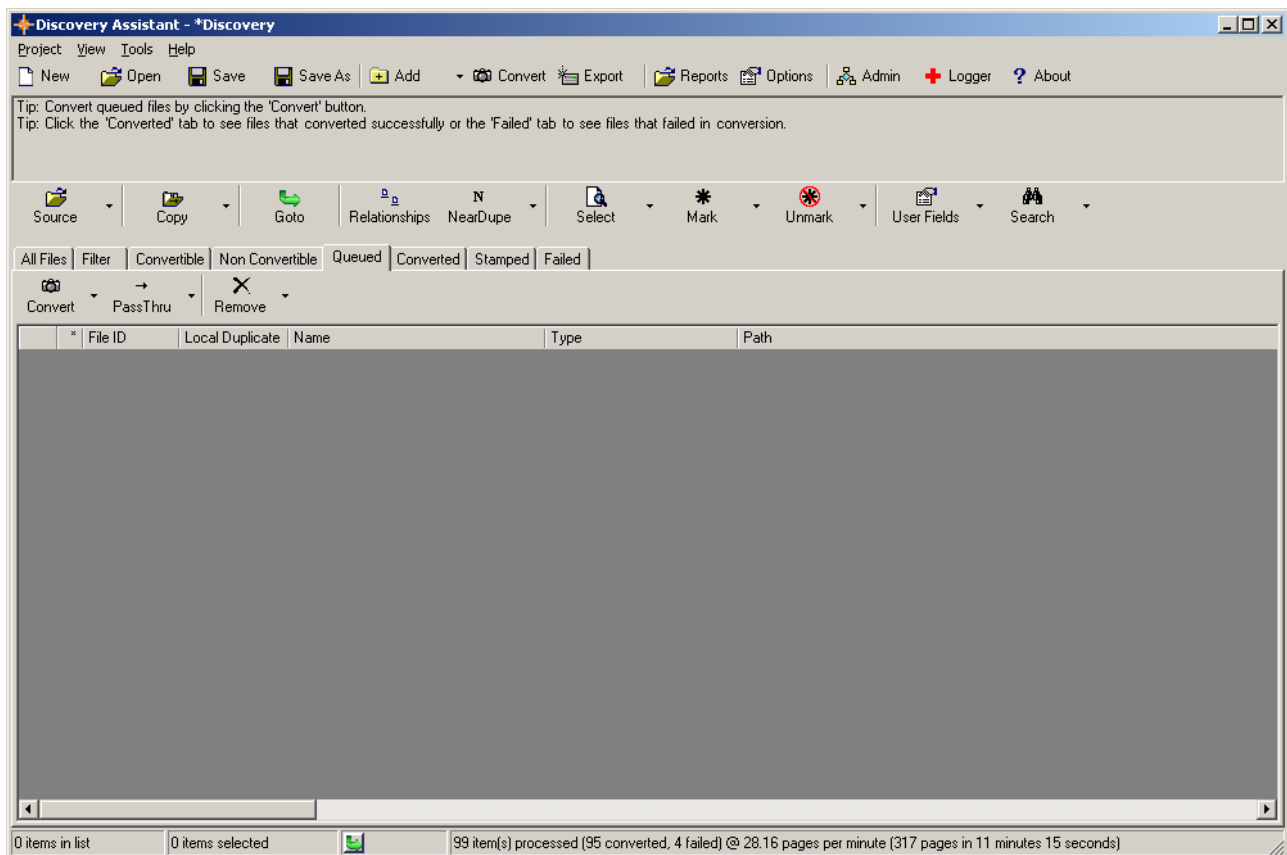 Allows you to assign a file type to the selected items (or all items if no items are selected).

 Removes the selected items (or all items if no items are selected) from the project. It does not delete them from the disk.
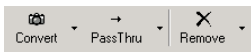
 Creates a Detail or Summary Report in a CSV file.

The Queued Files Tab lists all files in the open project that are queued for conversion.

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Status
  - N.A. - indicates that the item is not convertible and is not a container of other items.
  - Container - indicates that the item is a PST file. It is not, itself, convertible but may contain items that are.
  - Skipped - item is a duplicate of another convertible item that does not have a status of Skipped
  - Pending - item is convertible but has not been queued or converted yet.
  - PassedThru - item has been passed through the converted tab.
  - Queued - item has been queued for conversion.
  - Converted - item has been converted successfully.
  - Failed - item failed conversion or failed to be extracted from its parent file.
  - ERROR - an error occurred extracting files from this file (e.g., PST, ZIP, etc.)
- Path - The full path to the item. For items contained in PST, ZIP, MSG, EML, NSF, RAR, PDF and Office files, this path will indicate the parent container.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

  Automatically queues and converts the selected items (or all items if no items are selected).

  Passes through the selected files (or all items if no items are selected) the normal comversion process by converting the metadata to image instead of the file contents.

  Removes the selected items (or all items if no items are selected) from the project. It does not delete them from disk.

Converted Tab

The Converted Files Tab lists all of the items that have been successfully converted.

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- QC Pass - If the QC (Quality Control) has passed, it shows "True", else if QC has failed, it shows "False", else blank is shown if QC has not yet been on that item.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will say "True", else it will be blank.
- Bates Number - If the item has been Bates stamped, this column will contain the starting Bates number, else "UNASSIGNED".
- BatesGroupRange - If Bates numbers have been assigned, for items that are parents with one or more children items, this contains the range of Bates numbers from the first page of the top-level parent to the last child's page. Otherwise, this contains blanks.
- DocID - The document ID assigned to the item, else "UNASSIGNED".
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Pages - The number of pages in the output of the current item.
- Removed Pages - If deblanking was done on the file, the number of pages that were removed, else "unknown".
- OCR'ed - "True" if the item has had OCR (Optical Character Recognition) applied to it, else blank.
- Language - If the Language option .

- Status
    - N.A. - indicates that the item is not convertible and is not a container of other items.
    - Container - indicates that the item is a PST file. It is not, itself, convertible but may contain items that are.
    - Skipped - item is a duplicate of another convertible item that does not have a status of Skipped
    - Pending - item is convertible but has not been queued or converted yet.
    - PassedThru - item has been passed through the converted tab.
    - Queued - item has been queued for conversion.
    - Converted - item has been converted successfully.
    - Failed - item failed conversion or failed to be extracted from its parent file.
    - ERROR - an error occurred extracting files from this file (e.g., PST, ZIP, etc.)
- Time Stamp - The date and time of the conversion.
- Output Type - Output file type.
- Output File - Full path to the output file.
- Hash Code - The MD5 hash code for the item.
- Modify Date - The last modified date and time of the item.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons
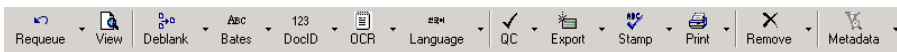
Requeue Requeues the selected items (or all items if no items are selected) for conversion.

View View the output file for the currently selected item.

Deblank Removes blank pages to save space and clean up the output file.

Bates Used to assign Bates numbers to the selected items (or all items if no items are selected).

DocID Used to assign DocID's to the selected items (or all items if no items are selected).

OCR Used to OCR selected items (or all items if no items are selected). User can define whether they want to OCR all the files or only those without text.

Language Determines the language of the selected items (or all items if no items are selected) and sets the language in the language column.

QC Opens a window to manually check files for quality.

Export Exports selected items (or all items if no items are selected).

Stamp Bates Stamp the selected items (or all items if no items are selected).

Print Send output from the selected items (or all items if no items are selected) to the printer.

Remove Removes the output of the selected items (or all items if no items are selected) and no longer shows them in this tab.

Metadata Indexes and provides filter options for the Metadata of the files in this tab.

Stamped Tab

The Stamped Files Tab lists all items that have been Bates Stamped.

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will say "True", else it will be blank.
- Bates Number - If the item has been Bates stamped, this column will contain the starting Bates number, else "UNASSIGNED".
- DocID - The document ID assigned to the item, else "UNASSIGNED".
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Pages - The number of pages in the output of the current item.
- Output Type - Output file type.
- Stamped File - Full path to the stamped output file.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

View   View the output file for the currently selected item.

Print   Send output from the selected items (or all items if no items are selected) to the printer.

Export   Exports selected items (or all items if no items are selected).

Remove   Removes the stamped output of the selected items (or all items if no items are selected) and no longer shows them in this tab.

Failed Tab

The Failed Files Tab lists all items that have failed during conversion.

Data Grid Columns

- * - This shows "*" if the item has been marked, else blank.
- File ID - This is a unique number assigned to each item in the project. Numbers are assigned in ascending order as items are scanned or dropped into the project. The File ID may not always start at 000001, but it will always be sequential.
- Local Duplicate - If the item belongs to a set of duplicates in this project, this column will say "True", else it will be blank.
- Name - Typically, this will be the filename without path information as shown on the disk. For messages contained in PST files, this column will contain the message subject.
- Type - This column contains the file type description as provided by the operating system.
- Return Code - Return code from the conversion engine, with a negative value indicating that an error occurred.
- Failure Reason - Reason for conversion failure.
- Path - The full path to the item. For items contained in PST, ZIP, MSG, EML, NSF, RAR, PDF and Office files, this path will indicate the parent container.
- User, Near Duplicate and E-Mail Thread Fields - click here for details.

Toolbar Buttons

Requeue   Requeue the selected items (or all items if no items are selected).

PassThru   Passes through the selected files (or all items if no items are selected) the normal comversion process by converting the metadata to image instead of the file contents.

Remove   Removes the selected items (or all items if no items are selected) from this tab.

Summary Report   Creates a Summary Report in a .txt file.

Reports   Creates a Detail or Summary Report in a CSV file.

Project Options

To view or change the project options, click on the Options button of the main toolbar.
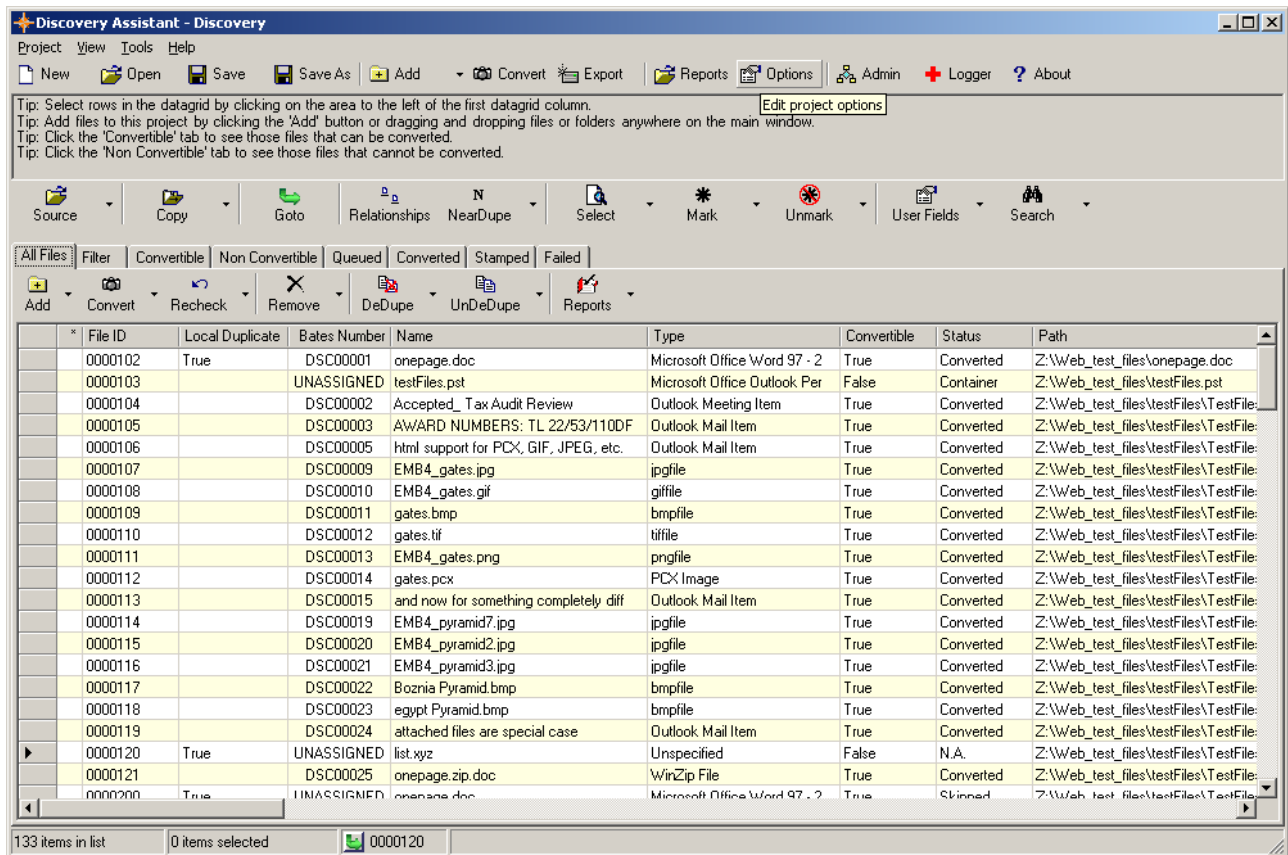
You will then be presented with the following dialog.

Save Profile - Save the current settings profile into an XML file.
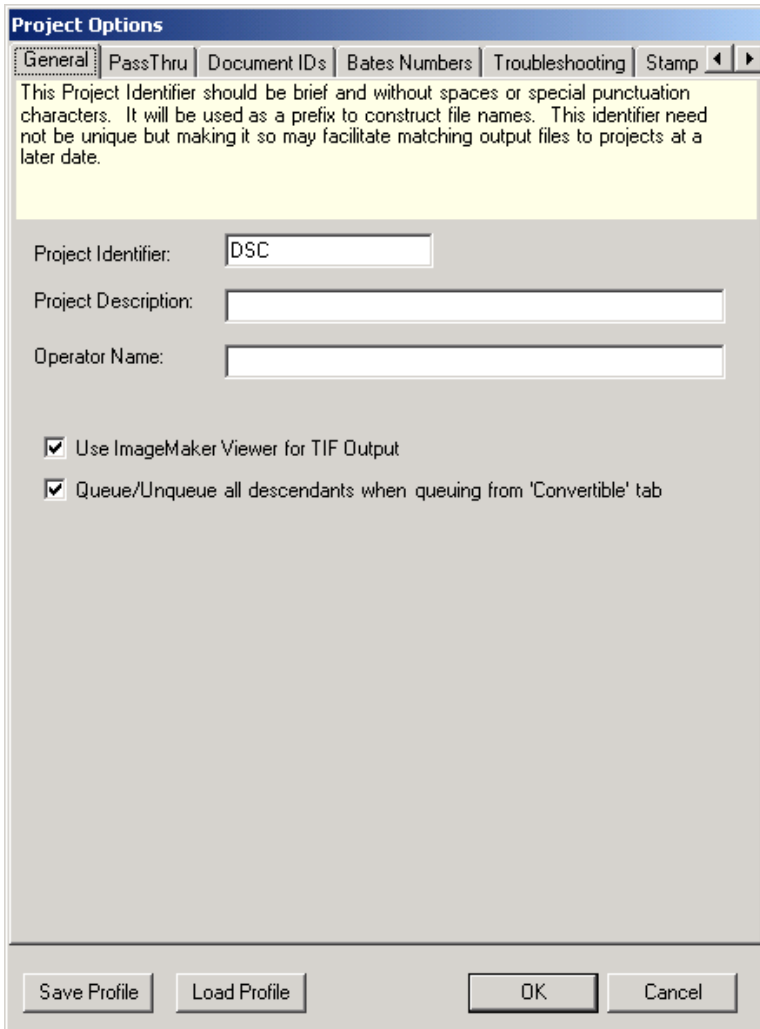
Load Profile - Load a stored settings profile XML file to have its values override the current settings values.

For information on each category of project options, click on the links below.

- General Options
- PassThru Options
- Document IDs Options
- Bates Numbers Options
- Troubleshooting Options
- Stamp Options
- Outlook Options
- Embedded Options
- De-Duping Options
- LotusNotes Options
- Output Options
- Search Options
- Scan Options
- Export Options
- Distrbuted Processingr Options
- AutoSave Options
- OCR Options

- [Print](#) Options
- [TIFF](#) Options
- [Deblank](#) Options

General Options

This tab provides several general options applicable in multiple parts of the program.

You can enter a Project Identifier. This Project Identifier should be brief (3 or 4 alphabetic characters) and without spaces.&mbsp; It will be used as a prefix to construct file names. This identifier need not be unique but making it so may facilitate matching output files to projects at a later date. Initially this Project Identifier is created for you from the Project File Prefix which is the first 3 consonants of the project name.

Project Description - This is a description of the project, with it being included in all of the reports that are created.

Operator Name - This is the name of the person who primarily will be dealing with the project, with this name being included in all of the reports that are created.

Use ImageMaker Viewer for TIF Output - If this is checked, then any TIFF output flies that have been created and that you request to be viewed (e..g., by clicking on the lower-left image in the Converted tab) will be viewed by the ImgView. Otherwise, the program that is registered on that machine to view TIFF files on that machine is used instead, e.g., the Windows Picture and Fax Viewer.

Queue/Unqueue all descendants when queuing from 'Convertible' tab - If this is checked, then all of the descendants of each item being queued from the Convertible tab will be queued, unless all of the files are queued from that tab by explicitly choosing that drop-down menu. In that case, or if this option is not selected, only the requuested items are queued for conversion.

De-Duping Options
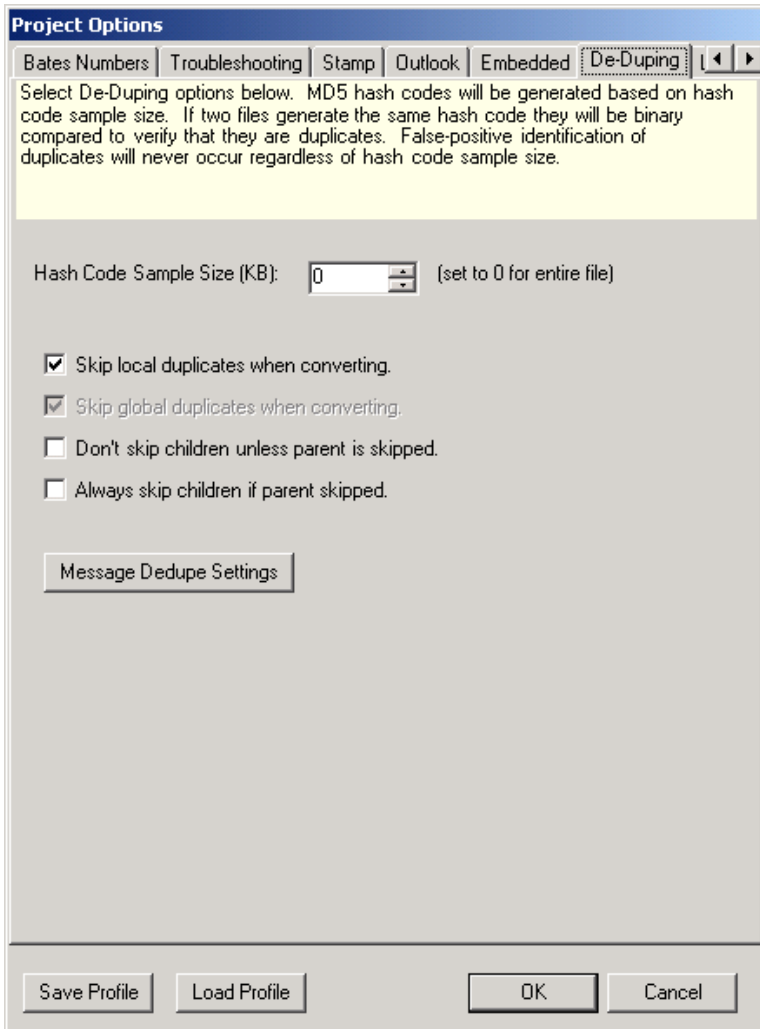
This tab lets you choose how duplicate files are handled in Discovery Assistant. Discovery Assistant generates an MD5 hash code for each item as it is scanned into the project. If two hash codes match then the items are binary compared to determine if they are identical; false positives are therefore eliminated.

Hash Code Sample Size - Determines the amount of data read from each item to generate a hash code. Increasing this value may reduce the number of full file binary comparisons required. Decreasing this value may speed up hash code processing. We recommend leaving this at its default value.

You can choose how Discovery Assistant treats duplicate files during conversion:

- Skip local duplicates when converting - If checked, change the status of all local duplicate files (i.e., files that are duplicated in just the current DA project) as "Skipped", apart from the first item or it is prevented by another one of the options provided in this dialog. This will prevent these duplicate files from being queued for conversion and, thus, they will not be converted.
- Skip global duplicates when converting - This option is only available if there is a global project associated this DA project (see ????? for details). If it is available and checked, DA will change the status to "Skipped" for all files that are duplicated in other projects associated with its global project file, apart from the first one found, it is a file that was marked using the De-Nist option (see ????? for details), or it is prevented by another one of the options provided in this dialog. As with

local duplicates, this will prevent these duplicate files from being queued for conversion and, thus, they will not be converted.

- Don't skip children unless parent is skipped - If checked, then an item will only possibly be marked as skipped if it has no parent, its parent is a PST or NSF file, or its parent is marked as being skipped.
- Always skip children if parent is skipped - If checked, then an item will always be marked as skipped if it has a parent, other than a PST or NSF file, that is marked as being skipped.

Message Dedupe Settings - Specify which e-mail fields are to be used for checking for duplicate e-mail files. See Dedupe Settings for details.

Embedded Options
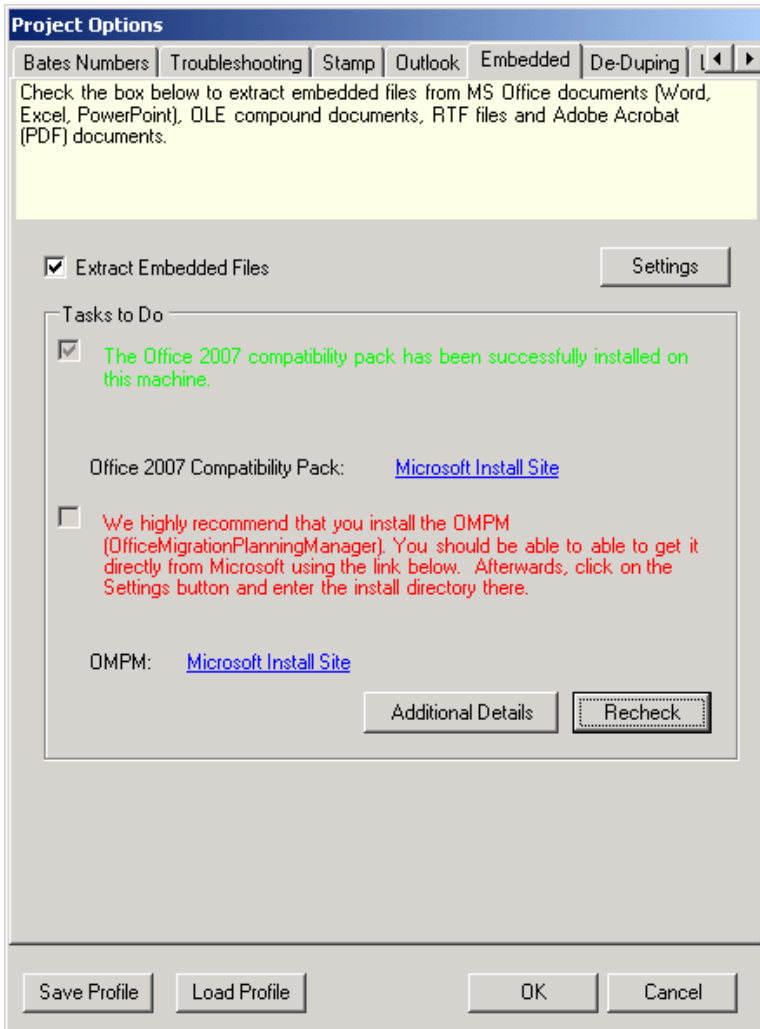
This tab allows the user to set whether the want to extract embedded files in various types of files, especially Microsoft Office documents and PDF files.

Clicking the Settings button will popup a dialog allowing you access to the extraction settings.

Tasks to Do - This groups the tasks that you should do to be able to fully and most reliably be able to extract embedded files. Note that links are provided to be able to download these free programs directly from Microsoft.
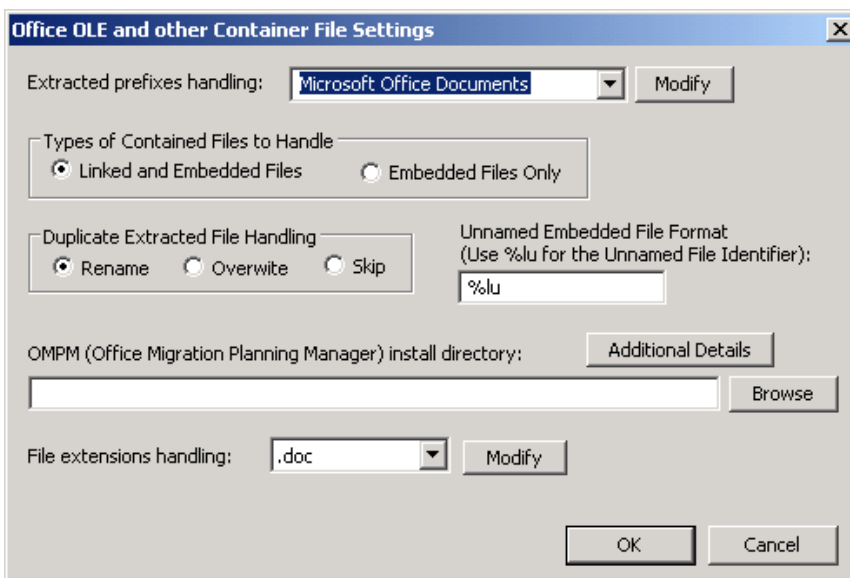
- Office 2007 Compatibility Pack - This allows you to be able to generally read and write Office 2007 (or higher) format files using an earlier version of Office (i.e., 2000, 2002/XP and 2003). Since there are limitations in being able to enumerate and extract embedded files in earlier versions of Office formats that are not present in the Office 2007 formats, DA will be able to use this to convert older Office format files to more reliably extract embedded files. Note that you should install this even if you have Office 2007 (or higher) since DA can use the compatibility pack to convert files more quickly, and sometimes more reliably (e.g., some older file formats), than if it needs to use Office directly.
- OMPM - This acronym stands for "Office Migration Planning Manager". Although it is only strictly required for Office 2000 Excel and PowerPoint system file conversions, it allows DA to be more robust and efficient in the conversion process for Office 2000, Office 2002/XP, Office 2003 and even Office 2007 or higher systems. Note that the OMPM requires the Office 2007 compability pack to be installed. Also, if you are planning to migrate to a later version of Office, some of its tools may have additional uses for you.
- Additional Details - This will show you a dialog with more detailed information regarding what the additional tasks are used for and why you should do them.
- Recheck - When the "Embedded Options" dialog is initially shown, it will have the checkboxes checked or unchecked, and appropriate description, based on the current status. Clicking this button will cause DA to recheck the status of the items and update the display to match the current status.

PassThru Options
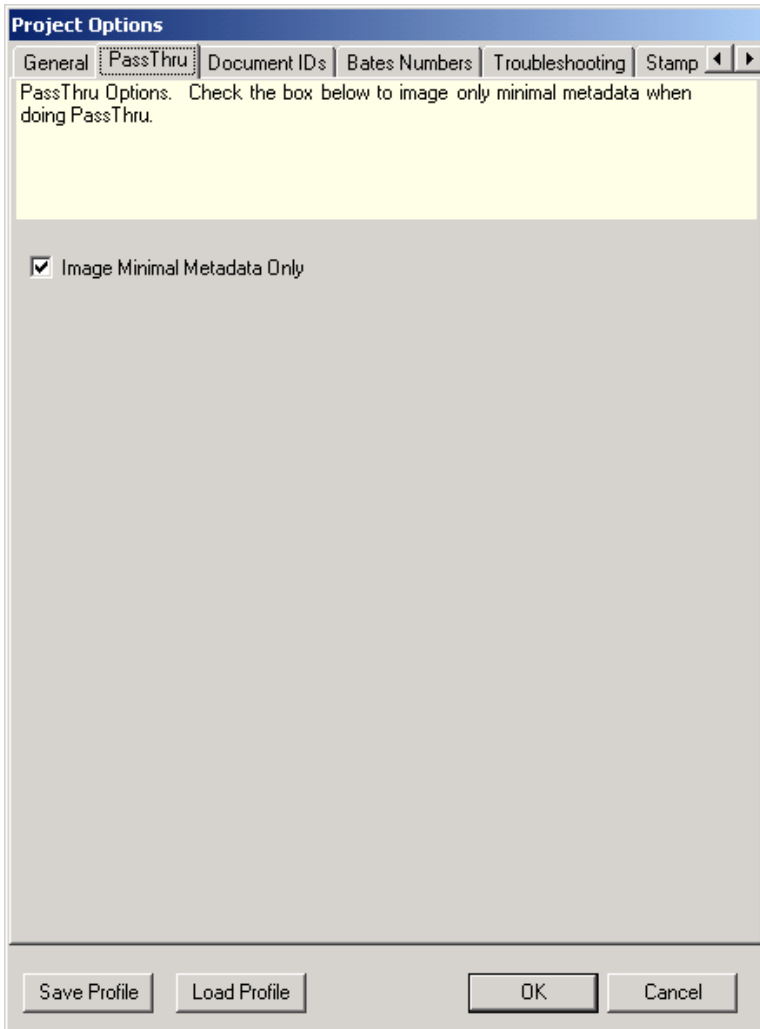
This tab allows the user to set certain default options involved with using the PassThru option to create an image placeholder for items.

TIFF Options

Project Options

General | De-Duping | Embedded | PassThru | TIFF | Output | Scan | Distribu ◄ ►

This controls how added TIFF files are treated.

Treat TIFFs As Containers

☐ Burst Child TIFF Pages from Parent, Container TIFFs

Prefix of Burst Child TIFF Pages: EMB6_

Save Profile | Load Profile | OK | Cancel

Output Options
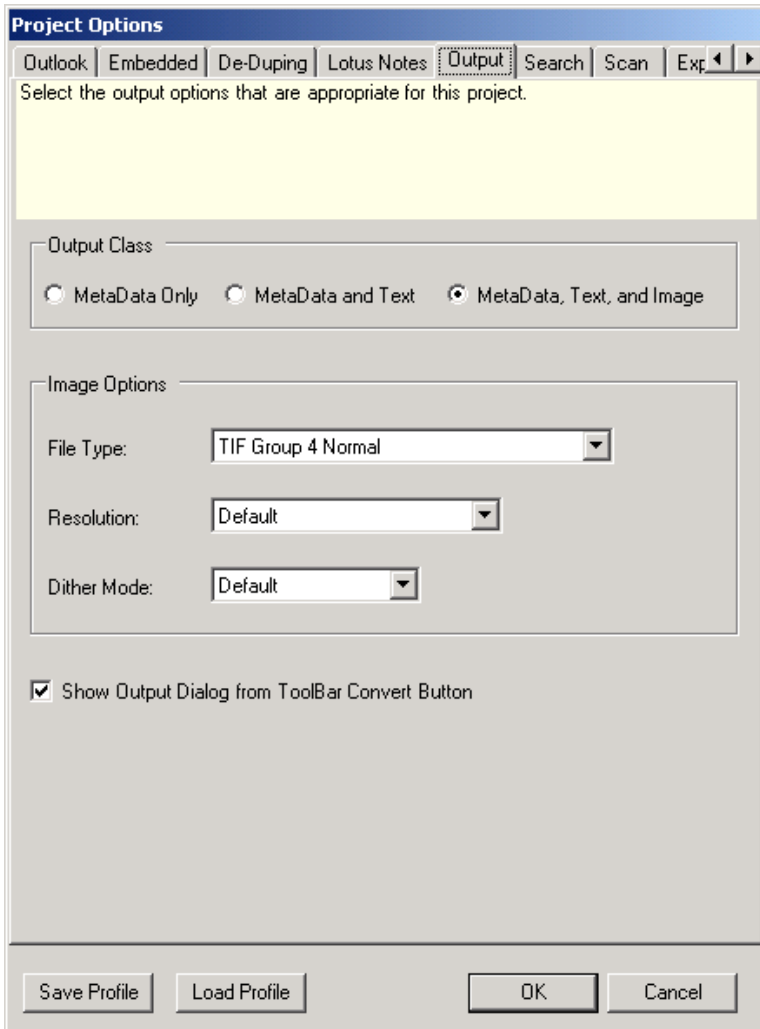
MOVE THIS TO A SEPARATE SECTION!!!!

This tab lets you select options that affect how conversions are performed.

Output Class - This affects which output files are generated during conversions:

- MetaData Only - This will only create the metadata file containing certain system and internal data of the file.
- MetaData and Text - This will also generate extracted text from the file.
- MetaData, Text and Image - This will convert the file to the requested image file format (see "Image Options" below) and, usually, use its associated text file. However, in some cases, such as PDF and XPS files, as well as HTML files for Internet Explorer 9 or higher, the text is extracted directly from the file itself instead. This is done in some cases because the generated text from the print driver is not correct.

Image Options - This controls how the image files, if any, will be generated:

- File Type - The types of generated image files that are available depend on the ImageMaker print driver that is being used on the system. The current supported values, divided into 4 groups, are: Monochrome File Types:
  - o TIF Group 3

- o TIF Group 3 (2d)
- o TIF Group 4
- o TIF Packbits
- o PCX
- o DCX
- o Windows Bitmap
- o PNG
- o PDF

Custom File Types:

- o XAR
- o Voxem
- o WinFax
- o PostScript
- o PCL

8 Bit File Types:

- o TIF Packbits
- o Windows Bitmap
- o GIF
- o PDF Colour
- o PDF Grey
- o PNG Colour
- o PNG Grey
- o TIF Packbits Grey

24 Bit File Types:

- o TIF Packbits
- o Windows Bitmap
- o JPEG
- o TIF Uncompressed
- o PNG
- o PDF
- o TIF JPEG
- o TIF JPEG7
- Resolution - This controls how high quality, through the number of horizontal pixels and vertical pixels that are used. The supported values are:
  - o Default - depends on the current settings of the print driver for the selected printer
  - o Fine Fax - 204 x 196
  - o Normal Fax - 204 x 98
  - o 100x100
  - o 200x200
  - o 300x300
  - o 400x400
- Dither Mode - This controls how high quality, through the number of horizontal pixels and vertical pixels that are used. The supported values are:
  - o Default - depends on the current settings of the print driver for the selected printer
  - o Fine Fax - 204 x 196
  - o Normal Fax - 204 x 98
  - o 100x100

- 200x200
  - 300x300
  - 400x400
- MetaData, Text and Image - This will convert the file to the requested image file format (see "Image Options" below) and, usually, use its associated text file. However, in some cases, such as PDF and XPS files, as well as HTML files for Internet Explorer 9 or higher, the text is extracted directly from the file itself instead. This is done in some cases because the generated text from the print driver is not correct.

Scan Options

Project Options

Outlook | Embedded | De-Duping | Lotus Notes | Output | Search | Scan | Exp ◄ ►

Choose categories of files to be excluded or included when scanning your harddrive.

When scanning files/directories...

☑ Exclude Executable Files
☐ Exclude Hidden Files
☐ Exclude Outlook Attachments
☑ Exclude System Files
☑ Include Subdirectories

Save Profile | Load Profile | OK | Cancel
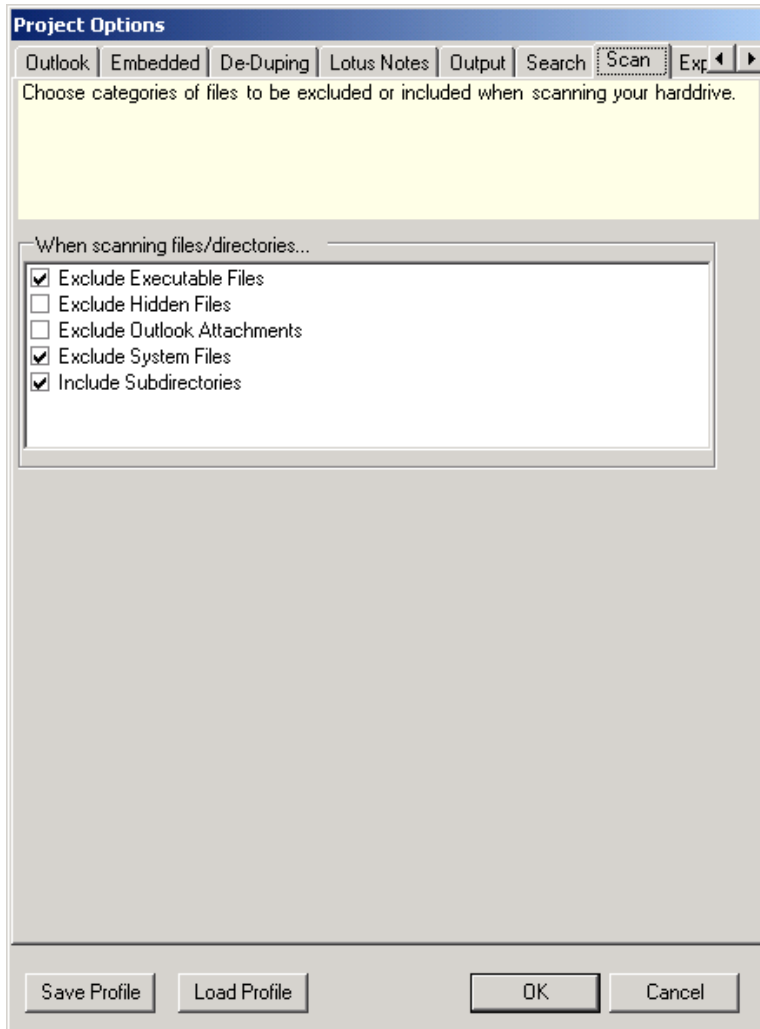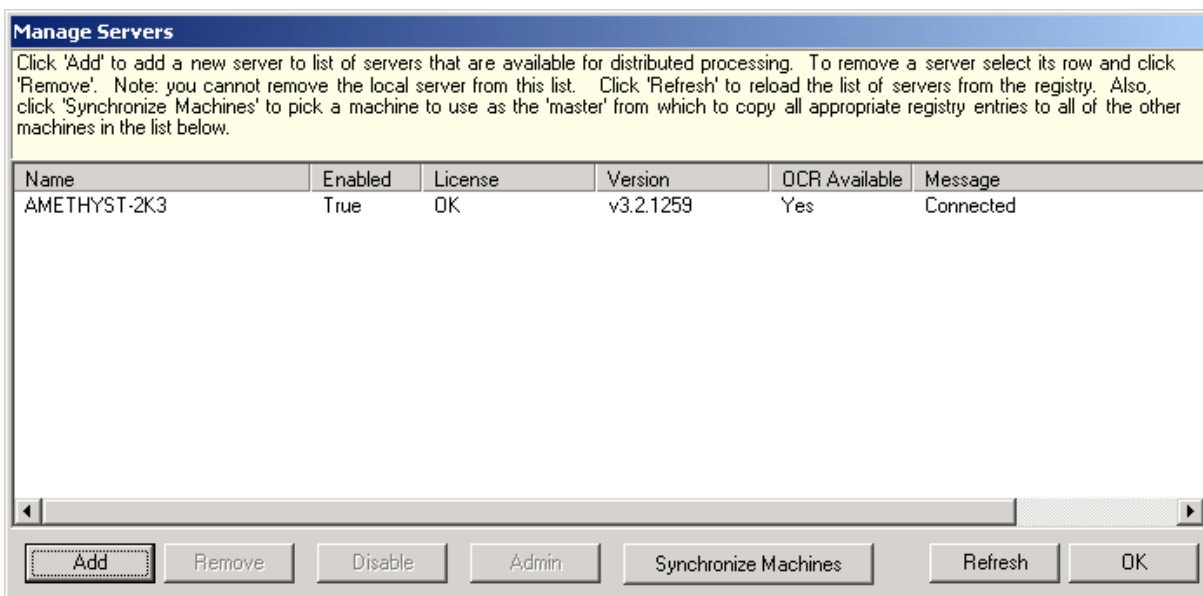
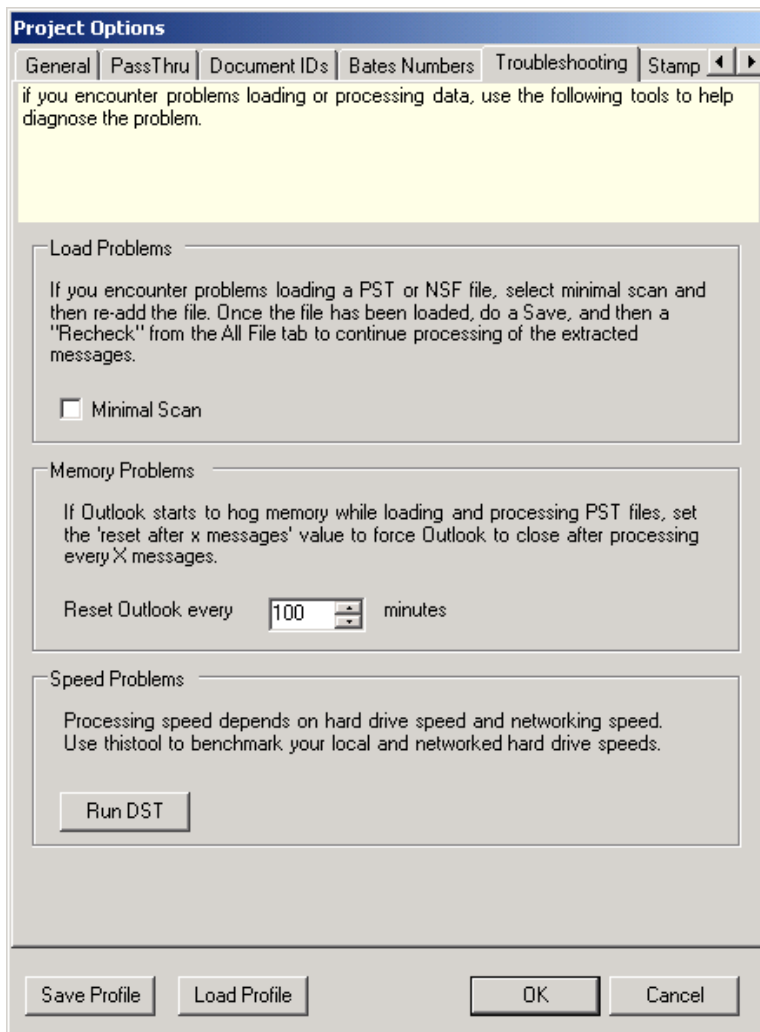## Distrubuted Processing Options
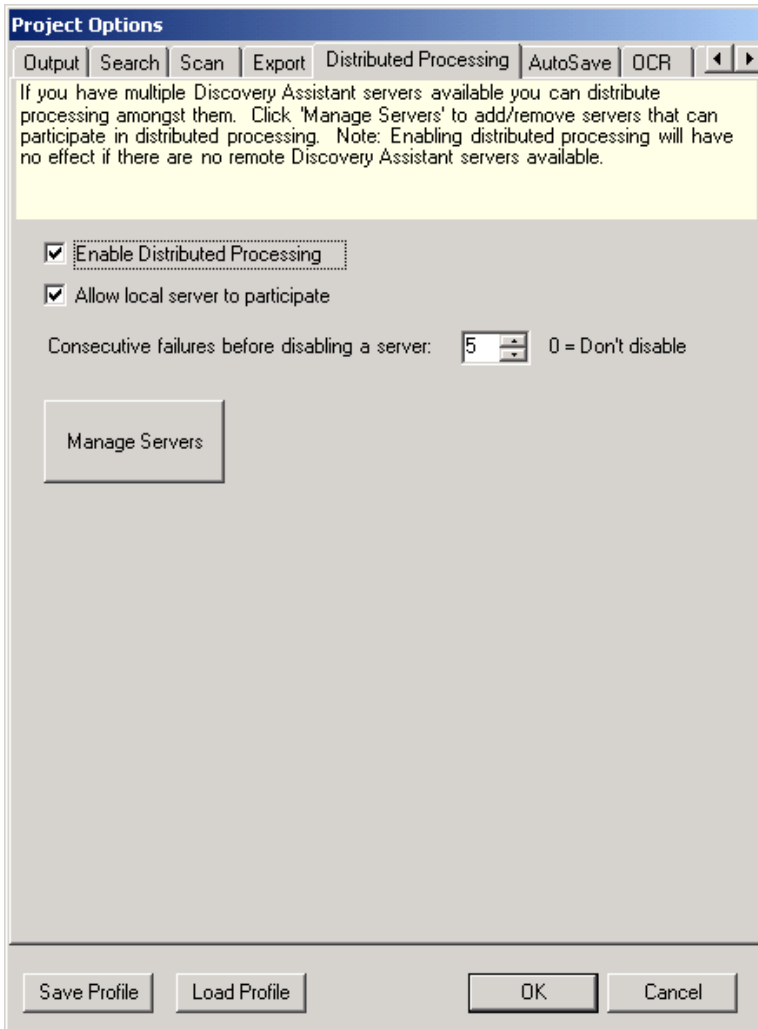
This tab lets you manage the distribted processing options. The "Manage Servers" button allows to find XDC conversion servers in your network.
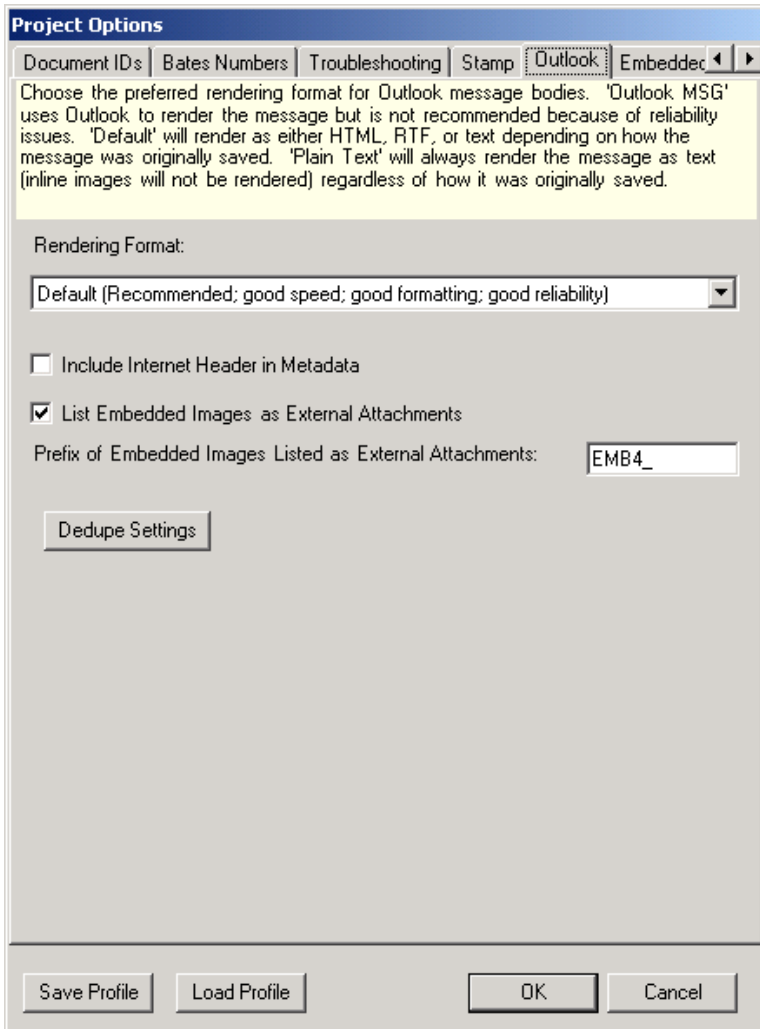
Troubleshooting Options

This tab allows the user can set various options that help with diagnosing and resolving general problems such as with loading, files, memory and hard drive speed.

Outlook Options

This tab lets you select options related to the handling of Outlook messages.

First, you can choose the default rendering format used for Outlook messages:

- Outlook MSG - Discovery Assistant uses Outlook to print the message. Formatting is exactly as printing from Outlook, but it may result in loss of information if the email message is originally in HTML format. Use of this rendering format is not recommended due to its generally slow speed and low reliability.
- Default - Discovery Assistant extracts the message to either HTML or RTF before converting. Formatting is good with no loss of information. This rendering format is recommended if output formatting is important.
- Plain Text - Discovery Assistant extracts the message to text before converting. Conversion is very fast and reliable, but formatting is plain. This rendering format is recommended if speed and reliability is important.

Include Internet Header in Metadata - If checked, this will add an extra field in the metadata file to include the header data that internally occurs at the start of each e-mail message. This is useful for checking for spam and for doing forensics, such as investigating if somebody is trying to disguise an e-mail.

List Embedded Images as External Attachments - If checked, this will extract embedded images, that are not usually listed as being attachments, to separate files as attachments. Otherwise, they are not extracted at all.

Prefix of Embedded Images Listed as External Attachments - If the "List Embedded Images as External Attachments" option is checked, this supplied string will be used as a prefix in the DA Name field before each embedded image that has been extracted as an external attachment. This allows you to easily identify which attachments are embedded images, as well as group them all together by sorting the project rows on the Name field.

Dedupe Settings - Specify which e-mail fields are to be used for checking for duplicate e-mail files. See Dedupe Settings for details.

OCR Options

Project Options

Search | Scan | Export | Distributed Processing | AutoSave | OCR | Print | ◄ | ►

Application and parameters to be used when OCR'ing images. If 'Use Microsoft Office Document Imaging Library' is checked Discovery Assistant will use this library if installed otherwise it will use the specified Command-line OCR program.

OCR Package

☑ Use Microsoft Office Document Imaging Library if available.

Command-line OCR Program

Application: `D:\Master\XDC\DiscoveryAssistant\bin\Debug\OCRCmd`

Parameters: `"%INPUTFILE%" "%OUTPUTFILE%"`

Smart OCR

☑ OCR only items without text (when OCR'ing more than one item)

☑ Show OCR Dialog from Toolbar OCR Button

Distributed Processing

☐ Allow Distributed OCR

☑ Allow local server to participate   Manage Servers

Save Profile | Load Profile | OK | Cancel

Print Options

This tab lets you select the printer to send converted items to when you press the "Print" button from the Converted or Stamped tabs.

Export Options
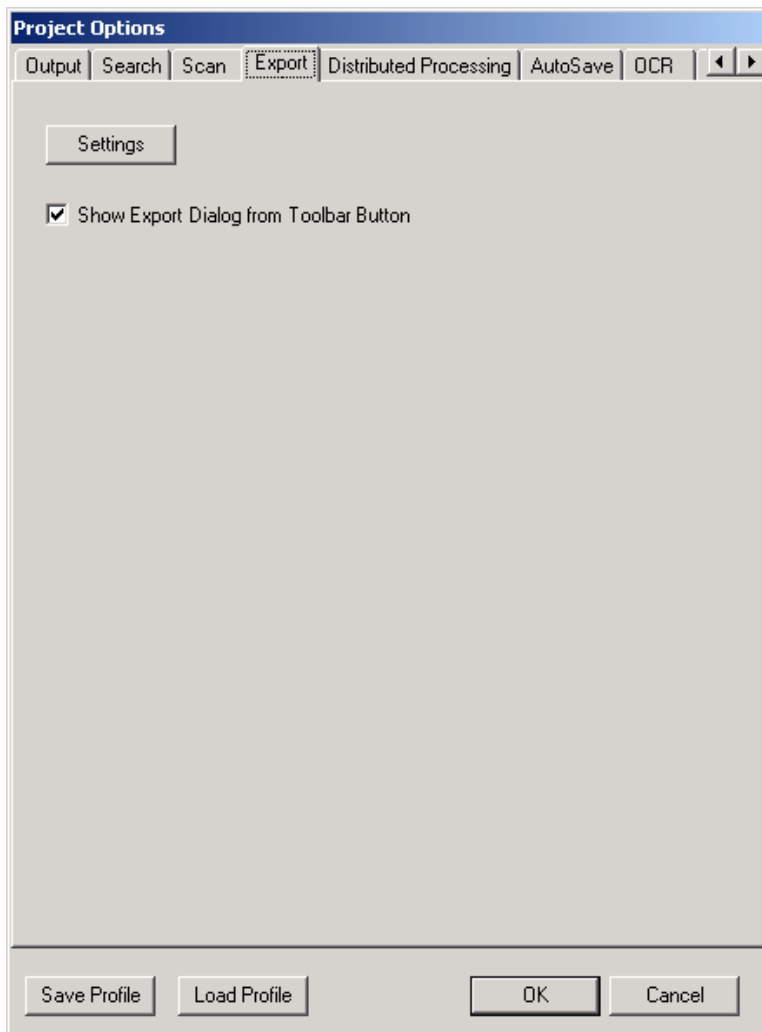
The Export Options Tab allows you to specify how converted or stamped items are exported to external systems. There are three choices for Export Type.

1. Comma Separated Text
2. Summation DII Class I
3. Concordance
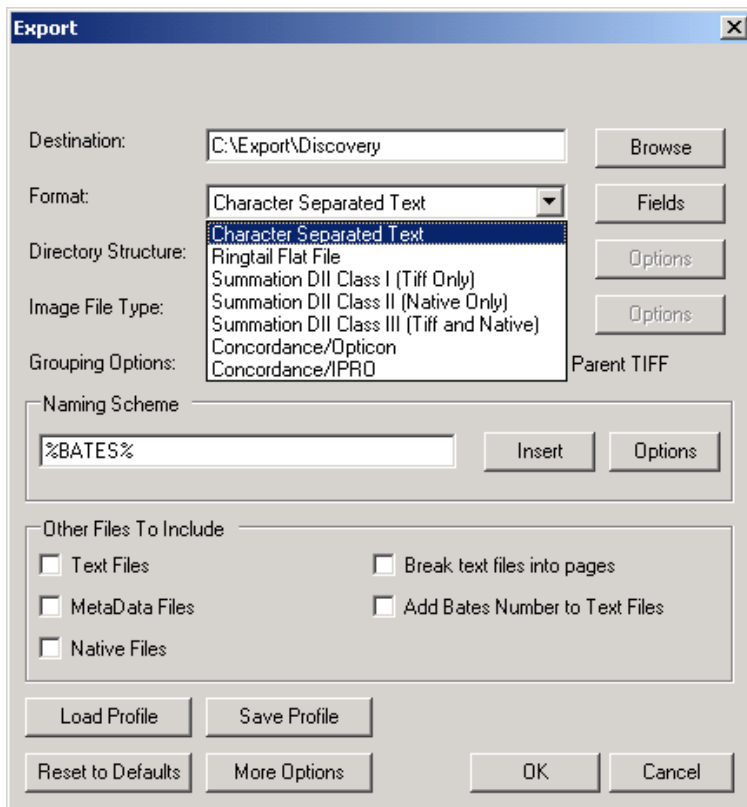
Depending on which Export Type is selected some options may not be available.

If Export Type selected is Comma Separated Text then you have a choice of three Export File Type.

1. Default - Whatever format the converted file is in.
2. PDF - TIF image within PDF.
3. Searchable PDF - Native PDF (requires GhostScript and the Discovery Assistant PostScript AddOn. Converted files must be in PostScript format).
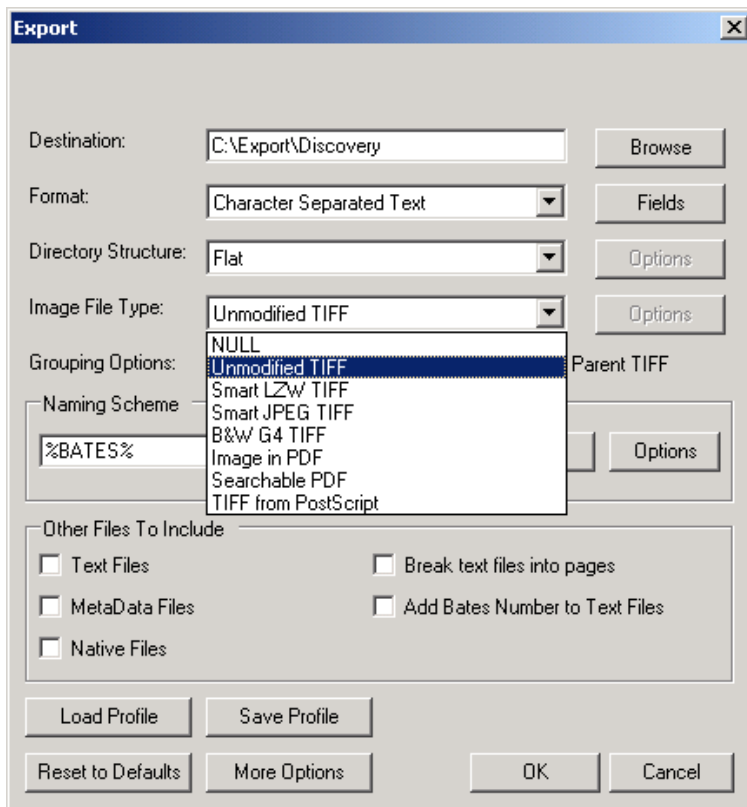
The Directory Structure drop-down list-box allows you to choose the directory structure created within the export folder. You have three choices.

1. Flat - All exported files are written to the Export Folder.
2. Mirror Source - A directory structure matching the directory structure of the input files is created under the Export Folder.
3. Volume/Box - Exported files are organized into Volume and Box folders under the Export folder.
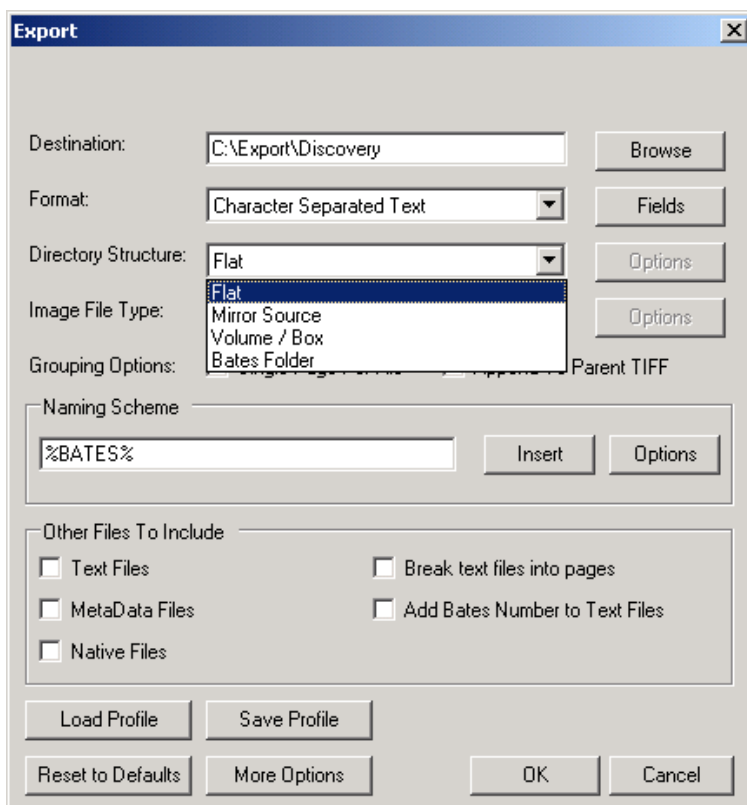
Choosing the Volume/Box directory structure allows you to limit the size or number of pages that are exported to each Volume and Box folder.

The Output File Naming Scheme allows you to control the names given to exported files. When exporting to a flat or Volume/Box directory structure it is imperative that output files have a unique name. You can ensure this by making sure that the field FileID is part of the naming scheme. If you are exporting single-page-per-file items it is imperative that you include the Page field as part of the naming scheme otherwise subsequent pages will overwrite the first page of an exported item. You can enter any text you want as part of the naming scheme and include any of the fields provided by the Insert Field button.

Character Separated Text Export Fields

The following is a list of fields that can be exported. Fields with a check mark next to them are enabled and will be exported. You can edit the Field Title by clicking on it once to select then a second time to edit. Click 'Report' to create a CSV file containing a list of all available fields, their type and description.

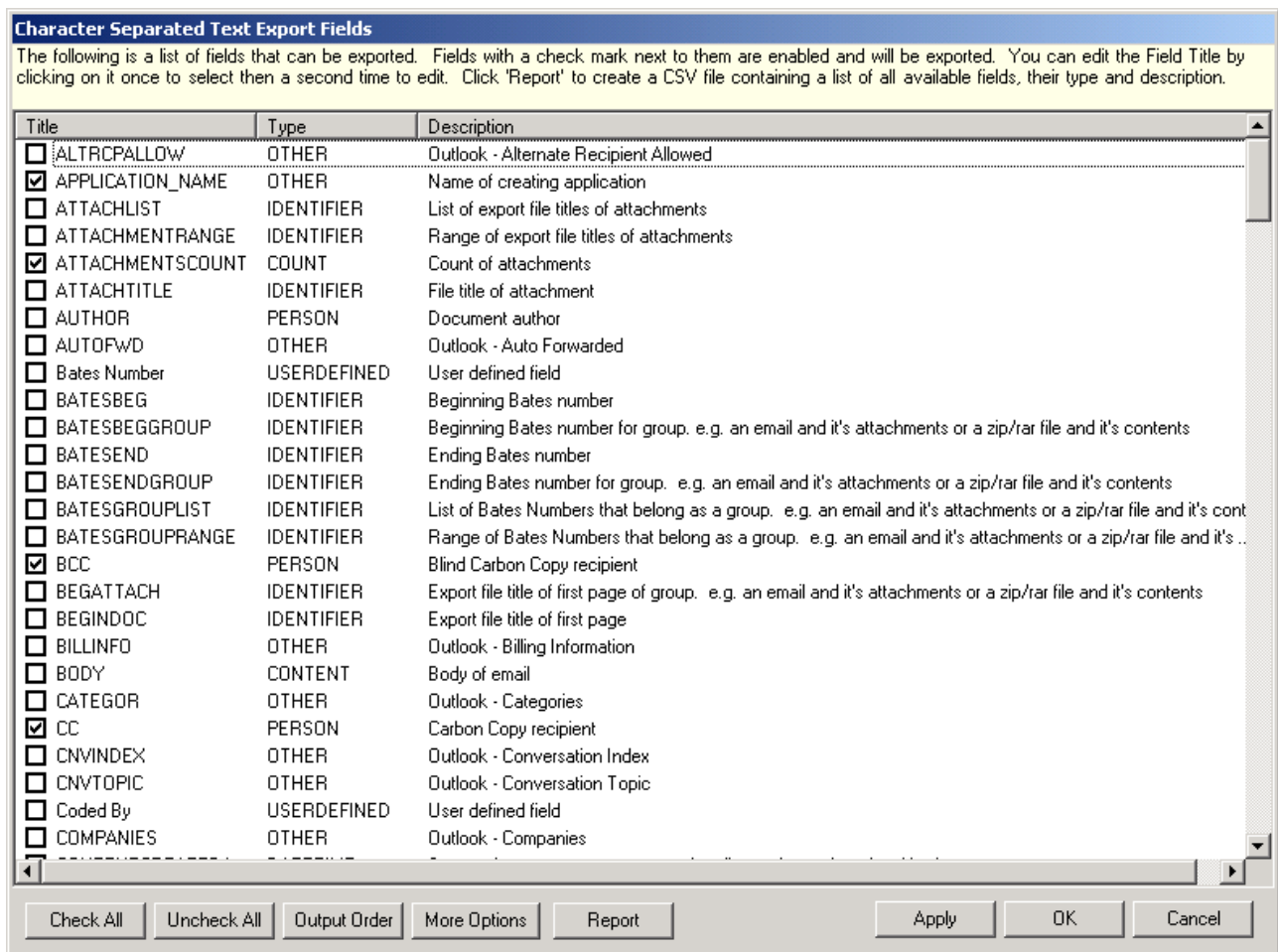| Title	Type	Description
☐ ALTRCPALLOW	OTHER	Outlook - Alternate Recipient Allowed
☑ APPLICATION_NAME	OTHER	Name of creating application
☐ ATTACHLIST	IDENTIFIER	List of export file titles of attachments
☐ ATTACHMENTRANGE	IDENTIFIER	Range of export file titles of attachments
☑ ATTACHMENTSCOUNT	COUNT	Count of attachments
☐ ATTACHTITLE	IDENTIFIER	File title of attachment
☐ AUTHOR	PERSON	Document author
☐ AUTOFWD	OTHER	Outlook - Auto Forwarded
☐ Bates Number	USERDEFINED	User defined field
☐ BATESBEG	IDENTIFIER	Beginning Bates number
☐ BATESBEGGROUP	IDENTIFIER	Beginning Bates number for group. e.g. an email and it's attachments or a zip/rar file and it's contents
☐ BATESEND	IDENTIFIER	Ending Bates number
☐ BATESENDGROUP	IDENTIFIER	Ending Bates number for group. e.g. an email and it's attachments or a zip/rar file and it's contents
☐ BATESGROUPLIST	IDENTIFIER	List of Bates Numbers that belong as a group. e.g. an email and it's attachments or a zip/rar file and it's cont
☐ BATESGROUPRANGE	IDENTIFIER	Range of Bates Numbers that belong as a group. e.g. an email and it's attachments or a zip/rar file and it's ..
☑ BCC	PERSON	Blind Carbon Copy recipient
☐ BEGATTACH	IDENTIFIER	Export file title of first page of group. e.g. an email and it's attachments or a zip/rar file and it's contents
☐ BEGINDOC	IDENTIFIER	Export file title of first page
☐ BILLINFO	OTHER	Outlook - Billing Information
☐ BODY	CONTENT	Body of email
☐ CATEGOR	OTHER	Outlook - Categories
☑ CC	PERSON	Carbon Copy recipient
☐ CNVINDEX	OTHER	Outlook - Conversation Index
☐ CNVTOPIC	OTHER	Outlook - Conversation Topic
☐ Coded By	USERDEFINED	User defined field
☐ COMPANIES	OTHER	Outlook - Companies

[Check All] [Uncheck All] [Output Order] [More Options] [Report]   [Apply] [OK] [Cancel]

Bates Numbers Options

This tab allows the user to set the default settings for when the Bates Number button in the Converted tab is pressed. These values can also be set when the Bates Number button is pressed.

Search Options
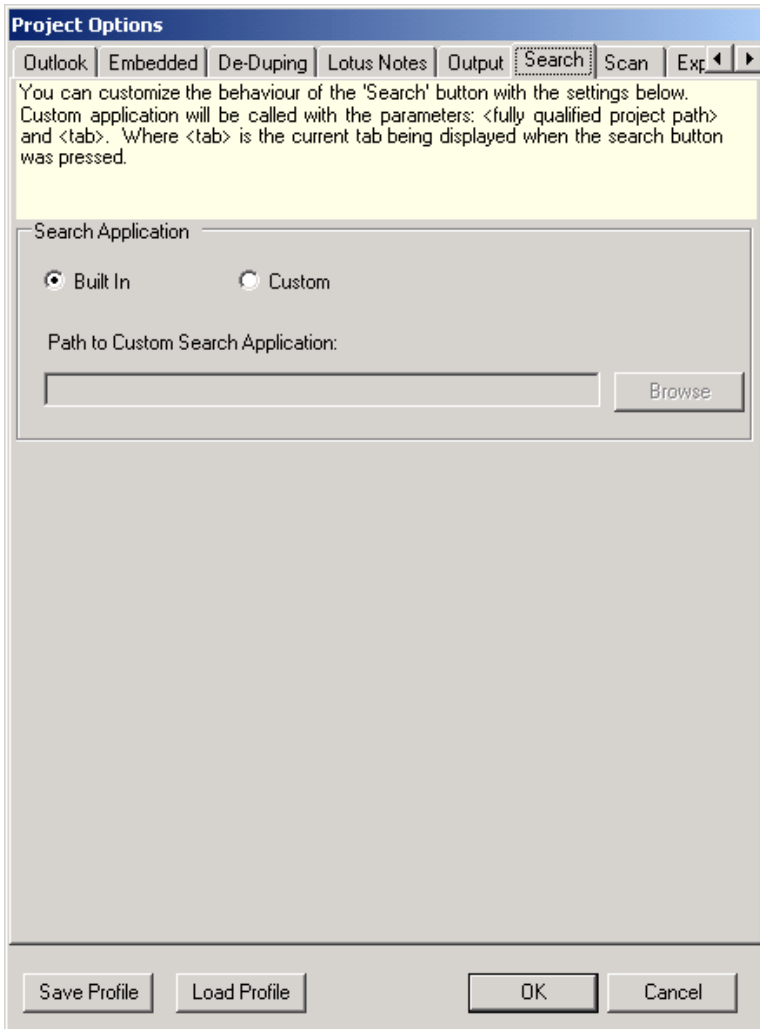
The Search Options alllows the user to choose whether you wish your Search from the Button in the Universal Tab to be done by the built-in application which is currently under-development or a custom program.

Document IDs Options

In the Document IDs Options tab, the user can set the defaults settings for when the DocID button in the Converted tab is pressed.

Deblank Options

???.
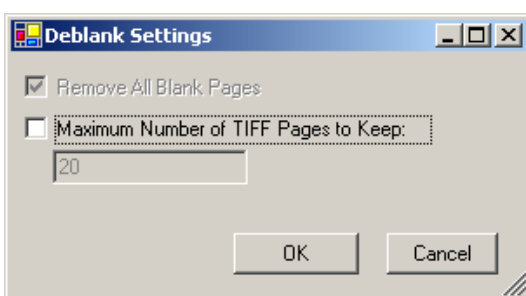
AutoSave Options
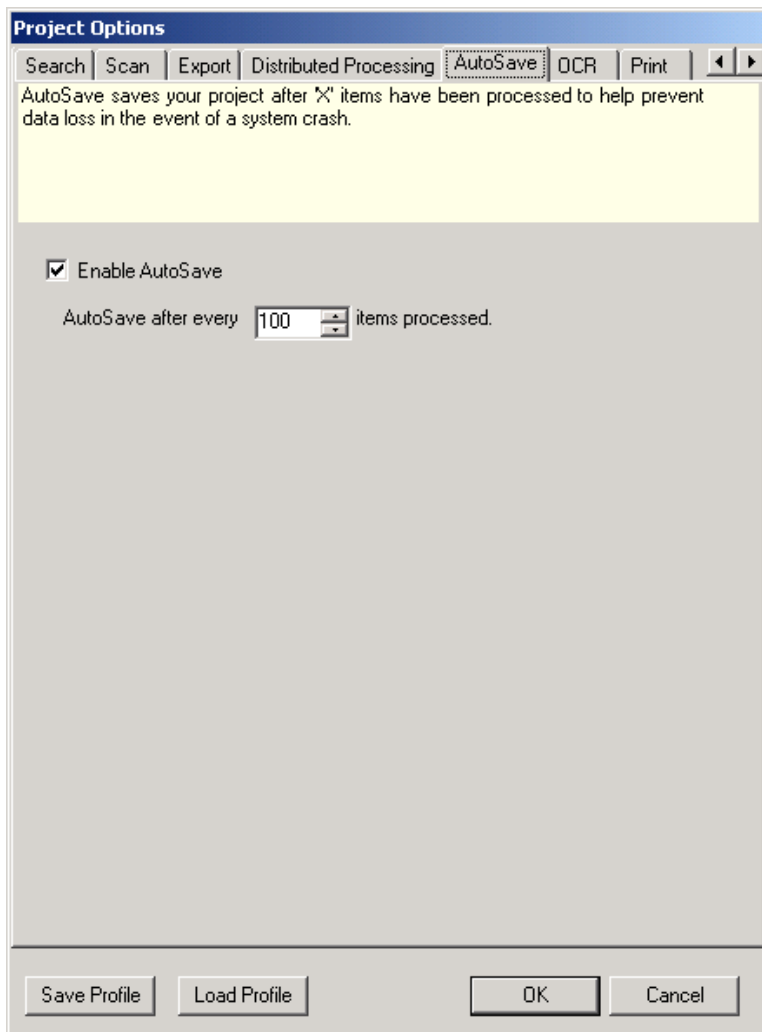
AutoSave saves your project after every 'n' conversions to help prevent data loss in the event of a system crash. You can choose to enable or disable this feature and you can set the number of conversions.

Stamp Options

Stamp Settings

Margin values given in hundredths of an inch.

Top Margin: 25   Left Margin: 25

Bottom Margin: 10   Right Margin: 25

Selected Stamp

Insert Field   Choose Font

☐ Apply to First Page Only

Watermark
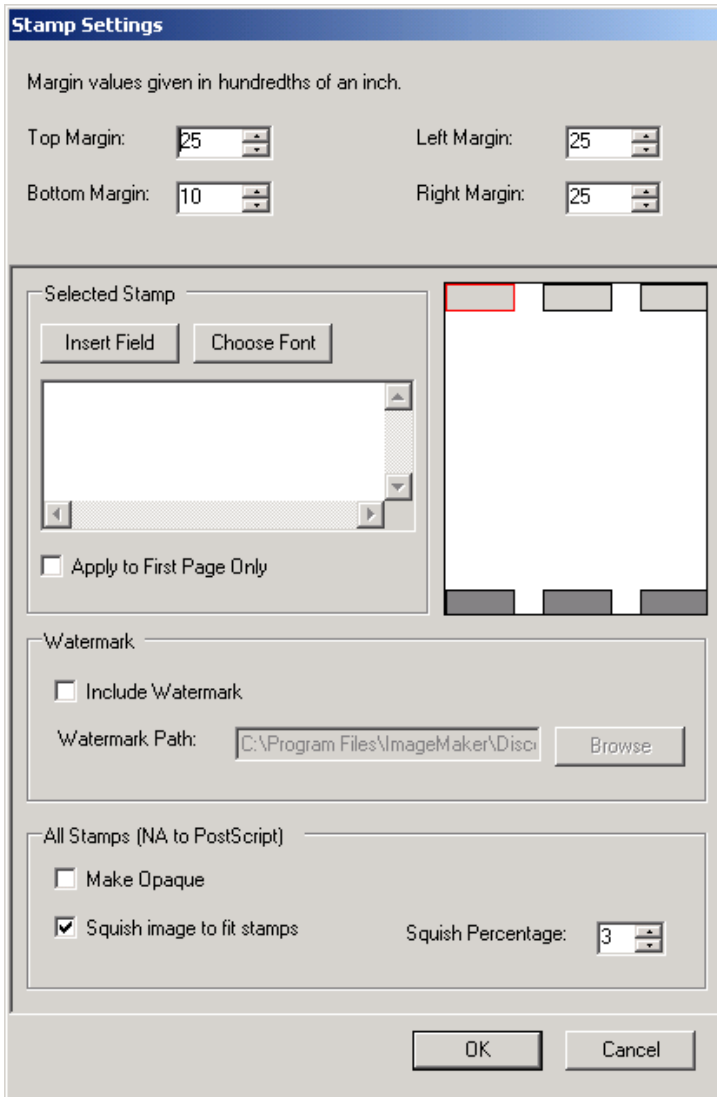
☐ Include Watermark

Watermark Path: C:\Program Files\ImageMaker\Disc   Browse

All Stamps (NA to PostScript)

☐ Make Opaque

☑ Squish image to fit stamps   Squish Percentage: 3

OK   Cancel

Lotus Notes Options

This tab allows the user to set if the documents from Lotus Notes files will be extracted as plain text or in RTF (Rich Text Format).

NOTE: For further information see Readme file accessible from Start/All Programs/ImageMaker DiscoveryAssistant.

To install

1. Take a quick look at the User Manual
2. Install the Microsoft .NET Framework Version 1.1 Redistributable Package:
   http://www.microsoft.com/downloads/details.aspx?FamilyId=
3. Install the latest Microsoft Office updates:
   http://office.microsoft.com/OfficeUpdate/default.aspx
4. Install the DiscoveryAssistant application (unzip XDCAssistant.zip and run Setup.exe).
5. If you are planning on converting PST or MSG files, you must first run the ClickYesSetup.exe system tray program. This supports the automatic closing of outlook security dialogs. (see more information below on the YesClick program).
6. Call your technical contact (Ken Davies) at ImageMAKER Development for a walkthrough instructions.
   (604) 525-2108. Pacific Standard Time.

If converting PST and MSG files, we recommend the following additional steps:

7. Outlook puts up a security dialog box that needs to be dismissed on every MSG conversion. The drop we've sent you contains a 'ClickYesSetup' application that can be installed (Start / Programs / Discovery Assistant) that will dismiss the dialog box. This app places itself in the system tray. You can click on it to set default values, including 'un-suspend'. More details at: http://www.express-soft.com/mailmate/clickyes.html
8. Conversion will go quicker if you have the Outlook client open. This reduces the time we take opening and closing Outlook.
9. There is tremendous variety in PST/MSG files. Our current working methodology is if you do have a conversion failure, we can likely comment on (and fix) the problem by getting you to send us a log files. Log files can be generated by running imgLOG.exe before loading files to be converted. (Start / Programs / Discovery Assistant / imgLog.exe).

System Requirements

- Windows 2000, Windows XP, Windows 2003 (client, server, or WTS)
- 1 gig of ram (can get by with 500 megs)
- Lots of hard drive space (for output files)
- Microsoft Office, Acrobat, IE6, and any other file type application pre-installed.

Comfortably handles the conversion of up to 100,000 files per 'project'. For example, if you have one million files to convert, then our recommendation is to break them down into 10 separate projects.

Supported File Types

We support file formats based on the file association settings. When files are imported, we identify which can be converted, and which can't. A list of supported file associations can be found by selecting Admin / Configure / Documents.

If there is an associated file type that can't be converted, then provide us with the particulars, and we'll do some investigation.

To support conversion of PDF to TXT

PDF 2 TXT conversion software. Web: www.verypdf.com

Required to support outputting TXT contents of PDF files. To be used in Discovery Assistant.

To install the PDF2TXT v3.0 demo software:
- run the pdf2txt_setup.exe application
- you registration code is: 53JHG$$@J3890J83

Make a note as to where the EXE is installed, e.g.
D:\master\XDC\Disklayout\pdf2txt\bin\pdf2txt.exe

In the Discovery Assistant application, under Admin, in the Config area, go to the Acrobat Options tab and add the following string:

D:\master\XDC\Disklayout\pdf2txt\bin\pdf2txt.exe "%INFILE%" "%OUTFILE%" -silent -first %STARTPAGE% -last %ENDPAGE%

If the output is multi-pages per file, we specify 0 for the start page, and #pages for the end page.
If the output is single-page per file, we call the application multiple times, once for each page in the source file.

This same string is stored in:
HKLM\software\imagemaker\xdc\settings\pdf2txtCmd

MSG and PST handling

Discovery Assistant excels at handling MSG and PST file formats.

Some things to do to ensure your system is running efficiently:

- Install the YesClick software to auto-close the Outlook security dialog. (see 'Handling Outlook Security Issues below).
- If you've installed Office XP, be sure to disable the Speech and Handwriting Recognition software, as this eats system resources. To disable, go to:

  "Control Panel"
  "Add/Remove Programs"
  "Microsoft Office," click on the "Change" button
  browse to "Office Shared Features," "Alternative User Input," and select for Speech and Handwriting Recognition (both) "Not available" from the drop-down box.

Very neat trick to extract multiple MSG files from a PST file for testing, debugging, message ordering etc...:

- Open Outlook mailbox
- Sort messages

- Select multiple messages, and then from the outlook menu, choose 'copy'.
- Open Windows Explorer and create a new directory.
- Select 'paste'.

The files are written to the output directory in the same order as they are listed in Outlook. To then convert these files, 'drag' the msg files from the Explorer interface into the All Files' tab of Discovery Assistant. Message order will be maintained.

Handling Outlook Security issues

When processing PST and MSG files, a recent Outlook security patch puts up a dialog that we haven't yet figured out how to close: "a program is trying to send an email with Outlook or access its address book".

We've found a third party product that can close this dialog, that installs as a system tray application. Details (and the latest download) are at: http://www.express-soft.com/mailmate/clickyes.html
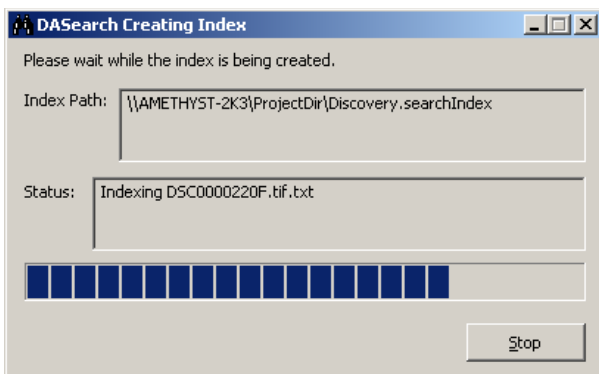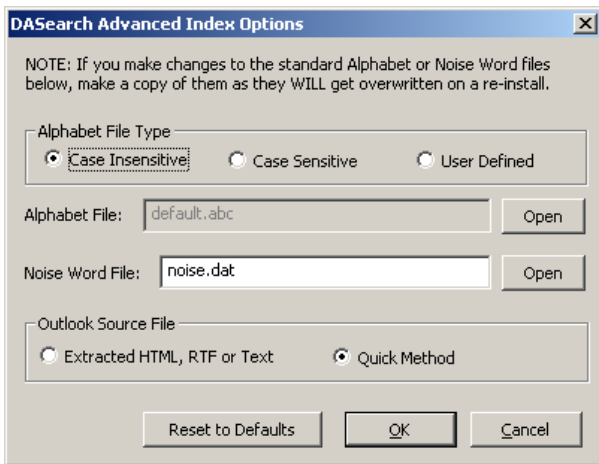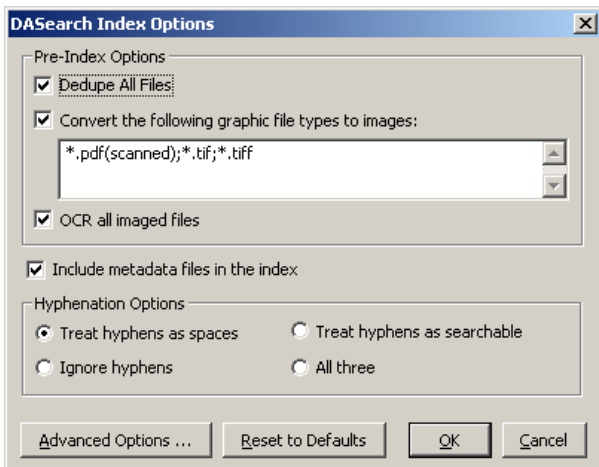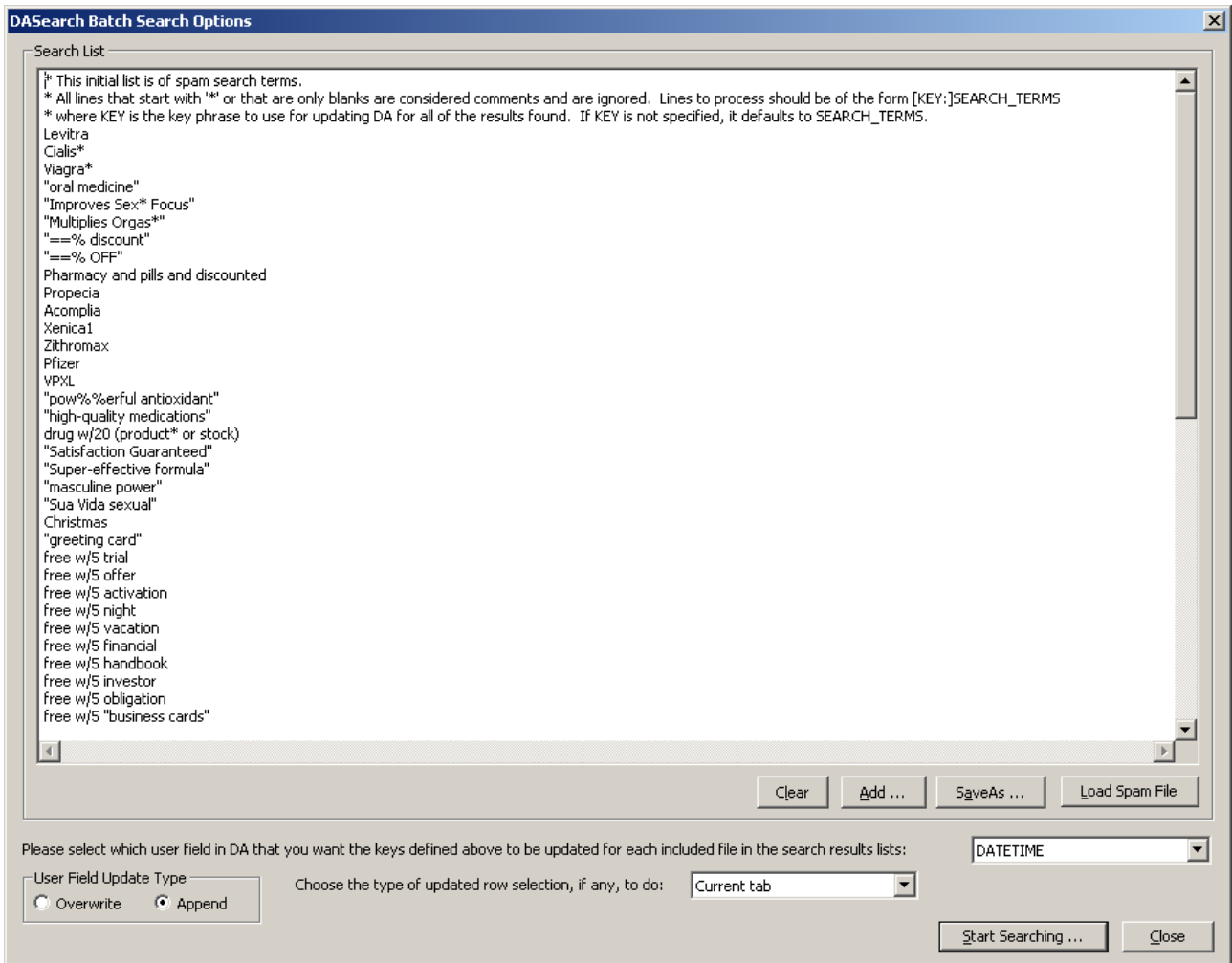
How to Install and Launch

Run the installation packagen ClickYesSetup.exe and follow the instructions (silent mode setup is available for v.1.0.7 and later). After this just double-click the ClickYes shortcut icon on your Desktop to launch the program.

How to Use Express ClickYes

- After being launched Express ClickYes sits in the taskbar and does what it should.
- To resume/suspend program execution double-click its taskbar tray icon.
- Right-click its tray icon, to get the popup menu, and select some additional options.
- To exit ClickYes, right-click its icon and select Exit in the popup menu.

Search

DASearch (186 files)

Word List:
```
grayed (2)
great (26)
greater (19)
greatest (1)
green (6)
greenland (12)
greenlanders (1)
greenlandic (2)
grid (36)
gron (7)
gronland (1)
gros (1)
gross (22)
```

Search Features
- ☐ Stemming
- ☐ Phonic
- ☐ Fuzzy searching   0

Max. # of files to retrieve: 5000

☐ Debug logging   View Log ...   Clear Log

Search Examples
Boolean logic is used, e.g., and, or, not:
    apple tart
    "maple butter tart" or "pear pie"
    "cherry pie" and "orange drink"
    "apple tarts" and not "dessert option"

Search Request:
Greenland

Status:

Search Options ...   Batch Search ...   Search   Stop   Help   Close

---

DASearch Batch Search Options

Search List
```
* This initial list is of spam search terms.
* All lines that start with '*' or that are only blanks are considered comments and are ignored.  Lines to process should be of the form [KEY:]SEARCH_TERMS
* where KEY is the key phrase to use for updating DA for all of the results found.  If KEY is not specified, it defaults to SEARCH_TERMS.
Levitra
Cialis*
Viagra*
"oral medicine"
"Improves Sex* Focus"
"Multiplies Orgas*"
"==% discount"
"==% OFF"
Pharmacy and pills and discounted
Propecia
Acomplia
Xenica1
Zithromax
Pfizer
VPXL
"pow%%erful antioxidant"
"high-quality medications"
drug w/20 (product* or stock)
"Satisfaction Guaranteed"
"Super-effective formula"
"masculine power"
"Sua Vida sexual"
Christmas
"greeting card"
free w/5 trial
free w/5 offer
free w/5 activation
free w/5 night
free w/5 vacation
free w/5 financial
free w/5 handbook
free w/5 investor
free w/5 obligation
free w/5 "business cards"
```

Clear   Add ...   SaveAs ...   Load Spam File

Please select which user field in DA that you want the keys defined above to be updated for each included file in the search results lists: DATETIME

User Field Update Type
○ Overwrite   ● Append

Choose the type of updated row selection, if any, to do: Current tab

Start Searching ...   Close

82

Advanced Features

- [Filtering Project Files](#)
- [Metadata Filtering](#)
- [Near Duplicate Detection](#)
- [Quality Control](#)
- [Relationships](#)
- [Search](#)

Filtering Project Files

Quality Control

Accept Reject | |<< << < > >> >>| Source OCR Replace Annotate Metadata

Text | Metadata | Custom Metadata

Sheet 1

Name	Value		Default Value	
Coded By		▼		▼
Reviewed By		▼		▼
Review Status		▼		▼
Privilege Status		▼		▼
Document Type		▼		▼
Privilege		▼		▼
Document Date				
Est Date				
Doc Title				
Author				
Author Org				
Addressee				
Addressee Org				
CC				
BCC				
Language		▼		▼
Location Catalog				
Call Number				
Publication				
ISBN URL				
Box Number				
File/Record #				
Folder Name				
Client Information				
Dept/Box/Container #				
Search Results				
NearDuplicate Group				
Doc Characteristics				
QC Pass		▼		▼
Assigned To				
DOCID				
File Number				

Manage Fields | Set All Default | Done | Cancel

Accept All | Accept Remaining | Reject All | Reject Remaining | Exit

□ Auto Accept   Page Flip Speed: Page Fit Mode: Window ▼

|< < >|| > >|

| Item 1 of 1 | 0 accepted; 0 rejected | [0000001] 3sheets.xls | Page 1 of 3 | Status: Unassigned |

Metadata Filtering

DAFilter (95 files)

Filter Features

☐ Stemming   ☐ Phonic   ☐ Fuzzy searching   [slider] 0

Max. # of files to retrieve:
5000

☐ Debug logging   View Log ...   Clear Log

| Regular | Custom |

File Creation Date: [All files, regardless of their creation date ▼]

File Last Mod. Date: [All files, regardless of their last mod. date ▼]

Date Sent Field: [All files, regardless of their date sent ▼]

Date Received Field: [All files, regardless of their date received ▼]

All Date Fields: [All files, regardless of any listed dates ▼]

File Size: [All files, regardless of their size ▼]

NOTE: The following string fields use Boolean (and, or, not, ...) logic, e.g., "apple juice" and oranges

Subject: [   ]   To: [   ]

CC: [   ]   BCC: [   ]

From: [   ]   Type: [   ]
Text to filter on in the BCC field

Name: [   ]   Body: [   ]

Author: [   ]   Attachment: [   ]

Importance: [   ]   Sensitivity: [   ]

All Listed: [   ]

Status: [   ]

Filters Used - Regular: 0; Custom: 0   Filter Options ...   Clear Filters   Search Text ...   Filter   Stop   Help   Close

Filter Results

File ID	Name	Hits	Sent	Received	Modified	Subject	Size
☑ 0000205	mommy's little secret	2	2005-06-08	2005-06-08	2005-06-08	mommy's little secret	90,777
☑ 0000198	encrypted zip file	2	2006-05-02	2006-05-02	2006-05-02	encrypted zip file	11,413
☑ 0000192	test with XLS in the middle.	2	2006-11-09	2006-11-09	2006-11-09	test with XLS in the middle.	21,779
☑ 0000190	Re: can you send me the sample MSG file…	2	2008-09-04	2008-09-04	2008-09-04	Re: can you send me the sam…	255,167
☑ 0000188	RE: method example	2	2006-12-22	2006-12-22	2006-12-22	RE: method example	50,089
☑ 0000158	Test Files XLS, TIF, DOC, PDF	2	2005-06-08	2005-06-08	2010-12-21	Test Files XLS, TIF, DOC, PDF	12,302,461
☑ 0000154	3 attachments.	2	2006-01-03	2006-01-03	2006-01-03	3 attachments.	507,493
☑ 0000145	embedded PDF, RTF, and DOC	2	2011-06-29	2011-06-29	2011-06-29	embedded PDF, RTF, and DOC	339,394
☑ 0000140	embedded images	2	2011-06-29	2011-06-29	2011-06-29	embedded images	1,364,703
☑ 0000134	Re: Command line parameters for ODCBa…	2	2010-04-06	2010-04-06	2010-08-11	Re: Command line parameter…	75,376

Source File Preview of 0000205 (mommy's little secret)

```
MessageClass :IPM.Note
From:Ken Davies
<EX:/O=IMAGEMAKER/OU=IMAGEMAKER_/CN=RECIPIENTS/CI

To:Ken Davies
</O=IMAGEMAKER/OU=IMAGEMAKER_/CN=RECIPIENTS/CN=F

Cc:
Bcc:
Subject:mommy's little secret
Sent:08/06/2005 1:59:45 PM #2005-06-08T13:59:45.0000000-07:00
Received:08/06/2005 1:59:45 PM #2005-06-08T13:59:45.0000000-
07:00
Number of Attachments:1
Attachment (1) Display Name:mommy.html
Attachment (1) File Name:mommy.html
```

Status: Showing: \\AMETHYST-2K3\ProjectDir\Discovery.tmp\DSC0000205C.mtf.txt

Metadata File Preview of DSC0000205F.mtf

```
Type:Outlook Mail Item
ID:000000006F0D91722339DC47BECE0ACF575

MessageClass:IPM.Note
Folder Path:Personal Folders\TestFiles
From:Ken Davies
<EX:/O=IMAGEMAKER/OU=IMAGEMAKER_/(

To:Ken Davies
</O=IMAGEMAKER/OU=IMAGEMAKER_/CN=

Subject:mommy's little secret
Date Created:2005-06-08T13:59:52.0000000-
07:00
Date Modified:2005-06-08T13:59:52.0000000
```

Status: Showing 2 of 2 hits

View Hits: First | Previous | Next | Last | Goto …

Filter Results Status: 19 files with 37 hits (19 total files with 37 total hits)

Open File … | Check/Uncheck … | Update DA … | Close

DAFilter (95 files)

Filter Features

☐ Stemming   ☐ Phonic   ☐ Fuzzy searching |‑‑‑‑‑‑‑‑‑‑‑| 0

Max. # of files to retrieve:
5000

☐ Debug logging   [View Log ...]   [Clear Log]

Regular | **Custom**

Name	Type	Filter
Coded By	DropDown	
Reviewed By	DropDown	
Review Status	DropDown	
Privilege	DropDown	
Privilege Status	DropDown	
Relevance	DropDown	
Client Information	DropDown	
ESI Client Info	DropDown	
UniqueID	Display	
DOCID	Display	
Document Type	DropDown	
Document Date	Date	
Create Date	Unknown	
Mod Date	Unknown	
Date Sent	Unknown	
Date Recd	Unknown	
Est Date	Date	
Doc Title	Single Line	
Subject	Single Line	
Doc Characteristics	Single Line	
Author	Single Line	
Author Org	Single Line	
Addressee	Single Line	
Addressee Org	Single Line	
TO	Single Line	
FROM	Single Line	
CC	Single Line	
BCC	Single Line	
Language	DropDown	
Search Results	Display	

NOTE: The single and multi-line fields use Boolean (and, or, not, ...) logic, e.g., morning or "late afternoon"

[Clear]   [Modify]

Status: | dtSearch filtering complete -- 19 files with 37 hits

Filters Used - Regular: 1; Custom: 0

[Filter Options ...]   [Clear Filters]   [Search Text ...]   [Filter]   [Stop]   [Help]   [Close]

Near Duplicate Detection

DANearDuplicates (75 files)

Precision: General (9 words)

Min. Match Acceptance Percent: `50`   Min. Contained Acceptance Percent: `90`

File of Boilerplate Strings to Not Search On

`C:\DiscoveryProjects\Projects\boilerplate.txt`   [Browse]   [Edit]

Automatic Boilerplate Detection Minimum Number of Files Found (0 to disable): `100`

FileID: `106`   Name: `html support for PCX, GIF, JPEG, etc.`

Near Duplicate Search Text

```
From:               Ken Davies
Sent:               May 16, 2005 12:51 PM
To:                 Ken Davies
Subject:            html support for PCX, GIF, JPEG, etc.
Attachments:        gates.gif; gates.bmp; gates.tif; gates.png; gates.pcx

HTML in MSG.

Jpeg:


GIF:
♀
```

[Load File ...]   [SaveAs File ...]

Search Status: Near duplicates search complete -- 10 files with 575 hits

Results Status:

[Advanced Options ...]   [Find Near Duplicates]   [Stop]   [Help]   [Close]

Near Duplicates Results for 0000106 (html support for PCX, GIF, JPEG, etc.)

File ID	Name	Max. Run	Word Hits	Phrase Hits	Phrase Count	% Match	% Contain	Mod. Date	Size
☑ 0000106	html support for PCX, GIF, JP...	89	91	89	89	100	100	11/01/2012	1,348
☑ 0000158	Test Files XLS, TIF, DOC, PDF	54	64	58	89	65	65	11/01/2012	1,636
☑ 0000198	encrypted zip file	54	64	58	69	65	65	11/01/2012	1,055
☑ 0000154	3 attachments.	54	63	57	76	64	64	11/01/2012	1,252
☑ 0000192	test with XLS in the middle.	54	63	57	74	64	64	11/01/2012	916
☑ 0000140	embedded images	54	63	57	71	64	64	11/01/2012	1,366
☑ 0000125	Chinese	54	59	55	71	60	60	11/01/2012	1,591
☑ 0000152	Single Attachment.	54	56	54	60	60	60	11/01/2012	863
☐ 0000190	Re: can you send me the sam...	54	59	55	220	25	61	11/01/2012	2,340
☐ 0000224	RE: ImageMAKER Discovery A...	54	75	61	576	10	62	11/01/2012	5,035

Extracted Text Preview of 0000106 ()

From: Ken Davies
Sent: May 16, 2005 12:51 PM
To: Ken Davies
Subject: html support for PCX, GIF, JPEG,
etc.
Attachments: gates.gif; gates.bmp; gates.tif;
gates.png; gates.pcx

HTML in MSG.

Jpeg:

GIF:

Status: Showing 3 reference section markers

Extracted Text Preview of DSC0000106F.tif.txt

From: Ken Davies
Sent: May 16, 2005 12:51 PM
To: Ken Davies
Subject: html support for PCX, GIF, JPEG,
etc.
Attachments: gates.gif; gates.bmp; gates.tif;
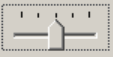gates.png; gates.pcx

HTML in MSG.

Jpeg:

GIF:

Status: Showing 3 matching result section markers

View Highlights: ☑ Missing (Orange)   ☑ Boilerplate (Blue)   ☐ Hit (Yellow)   ☑ Extra (Red)   View Markers: [First] [Previous] [Next] [Last] [Goto ...]

Results Status: 8 files with 523 hits (10 total files with 657 total hits)

[Open File ...]   [Check/Uncheck ...]   [Reshow All]   [Use Current File ...]   [Update DA ...]   [Close]

Mass Near Duplicates Options

Precision: [====slider====]   General (9 words)

Minimum match acceptance percent: [50]

Minimum contained acceptance percent: [90]

File of Boilerplate Strings to Not Search On
C:\DiscoveryProjects\Projects\boilerplate.txt   [Browse]   [Edit]

Auto. boilerplate detection min. # of files found (0 to disable):   [100]

Min. % match to group near dups. in the e-mail thread indices: [95]

Type of updated row selection, if any, to do: [Current tab ▼]

[Advanced Options ...]   [Reset to Defaults]   [OK]   [Cancel]

96

Relationships

Relationships of 0000106 – html support for PCX, GIF, JPEG, etc.

Click on a row in the list below to navigate to that item in the main Discovery Assistant window. Selected items in the list below can be selected/unselected in the main Discovery Assistant window by clicking the 'Select'/'Unselect' buttons. Click 'Back To Item' to navigate back to the original item. Click 'Source' to provide the options available to deal with the current (in this dialog) source item.

Back To Item | Source | Select | Unselect

Duplicates | Children | Ancestors | NearDuplicates | Contains | ThreadIndex

FileID	Name	Type	Status
0000106	html support for PCX, GIF, JPEG, etc.	Outlook Mail Item	Converted

Relationships of 0000106 – html support for PCX, GIF, JPEG, etc.

Click on a row in the list below to navigate to that item in the main Discovery Assistant window. Selected items in the list below can be selected/unselected in the main Discovery Assistant window by clicking the 'Select'/'Unselect' buttons. Click 'Back To Item' to navigate back to the original item. Click 'Source' to provide the options available to deal with the current (in this dialog) source item.

Back To Item | Source | Select | Unselect

Duplicates | Children | Ancestors | NearDuplicates | Contains | ThreadIndex

FileID	Name	Type	Status
0000107	EMB4_gates.jpg	jpgfile	Converted
0000108	EMB4_gates.gif	giffile	Converted
0000109	gates.bmp	bmpfile	Converted
0000110	gates.tif	tiffile	Converted
0000111	EMB4_gates.png	pngfile	Converted
0000112	gates.pcx	PCX Image	Converted

Relationships of 0000106 - html support for PCX, GIF, JPEG, etc.

Click on a row in the list below to navigate to that item in the main Discovery Assistant window. Selected items in the list below can be selected/unselected in the main Discovery Assistant window by clicking the 'Select'/'Unselect' buttons. Click 'Back To Item' to navigate back to the original item. Click 'Source' to provide the options available to deal with the current (in this dialog) source item.

[Back To Item]   [Source]   [Select]   [Unselect]

Duplicates | Children | Ancestors | NearDuplicates | Contains | ThreadIndex

FileID	Name	Type	Status
0000103	testFiles.pst	Microsoft Office Outlook P...	Container

---

Relationships of 0000106 - html support for PCX, GIF, JPEG, etc.

Click on a row in the list below to navigate to that item in the main Discovery Assistant window. Selected items in the list below can be selected/unselected in the main Discovery Assistant window by clicking the 'Select'/'Unselect' buttons. Click 'Back To Item' to navigate back to the original item. Click 'Source' to provide the options available to deal with the current (in this dialog) source item.

[Back To Item]   [Source]   [Select]   [Unselect]

Duplicates | Children | Ancestors | NearDuplicates | Contains | ThreadIndex

| FileID	Name	Type	Status	Ma...
0000152	Single Attachment.	Outlook Mail Item	Converted	060%
0000125	Chinese	Outlook Mail Item	Converted	060%
0000140	embedded images	Outlook Mail Item	Converted	064%
0000192	test with XLS in the middle.	Outlook Mail Item	Converted	064%
0000154	3 attachments.	Outlook Mail Item	Converted	064%
0000198	encrypted zip file	Outlook Mail Item	Converted	065%
0000158	Test Files XLS, TIF, DOC, PDF	Outlook Mail Item	Converted	065%
0000106	html support for PCX, GIF, JPEG, ...	Outlook Mail Item	Converted	100%

Relationships of 0000106 - html support for PCX, GIF, JPEG, etc.

Click on a row in the list below to navigate to that item in the main Discovery Assistant window. Selected items in the list below can be selected/unselected in the main Discovery Assistant window by clicking the 'Select'/'Unselect' buttons. Click 'Back To Item' to navigate back to the original item. Click 'Source' to provide the options available to deal with the current (in this dialog) source item.

| Back To Item | Source | | Select | Unselect |

Duplicates | Children | Ancestors | NearDuplicates | Contains | ThreadIndex

FileID	Name	Type	Status	Ma...
0000152	Single Attachment.	Outlook Mail Item	Converted	090%
0000106	html support for PCX, GIF, JPEG, ...	Outlook Mail Item	Converted	100%

FileID	Name	Type	Status	Thread Index
0000152	Single Attachment.	Outlook Mail Item	Converted	
0000106	html support for PCX, GIF, JPEG, ...	Outlook Mail Item	Converted	0000152-0000106

Admin

Admin is the XDC Server administrative program. XDC Server is the backend that Discovery Assistant uses for all its conversion tasks. The Admin program allows you to monitor conversions and modify XDC Server settings. You can invoke Admin either from your start menu or from the Tools menu or main toolbar of Discovery Assistant.
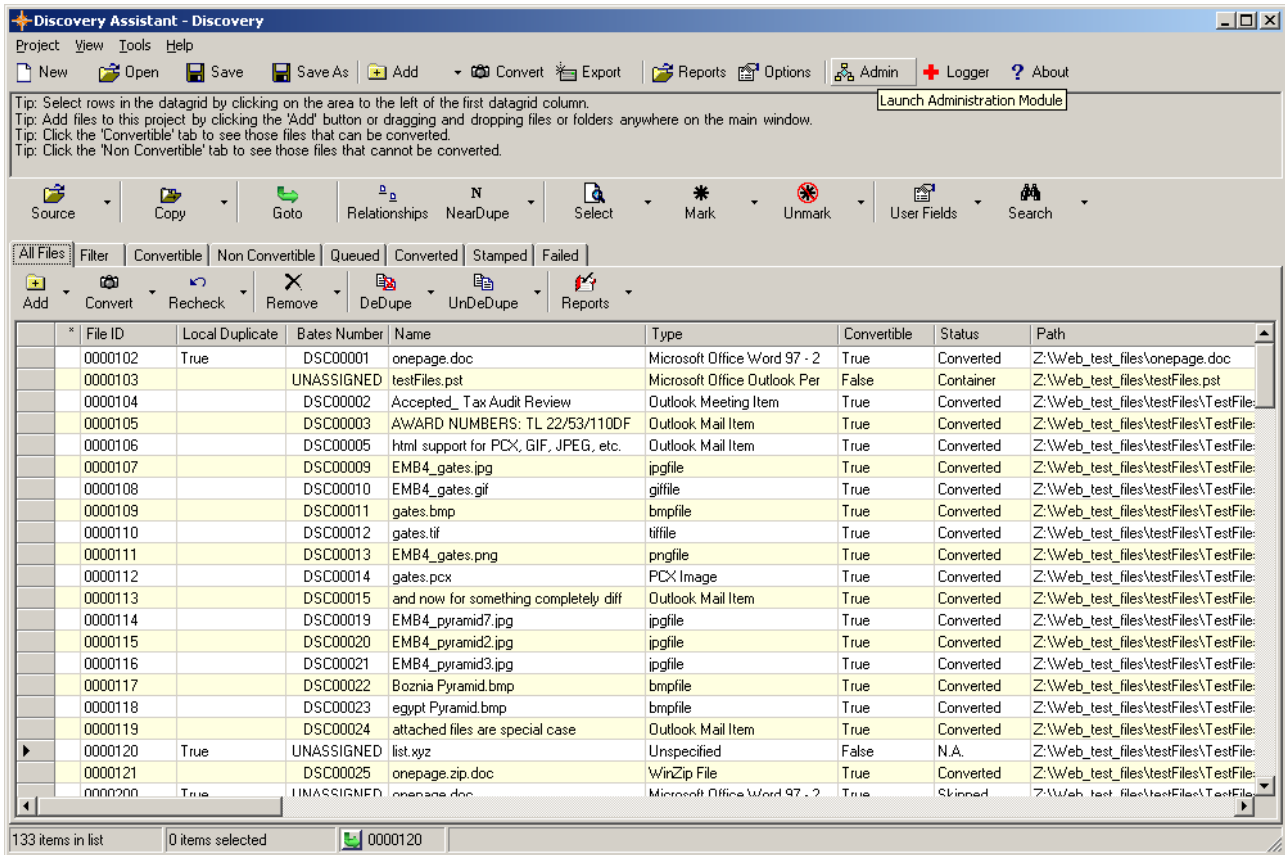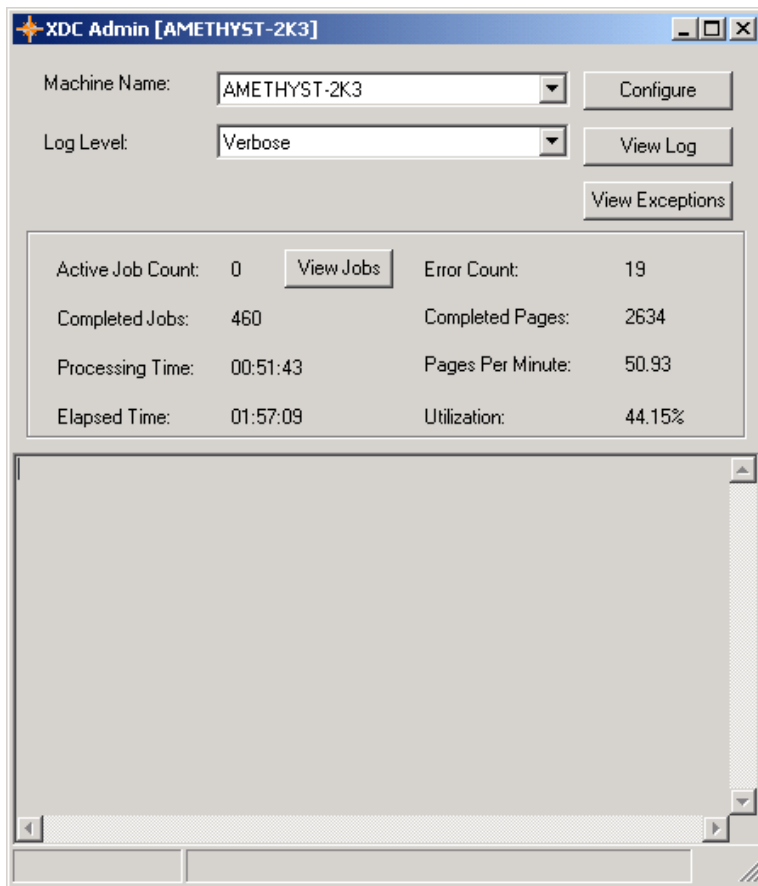
Admin Options

The Admin Options control the XDC server responsible for running Discovery Assistant's conversions. To open the Admin click the Configure button in the Admin dialog box.

TimeOuts

All time values except Max Pages are given in seconds. Zero in Total Job or Max Pages indicates an infinite value.

Total Job:	3600	Print Job Start:	60
First Page:	120	Next Page:	120
Max Pages:	0	Print Queue Clear:	60
Exclusive Access:	60	Printer Available:	60

Documents

Configure Server

Options | TimeOuts | Documents | Word Options | Excel Options | Powerf ◄ | ►

The following is a list of document types that are available on this system.

Extension	Name	Enabled
.323	H.323 Internet Telephony	Disabled
.386	Virtual device driver	Disabled
.61y	Unspecified	Disabled
.7z	WinZip File	Disabled
.8bf	Adobe Photoshop Plugin	Disabled
.8bi	Adobe Photoshop Plugin	Disabled
._sln	Microsoft Visual Studio Solution	Disabled
._sln60	Microsoft Visual Studio 6.0 Solution	Disabled
._sln70	Microsoft Visual Studio .NET 2002 Solut...	Disabled
._sln71	Microsoft Visual Studio .NET 2003 Solut...	Disabled
._sln80	Microsoft Visual Studio 2005 Solution	Disabled
._sln90	Microsoft Visual Studio 2008 Solution	Disabled
._vbxsln80	Microsoft Visual Basic Express 2005 Sol...	Disabled
._vbxsln90	Microsoft Visual Basic Express 2008 Sol...	Disabled
._vcppxsl...	Microsoft Visual C++ Express 2005 Solu...	Disabled
._vcppxsl...	Microsoft Visual C++ Express 2008 Solu...	Disabled
._vcsxsln80	Microsoft Visual C# Express 2005 Soluti...	Disabled
._vcsxsln90	Microsoft Visual C# Express 2008 Soluti...	Disabled
._vjsxsln80	Microsoft Visual J# Express 2005 Solution	Disabled
._vstasln80	Microsoft Visual Studio 2005 Tools for A	Disabled

Refresh List | Modify | New

OK | Apply | Cancel

Word Options

Microsoft Word preprocessing: Advanced settings

☐ Background printing

Print Field Codes: [Program default ▼]

Lock field codes

Header and footer: [Locked ▼]

Body: [Locked ▼]

Revision information: [Program default ▼]

Macros: [Never run ▼]

☐ Replace known date-sensitive macros with plaintext

☐ Fit tables to print area

Word-specific time-outs

Print thread timeout (from stop semaphore to thread kill) [milliseconds]: [500]

Application kill timeout (delay before sending WM_CLOSE to application windows) [milliseconds]: [5000]
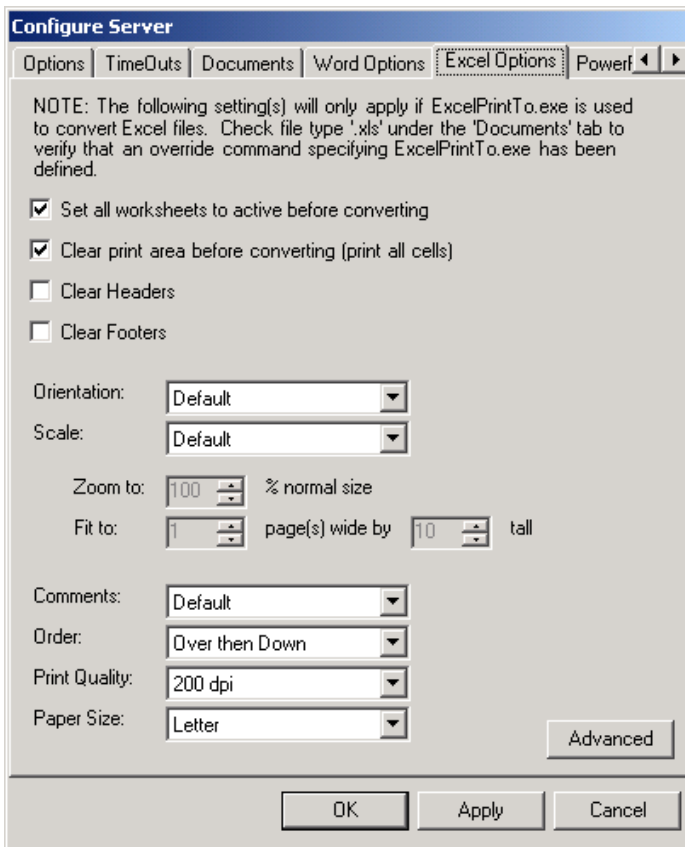
Process termination timeout (between WM_CLOSE messages and TerminateProcess call) [milliseconds]: [3000]

This version of the MSWord preprocessor is designed for use with Office 2000/Word 9 or later. Installed MSWord version: 12.0.

[OK]   [Cancel]

Excel Options

The Admin's Excel Options control the conversion of Excel Spreadsheets

Click the Advanced button to view more options.

Excel worksheet preprocessing: Advanced Settings

☐ Un-hide hidden worksheets before printing
☐ Un-hide hidden cells, columns, and rows
☐ Clear Headers
☐ Clear Footers
☐ Export formulas to text file
☑ Clear print area (prints entire sheet, not just specified area)
☐ Pause printer before conversion

Macros: [Never run ▼]

Recalculation: [On demand only ▼]

Column widths: [Global change (faster) ▼]

Column width threshold: [100]

Row heights: [Do not change ▼]

Grid lines: [Forced disabled ▼]

Row/column headings: [Default to Page Setup ▼]

┌─ Excel-specific time-outs ──────────────────┐
│ Print thread timeout (from stop semaphore   [2000] │
│ to thread kill) [milliseconds]:   │
│   │
│ Application kill timeout (delay before   [5000] │
│ sending WM_CLOSE to application   │
│ windows) [milliseconds]:   │
│   │
│ Process termination timeout (between   [3000] │
│ WM_CLOSE messages and   │
│ TerminateProcess call) [milliseconds]:   │
└──┘

This version of the Excel preprocessor is designed for use with
Office 2000/Excel 9 or later. Installed Excel version: 12.0.

[OK]   [Cancel]

Powerpoint Options

Configure Server

Documents | Word Options | Excel Options | PowerPoint Options | Acroba ◄ | ►

NOTE: The following setting(s) will only apply if PwrPntPrintTo.exe is used to convert PowerPoint files. Check file type '.ppt' under the 'Documents' tab to verify that an override command specifying PwrPntPrintTo.exe has been defined.
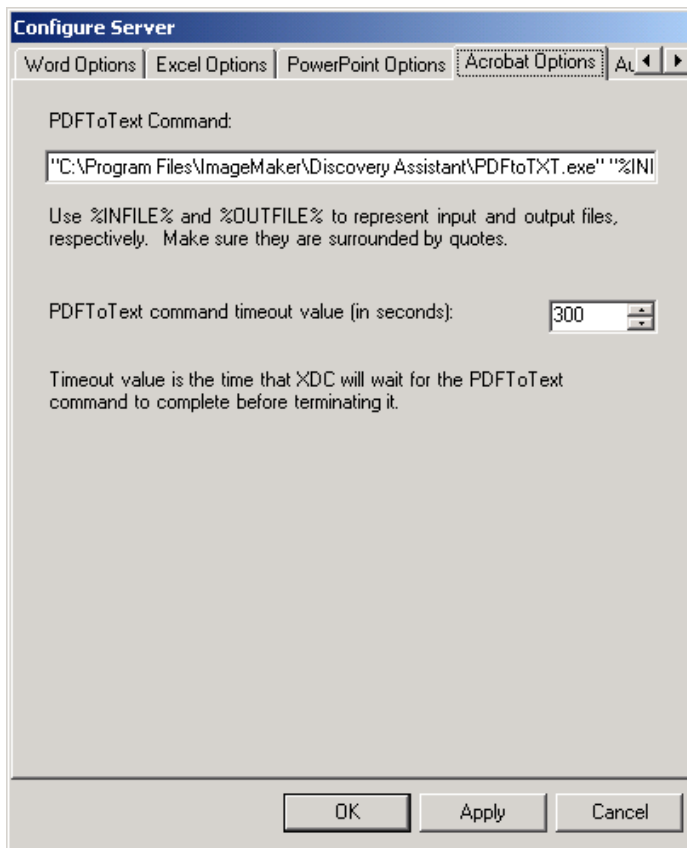
Scale to Fit Paper
- ○ Yes
- ○ No
- ◉ Default

Frame Slides
- ○ Yes
- ○ No
- ◉ Default

Print Hidden Slides
- ○ Yes
- ○ No
- ◉ Default

Output Type: [Default ▼]

Colour Type: [Default ▼]

[OK] [Apply] [Cancel]

Acrobat Options

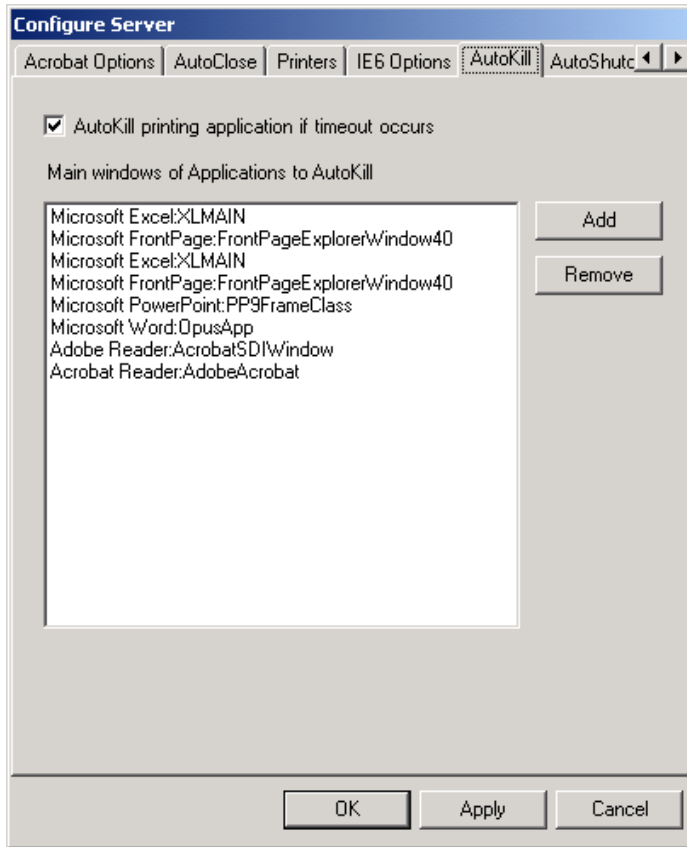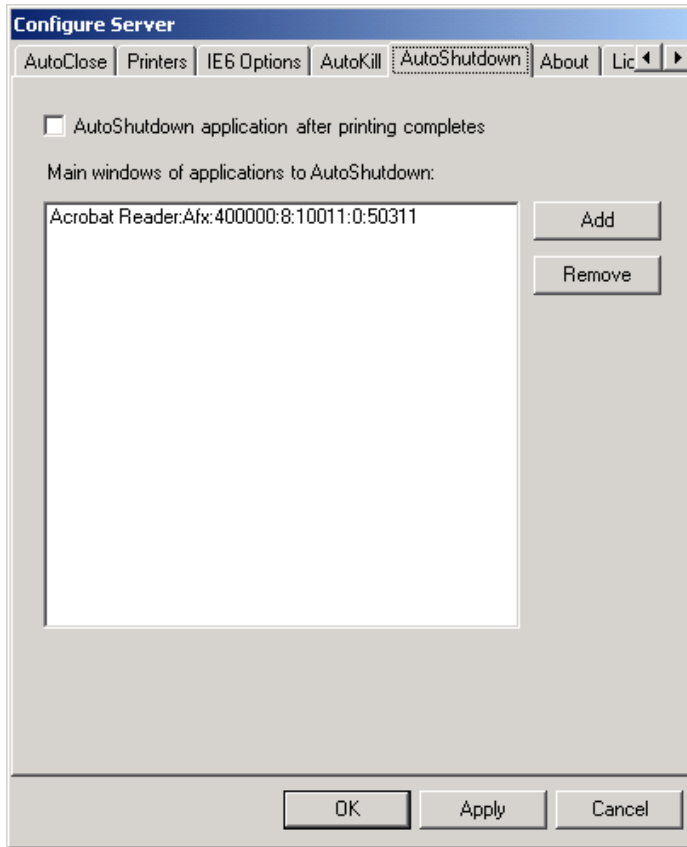## AutoClose
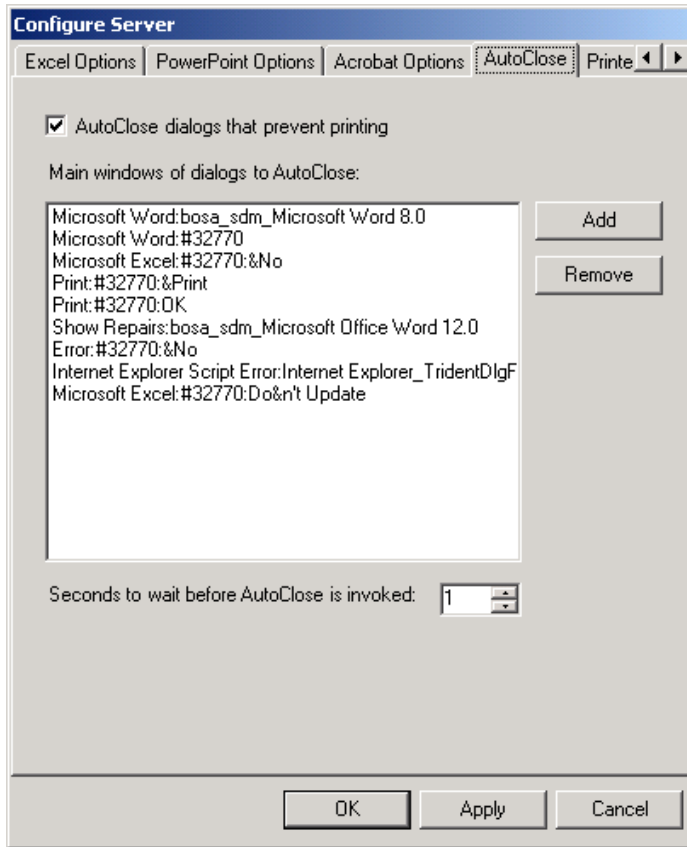
Printers
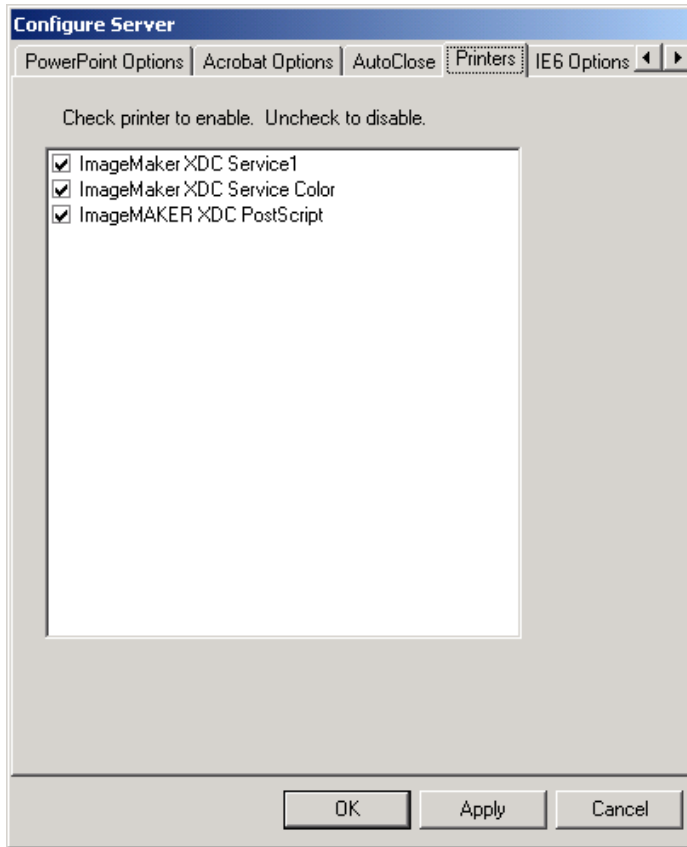
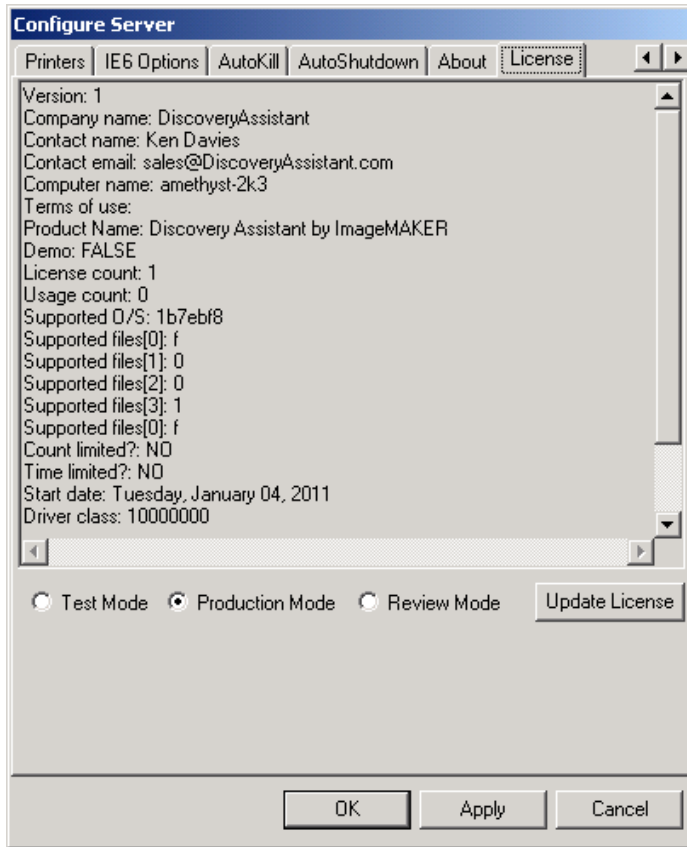## IE6 Options
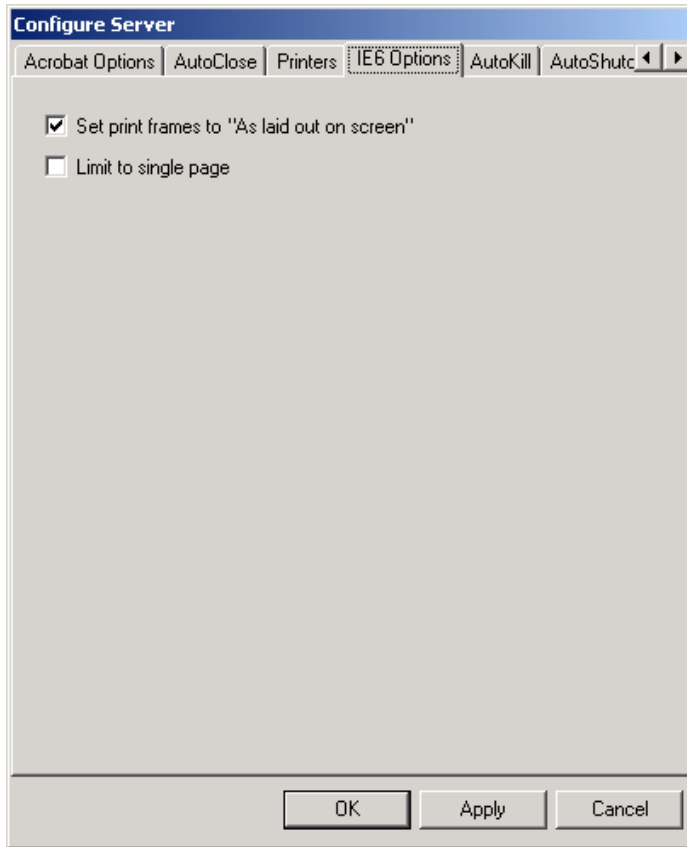
AutoKill

AutoShutdown

About

License

Active Jobs

The Active Jobs dialog displays the current active jobs. When there are no active jobs the dialog will be blank. If you want to cancel a job simply select the desired job with the mouse button and click the Cancle button.