End-to-End eDiscovery with LEADTOOLS Document Imaging





© 2018 LEAD Technologies, Inc.

www.leadtools.com

Introduction

When it comes to change, the desire for efficiency is surely at or near the top of the list of reasons. Some processes and industries are harder to change, especially those that have been around for a long time. Court systems in many countries are one of the oldest and most well established processes to ensure all-around fairness, even if it must sacrifice expediency. Thankfully, the legal industry has taken major strides towards adapting to the digital age with the evolution of eDiscovery and document imaging.

One major reason Electronic Discovery has become such a major part of the legal industry is the fact that so much data is natively generated and stored digitally. This electronically stored information (ESI) comes from a wide variety of communications (e.g. email, text messaging) and file formats. In addition to being a practical necessity, the benefits of eDiscovery are quickly expanding it into the preferred method of discovery. Rather than juggling two methods of discovery for different information sources, many legal counsels convert their traditional paper documents into ESI so all case information can follow the same process.

EDRM – Electronic Discovery Reference Model

Speaking of process, the Electronic Discovery Reference Model (EDRM) is a descriptive paradigm for how eDiscovery generally works in each case. Each stage in the EDRM is fluid and can be repeated or refined as the case evolves. The overarching goal and outcome is to take a huge amount of ESI and cull it down to what will actually be used in court.

Document Imaging technology, and in particular the technology offered in LEADTOOLS, has a role in nearly every stage of the EDRM. Some might be obvious, such as scanning paper documents into digital formats and using optical character recognition (OCR) to make the documents searchable. However, in a highly-competitive industry where time and funds can quickly get exhausted and turned into a lost case, having every available tool at one's disposal is vital to success in the courtroom. Before diving into the specific imaging technologies, the following overview of the EDRM will help enlighten how and where each technology fits in.



Figure 1: EDRM diagram (adapted from edrm.net)

Information Governance/Management

Information Governance, or Information Management ensures that proper information is saved and that it's stored and organized well.

Identification

This second phase in the EDRM is the first active step in the legal process, which throws a wide net to gather every bit of information that has any potential relevance to the case.

Preservation and Collection

After documents are identified, they must be preserved. Preservation applies a legal hold on ESI that requires the documents to not undergo any changes throughout the remainder of the case. Collection is the physical gathering of all ESI by a client into a transferable medium for sharing with their legal counsel.

Processing, Review, and Analysis

Removing duplicates and normalizing documents into decided-upon format(s) are the major tasks accomplished during Processing. Review takes a top-level look at the relevance of the ESI and Analysis is then a deep-dive look into the ESI to what is precisely relevant.

Production

Before they can appear in court, opposing sides must share their ESI with one another and decide which ESI is usable in the case.

Presentation

As implied by the name, the ESI is finally Presented in court.

LEADTOOLS at Work in eDiscovery Applications

LEADTOOLS Document Imaging SDKs have a wide gamut of imaging technologies perfectly suited in any eDiscovery application. The simplest of single-service specialists, end-to-end commercial ECMs, and everything in between will find everything needed to add world-class imaging technology to their applications.

Much the same way that the EDRM works as a general guide and process with steps that can be skipped or revisited, the imaging technology outlined below is not a mandatory set of features. For the most part, the order of these technologies follow the typical flow of use within an enterprise-level ECM but can be modified and reorganized to match the goals and creativity of any development team.

Scanning

Easily one of the most crucial elements of an ECM and eDiscovery application is the ability to digitize paper documents. Nothing is more efficient at getting high-quality digital replications of paper documents than scanning. Even if this is all that a company does to prepare for court, scanning provides massive savings in time and finances due to the simplified transportation and sharing of electronic documents. One USB stick can replace hundreds, if not thousands, of pounds of papers shipped and carried from one office to another and then to the courtroom.

LEADTOOLS includes high-level classes that make it very easy to acquire images from any scanner with a TWAIN driver or SANE backend. Consider the following snippet, which prompts the user to select a TWAIN source, then loads the acquired image into the viewer.

```
private void GetImageFromTwainSource()
{
    _twainSession.SelectSource(string.Empty);
    _twainSession.AcquirePage += new EventHandler<TwainAcquirePageEventArgs>(
        twainSession_AcquirePage);
    _twainSession.Acquire(TwainUserInterfaceFlags.Show);
}
private void twainSession_AcquirePage(object sender, TwainAcquirePageEventArgs e)
{
    rasterImageViewer1.Image = e.Image;
}
```

Document Cleanup and Preprocessing

Another hugely important feature when considering an imaging SDK is its ability to cleanup scanned images. There are two primary benefits to cleaning images, and each has huge trickle-down impacts on the entire eDiscovery process.

First, and probably the most obvious, is that the document itself is more readable. This is great for the human eye, but even better for the computer. Only a few pixels separate a lower-cased I, upper-cased L, and the number 1. The human eye can still read text with a strikethrough or a line caused by a crease in the paper, but even the best OCR engines will return gibberish.

Second, is storage space. Many compression algorithms accomplish their work by comparing neighboring pixels. This is especially true for black and white images that make up the majority of scanned documents. Performing image cleanup functions that remove dust speckles, hole punches, lines, borders, and the like have a profound impact on the length of runs and size of blocks comprised of a single color, allowing for very high compression ratios upwards of 92% of the dirty image's compressed size.

LEADTOOLS Mobile Imaging S XXX LEADTOOLS Imaging S XXX
LEADTOOLS Imaging Pro \$ XXX LEADTOOLS Document Imaging \$ XXX LEADTOOLS Recognition \$ XXX LEADTOOLS Document Imaging \$ XXX LEADTOOLS Medical Imaging \$ XXX LEADTOOLS Medical Imaging \$ XXX LEADTOOLS PACS Imaging \$ XXX LEADTOOLS Medical Imaging \$ XXX LEADTOOLS Medical Imaging Suite \$ XXX LEADTOOLS Medical Imaging Suite \$ XXX LEADTOOLS Multimedia \$ XXX Billing Address if different: BILL TO: NAME To the formation to the state of the state
name: John Smith ^{TTLE}
STREET
CITY STATE
ZIP COUNTRY
PHONE # FAX #
E-MAIL ADDRESS
SUB TOTAL
Add Shipping and Handling Charges

Figure 2: Dirty image cleaned with LEADTOOLS. Both are compressed with CCITT G4 but the clean image is only 12kb compared to 146kb

OCR

Optical Character Recognition is a must-have for any company wanting to step up their eDiscovery game. While scanning and cleaning images gets the job done and covers bases legally, OCR gives a competitive edge and is a major stepping stone into many other technologies.

Previously static images can come alive when converted to PDF, Microsoft Word, or any other number of searchable-text formats. Being able to search the contents of a digitally stored document is a monumental asset in eDiscovery, particularly during the Processing, Review, and Analysis phases. LEADTOOLS can create image-over-text PDFs, which are very useful in eDiscovery because the format retains the original image on top of a text layer, making it both searchable and virtually unchanged from the original. Even if TIFF is still the preferred file format within an ECM, OCR can infuse those static images with metadata that can be searched against or indexed for easier reference when needed.

LEADTOOLS makes OCR incredibly simple. Converting a source file on disk to a searchable PDF can be done with only three lines of code.

```
IOcrEngine ocrEngine = OcrEngineManager.CreateEngine(
    OcrEngineType.Advantage, false);
ocrEngine.Startup(null, null, null, null);
ocrEngine.AutoRecognizeManager.Run(@"C:\InputFile.tif", @"C:\OutputFile.pdf",
    DocumentFormat.Pdf, null, null);
```

Forms

Forms recognition is a specialized implementation of OCR that is finely tuned to extract specific pieces of information from documents rather than capturing the entire document. An organization may have a wide variety of forms ranging from tax forms submitted upon employment to invoices and bills sent to customers. By their very nature, forms have a lot of repeated information and perhaps the only relevant data is that which is filled in by the applicant, client, employee, etc. Therefore the most efficient and helpful means of archiving these forms within an ECM is to extract the data from those fields and either index them in a database or save them as metadata.

LEADTOOLS is able to handle both the classification and processing of forms. When classifying, LEADTOOLS will compare the incoming document against a repository of master templates (i.e. blank, unfilled versions of the form). Then once it finds a match, it will perform zonal OCR to extract the data from the filled out form.

r Form FICC-107_0 r Form FFC-107 Confidence 88%	CR_Riled	~			Owner or Controlling	P	
ederPage screteberge ■ Mit of to Station 123 Somewhere Rd				11 Individual' LEAD	11 Individual's last name, or organization's name LEAD 14 Address		
			d	14 Address			
			-		newhere Rd		
				15 City	16 State		
				Acity	NC		
		19 Telephone	number - (include area code)				
				(987) 6 5 4 - 3 2 1 0		
			-		tem if you completed item 21.		
Field Pat-III Address	Type Re Text 123	suit Confidence 3 Somewhere Rd 95	Bounding Rectangle D(=215.778, Y=1667.008, Width=2100.631, Height=61.664 pixels)	If the ow	mer or controlling person is an individu	ual	
rt-III Address rt-III Country	Text 123	3 Somewhere Rd 95	(X=215.778; Y=1667.003; Width=2100.631; Height=61.664 pxeds) (X=1596.692; Y=1784.832; Width=722.043; Height=51.136 pxeds)		· ·		
t-III Address t-III Country t-III State-1	Text 123	sult Confidence 3 Somewhere Rd 95 0 97 93	(X+215.778, Y=1667.008, Width+2100.631, Height+61.664 pxels)		· · ·	ual ass	
II Address II Country II State-1 II State-2 II Other - Text	Text 123 Text N Text N Text C Text C	3 Somewhere Rd 95 0 97	Dx-115:778, 17-1677.008, Width-2100 (2011, [http://s1.644 pound) Dx-155:682, V-1738.422, Width-72.031, [http://s1.118 pound) Dx-215:778, V-1389.216, Width-72.053, [http://s1.864 pound) Dx-215:778, V-1389.216, Width-72.051, [http://s1.864 pound) Dx-116:7783, V-2047.072, Width-461.551, [http://s1.861200000001 pound)	a 🗌 Dr	iver's license/state ID b 🗌 Pa		
III Address III Country III State-1 III State-2 III Other - Text III Issuing State	Text 123 Text N Text N Text C	3 Somewhere Rd 95 0 97	0-181-77, 1-1857-00, 141690170-1311, hegr-4, 654 (annu), 0-1956-632, 1-1714-832, Waln-722, 041, hegr-4, 514 (annu), 0-1957-73, 1-1939, 2449-73, 524, hegr-4, 51400000002, pane), 0-1967-73, 1-1939, 2449-73, 524, hegr-4, 54500000000, pane), 0-1847,733, 1-1939, 2448-743, 1449, hegr-4, 51200000001, pane), 0-1847,733, 1-1939, Waln-463, 154, hegr-4, 51200000001, pane), 0-1847,733, 1-1939, Waln-463, 154, hegr-4, 51200000001, pane), 0-1847,733, 1-132, 34, Waln-52, 213, hegr-4, 51200000001, pane), 0-1847,734, 1-132, 34, 34, 34, 34, 34, 34, 34, 34, 34, 34		iver's license/state ID b 🗌 Pa		
III Address III Country III State-1 III State-2 III Other - Text III Issuing State III Doing Business As	Text 123 Text N Text N Text C Text C Text 1 Text 1	3 Somewhere Rd 95 0 97	Dx215 721, Y=1657.08, Histor-2103.021, hep-r4-05.66 penel. Dx1986.024, Y=1784.022, Web-r22.024, hep-r4-11.15 penel. Dx1997.024, Y=130.024, Web-r20.024, hep-r4-11.05 penel. Dx1907.024, Y=120.024, Web-r20.024, hep-r4-64 for 00000000 penel. Dx1907.024, Y=120.027, Web-r405.227, Hep-r4-65 for 0000000001 penel. Dx1807.024, Y=120.024, Web-r405.218, Hep-r2-65 for 000000001 penel. Dx1807.024, Y=120.024, Web-r405.218, Hep-r2-65 for 000000001 penel. Dx1807.024, Y=120.024, Web-r405.218, Hep-r2-65 for 0000000001 penel. Dx187.024, Y=120.024, Web-r2-037.005, Hep-r4-64 152 penel.	a 🗌 Dr	iver's license/state ID b 🗌 Pa		
I Address II Country State-1 State-2 II Other - Text II Issuing State Doing Business As II First Name	Text 123 Text N Text N Text C Text C Text Ads Text Ads	3 Somewhere Rd 95 0 97 93 0 0 0 am 95 am 96	CV1117, 271, 41167, 2001, 2114,	a 🗌 Dr e ID nu	wer's license/state ID b Pa	ass	
Address Country State-1 State-2 Other - Text Issuing State Doing Business As First Name Individual's Last Name	Text 123 Text N Text N Text C Text C Text C Text Ada Text Ada Text LEA	3 Somewhere Rd 95 0 97 93 0 0 0 am 95 am 96	0.15187.774,15187.081.15187.091.1519.1519.1529.7516.654.0500.0 0.15195.6529.75174,15132.021447-22.0431,156.0500.0 0.151778,774,15132.0214587-3521,14624-0500.0000000000000000000 0.1510729,714397.02149.0549.0549.0549.0549.00000000000000000	a 🗌 Dr	iver's license/state ID b 🗌 Pa	ass	
HI Address HI Country HI State-1 HI State-2 HII Other - Test HI Issung State HI Issung State HI Individual's Last Name HI Individual's Last Nam	Text 123 Text N Text N Text C Text C Text Ada Text Ada Text Ada Text LEA Text LEA	3 Somewhere Pd 95 0 97 93 0 0 0 sm 95 sm 96 AD 98	CV1117, 271, 41167, 2001, 2114,	a 🗌 Dr e ID nu	wer's license/state ID b Pa	ass	
rt-III Address	Text 123 Text N Text N Text C Text C Text C Text Ada Text Ada Text Ada text LEA Text LEA	Somewhere 95 0 97 93 0 0 0 0 0 sm 95 AD 95 11549876 95 3456789 91	0-281-774, 1-1667-000 Webm-2012021, In-opt-4-16 646 Janual) 0-1596-052, 1-1744-1822, Webm-222 044, Negr-4-11 136 Janual) 0-25778, 1-1748-1822, Webm-2632, Heiget-46 B0000000000 panels 0-25874, 2-1749, 1-1820, Webm-2632, Heiget-46 B0000000000 panels 0-26874, 281, 1-1213, Webm-262, 281, Negr-47, 812 Janual) 0-27874, 281, 1-213, Webm-262, 281, Negr-47, 812 Janual) 0-2788, 0-1717, 1584, Webm-220, 281, Negr-47, 812 Janual) 0-2788, 0-1717, 1584, Webm-220, 281, Negr-47, 812 Janual) 0-2788, 0-1717, 1584, Webm-220, 281, Negr-47, 812 Janual) 0-1487, 271, 1-1568, 271, Webm-570, 284, Negr-47, 812 Janual) 0-2788, 0-1717, 1584, Webm-220, 281, Negr-47, 812 Janual) 0-2788, 0-1717, 1584, Webm-220, 281, Negr-47, 812 Janual) 0-2788, 0-1717, 1584, Webm-286, 786, Negr-47, 75000000000 panels 0-258, 2-1939, 744, Webm-286, 786, Negr-47, 5500000000000000000000000000000000000	a Dr e ID nu Part IV	mber Money Services and	as: Pi	
41 Address 41 Country 41 State-1 41 State-2 41 Other - Tool 41 Issuing State 41 Doing Business As 41 Frist Name 41 Individual's Last Name 41 Zip Code 41 EIN	Text 123 Text N Text N Text C Text Text Text Text Text Text Text Ads Text LEA Text 2.8 Text 1.2 Text 9.8	Somewhere 95 0 97 93 0 0 0 0 0 sm 95 AD 95 11549876 95 3456789 91	0.1415.771, 14167.108 Witter-2102131. https://d.664.down) 0.1456.682, 14167.108, 1412, Wath-22034, Hasgir-61.14500000002 pathal 0.14157, 174, 14139, 213, Wath-2034, Hasgir-64.14500000002 pathal 0.14167, 214, 1-4107, Wath-416, 214, Hasgir-64, 55000000002 pathal 0.14167, 214, 1-4207, 2143, Wath-2207, 216, Hasgir-64, 55000000000 pathal 0.14167, 214, 1-4307, 2144, Hasgir-64, 551, Hasgir-64, 547000000000 pathal 0.14167, 214, 1-14167, 2144, Wath-316, Hasgir-64, 547000000000 pathal 0.14168, 214, 1-1419, 2144, Wath-316, Hasgir-64, 750000000000 pathal 0.14168, 214, 1-1419, 2144, Wath-316, Hasgir-64, 750000000000 pathal 0.14164, 214, 1-1419, 2144, Wath-316, Hasgir-64, 750000000000 pathal 0.14164, 214, 1-1419, 214, Wath-316, Hasgir-64, 7500000000000 pathal 0.14164, 214, 1-1419, 214, Wath-316, Hasgir-64, 75000000000000 pathal 0.14164, 214, 1-1419, 214, Wath-316, Hasgir-64, 75000000000000000000000000000000000000	a Drue ID nu e ID nu Part IV 23 States an	wer's license/state ID b Pa mber Pa Money Services and d/or territories where the registrant, its	ass Pi s a	
II Address III Country III State-1 III State-2 III Other - Text III Individual's Last Name III Individual's Last Name III Do Code III Ini III Telephone	Text 123 Text N Text N Text C Text Text Text Ads Text LEA Text 28 Text 28 Text 28 Text 98 Text 65	Somewhere RI 55 0 97 33 97 97 93 98 95 MD 95 MD 98 13456789 91 7 92	0-015 72 11, -1167 708 114801701021 hegr-4 5 64 demi) 0-1956 429 1174 1128 (2014) hegr-4 51 118 (south) 0-1956 429 1174 -1138 (2014) hegr-4 51 hegr-6 58 (soutomout) panels 0-2088 229 1147 1195 129 (Mar-16 251 hegr-6 58 (soutomout) panels 0-2087 428 11-218 (Mar-16 251 hegr-6 58 (soutomout) panels 0-2184 229 11-228 (Mar-16 251 hegr-6 58 (soutomout) panels 0-2184 229 11-228 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 229 11-228 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 229 11-228 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 229 11-1597 72 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 229 11-1597 72 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 229 11-1597 72 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 229 11-1597 72 (Mar-16 251 hegr-6 158 (soutomout) panels 0-2184 259 11-1597 72 (Mar-16 251 hegr-6 158 (soutomout) panels 0-1084 259 11-1597 72 (Mar-16 251 hegr-6 158 (soutomout) panels 0-1084 259 11-1597 72 (Mar-16 251 38 hegr-6 158 (soutomout) panels 0-1084 259 11-1597 72 (Mar-16 251 38 hegr-6 254 (soutom)) 0-1084 259 11-1597 72 (Mar-16 253 18 hegr-6 254 (soutom)) 0-1084 259 11-1597 72 (Mar-16 253 18 hegr-6 254 (soutom)) 0-1084 259 11-1597 72 (Mar-16 253 18 hegr-6 254 (soutom)) 0-1084 259 11-1597 72 (Mar-16 253 18 hegr-6 254 (soutom))	a Drue ID nu e ID nu Part IV 23 States an	mber Money Services and	ass Pi s a	
Country State-1 State-2 Other - Text Ilsung State Doing Business As First Name Individual's Last Nam Zip Code EIN Telephone Telephone Rest	Text 123 Text N Text N Text C Text Text Text Ada Text Ada Text LEA Text 28 Text 12 Text 98 Text 55 Text 55	Sconvolven HU JS 0 97 0 93 0 0 0 0 am 85 NO 96 15.4 - 98.76 86 7.7 92 4.32.10 93 15.4 - 98.76 90	0.1415.771, 14167.108 Witter-2102131. https://d.664.down) 0.1456.682, 14167.108, 1412, Wath-22034, Hasgir-61.14500000002 pathal 0.14157, 174, 14139, 213, Wath-2034, Hasgir-64.14500000002 pathal 0.14167, 214, 1-4107, Wath-416, 214, Hasgir-64, 55000000002 pathal 0.14167, 214, 1-4207, 2143, Wath-2207, 216, Hasgir-64, 55000000000 pathal 0.14167, 214, 1-4307, 2144, Hasgir-64, 551, Hasgir-64, 547000000000 pathal 0.14167, 214, 1-14167, 2144, Wath-316, Hasgir-64, 547000000000 pathal 0.14168, 214, 1-1419, 2144, Wath-316, Hasgir-64, 750000000000 pathal 0.14168, 214, 1-1419, 2144, Wath-316, Hasgir-64, 750000000000 pathal 0.14164, 214, 1-1419, 2144, Wath-316, Hasgir-64, 750000000000 pathal 0.14164, 214, 1-1419, 214, Wath-316, Hasgir-64, 7500000000000 pathal 0.14164, 214, 1-1419, 214, Wath-316, Hasgir-64, 75000000000000 pathal 0.14164, 214, 1-1419, 214, Wath-316, Hasgir-64, 75000000000000000000000000000000000000	a Drue ID nu e ID nu Part IV 23 States an	wer's license/state ID b Pa mber Pa Money Services and d/or territories where the registrant, its	ass Pi s a	

Figure 3: Displaying an extracted field using LEADTOOLS Forms Recognition

PDF

Adobe Portable Document Format is by far the most popular file format for storing documents, and with good reason. PDFs are easy to read, searchable, full of metadata, and can use annotations and markup. Its use is so commonplace that any concern over being able to load the file due to needing some kind of reader or plugin is largely moot. With all of that being said, other than the initial creation of a PDF through OCR mentioned earlier, why should an eDiscovery application developer use LEADTOOLS?

LEADTOOLS includes a full-blown PDF SDK that can accomplish far more than just creating a searchable PDF through OCR. Searchable text may be the biggest advantage of PDFs within the eDiscovery process, but unfortunately, that feature is typically confined to the viewer application. In a case with thousands of documents to analyze, opening each individual file is both tedious and prone to error. LEADTOOLS allows developers to parse the PDF document structure for the text, metadata, annotations, hyperlinks, and more, making it possible to search an entire folder or disk drive of PDF files for relevant ESI in a fraction of the time.

Document Converter and Viewer

Document conversion is a very common need during eDiscovery because each case is different and every client and litigator will come to various terms for the sharing of documents. An ECM may store everything as PDF, but the agreed upon format for production could be DOC. LEADTOOLS includes a Document Converter library which can convert between a wide range of document formats including PDF, PDF/A, DOC/DOCX, XLS/XLSX, and PPT/PPTX *without the need for OCR*. This unique distinction is a huge benefit for eDiscovery because it satisfies the Preservation requirement while also simplifying Processing, Review, Analysis, and Production. The Document Converter can also take over 150 non-document formats like TIFF, JPEG, PNG, and BMP and use OCR to convert those images into PDF image-over-text, making it a perfect and programmer-friendly document normalization tool.

Closely related to the Document Converter, is the Document Viewer. On the surface, the Document Viewer, available for both .NET and JavaScript, implements many of the features that rich PDF viewing applications like Acrobat Professional contain: search and select text, follow bookmarks, add and edit annotations, and high-quality, vector-based zoom for smooth rendering at any scale factor. The Document Viewer really sets itself apart when combined with the Document Converter. Together, they offer on-thefly document normalization that can view, search, and annotate over 150 document and image formats. Cloud-based storage and viewing is growing in acceptance across many courtrooms, and the LEADTOOLS JavaScript Document Viewer is a perfect fit that can view virtually any document in its native format with all of the same features as a text-based PDF.

C:\Users\Public\Documents\LEADTOOLS Images\Leadtools.pdf - C# Document Viewer Demo		- 0 ×
File Edit View Page Interactive Annotations Preferences Help		
	CONTRACT ON A CONTRACT ON	
Reg 1 Page 1 Page 4 Page 5	Larget Audence: Software developers and programmers. Vertical Industries Server, Computer software, Banking, Insurance, Aeronautics, IT, Engineering, Healthcare, Manufacturing, Automotive, Government, Military, Education, Texities, Food and Drug, Household products. Developer Tools LEAD Technologies is the developer and publisher of LEADTOOLS, the award winning line of development tookist. LEADTOOLS is a family of comprehensive tookits designed to help programmers integrater rasker, document, medical, multimedia, vector and Internet imaging into their applications cguidely and easily. LEADTOOLS gives developers tho their applications cguidely and easily. LEADTOOLS gives developers the most flexibility and their applications cguidely. A developer and tookits designed Display, Special Effects, Scanning/Capture, Common Dialogs, Phimign, DEOM, Annotations, Forms Recognition, Document Classion, Public, POCOLS toobal twents, Public, Stores and the Integration developers toobat specification developers Work the best imaging technology available. Imaging Applications/Utilites LEAD TooLS and the best imaging technology available LEAD TooLS and the best imaging technology available LEAD TooLS and the competer and on the award- winning the common and on the context on the avard- winning the Chorlogy available	
Post-textselectionchanged operation Pre-RendrSelectorText operation Pre-RendrSelectorText operation Pre-RendrSelectorText operation Post-AnderSelectorText operation	codere developed usion industru standard as well as LEΔΩ's own normalan	· .

Figure 4: LEADTOOLS Document Viewer in .NET application showing page thumbnails, annotations, and searchable text.

Annotations and Markup

Annotations and Markup can show up in many stages of the EDRM. Annotations exist in a layer on top of the image or document and can infuse the document with additional information and encourage collaboration and communication. Sticky notes, arrows, and highlights can draw attention to important parts of the document.

Probably the most important annotations for the legal industry are bates stamps, redactions, and encrypt. Bates stamps have a long history in the courtroom of sequentially stamping paper documents for identification purposes. Digital ESI is no different and LEADTOOLS annotations make it easy to overlay consecutive numbers on documents as they get imported into an ECM or transferred into a new storage disk for Production. Redaction uses black rectangles to cover up sensitive information that should not be divulged in court. Digital redaction with annotations is beneficial since it doesn't require any changes be made to the original ESI. Redaction annotations can even be undone with passwords, providing different amounts of information based on a user's access tier. The encrypt object works similarly to redaction, but scrambles the image underneath of it rather than completely hiding it.

Bate Stamp Properties	Х
Properties	Preview Area
Stamps Container Stamps : Add Remove	000001 08/12/2016
Stamp Text 000001 09/12/2016	LEAD Technologies, lic. is a printelly held company founded in 1980 by More established to commence the second se
Change Font	Fast Facts Founded: LEAD was founded in 1990 and is a privately held company.
Changer on	Corporate Headquarters: LEAD is headquartered in Charlotte, North Carolina.
Alignment Horizontal Left Vertical Top V	Target Advince: Software developer and programmen: Vertred Industries Server, Compater noteware, Server, Tanvanon, Aurocaudos, IT, Expensenzy, Haelhocare, Banktourne, Automotive, Government, Millany, Education, Tealini, Faced and Dava, Rokandie perdoda.
Horizontal Left Vertical Top V	Developer Tools
Logo Properties Logo Image X Text Opacity 1 Angle 0 Postion	LEAD Technologies in the developer and publisher of LEATOOLS, the award- wrong line of development toolist. LEATOOCS is a tamy'of competentive toolist design of the programment integrate mater, decourse it, medical based of the programment integrate mater, and an operating integrate LEATOOLS are set development in equip the based in powerful integrate technology, offending development expositor for File Formats (1950), timping technology, offending development expositor for programment technology, offending development expositor, lange Database, Special Competencies, hunge Programment, Deatoon (LEATOOLS), based and Forms Recognition, Doornerer (Deano, yu, PDP). COR, Bercook, Database, and more. The result of 30 years of development, a LEATOOLS, tools the terrary tambit millions of lines of code part of application development. A LEATOOLS is not been terrarised to the state of the state been terrarise tools and the state of the
Stretch Logo X 0 Y 0 Width 0 Height 0	Imaging Applications/Utilities
Elements Add Number Add Date {{BatesNumber*6*1*1*1}} {{BatesDateTime*Local*MM/dd/yyyy}}	LEAD has also developed its over live of end-user imaginal applications and utilities. All applications were usult using LEAD/DISC SMox. The products regine from the award-winning LEAD/DISC Set Set for end conventing and protein driver to a a set of the set of
v	Fit O Normal
Ok	Cancel

Figure 5: LEADTOOLS Bates Stamp Properties Dialog

Virtual Printer

There are some instances where ESI is difficult to process because it is saved in a format only usable in a legacy application or custom-developed system. Thankfully, the vast majority of applications have some form of printing functionality, but many legal teams don't use this to their advantage. Virtual Printing, also known as print-to-file drivers, is an oftignored solution that helps prevent the hassle of printing to paper and then using a scanner to import it into an acceptable electronic format.

The LEADTOOLS Virtual Printer offers a unique catch-all solution by capturing the print job from any application and converting it to LEADTOOLS-supported formats, including searchable PDF, DOC, and TIFF. Adding annotations is also supported within the Virtual Printer, allowing developers to create workflows that can tackle several goals at once by both normalizing documents into appropriate ESI and bates stamping them along the way.

Conclusion

eDiscovery is a huge market for software developers, full of opportunities at any point of the legal process. Businesses both large and small have various requirements, from preventative and organizational measures, finding and analyzing ESI, creating ESI from paper, all the way to presenting their case in the courtroom. With excellent support for scanning, document cleanup, OCR, forms recognition, PDF, document conversion and viewing, annotations, and virtual printing, LEADTOOLS Document Imaging SDKs can handle every requirement software developers in the legal industry are regularly asked to implement.

LEADTOOLS offers an incredible value with its comprehensive family of toolkits for document, medical, multimedia, and raster imaging. For more information on how LEAD Technologies can image-enable applications and boost ROI, visit www.leadtools.com to download a free evaluation, or give us a call at +1-704-332-5532.

Sales: +1(704) 332-5532 sales@leadtools.com

Support: +1(704) 372-9681 support@leadtools.com

LEAD Technologies, Inc. 1927 South Tryon Street Suite 200 Charlotte, NC 28203



About LEAD Technologies

With a rich history of over 25 years, LEAD has established itself as the world's leading provider of software development toolkits for document, medical, multimedia, raster and vector imaging. LEAD's flagship product, LEADTOOLS, holds the top position in every major country throughout the world and boasts a healthy, diverse customer base and strong list of corporate partners including some of the largest and most influential organizations from around the globe.





© 2018 LEAD Technologies, Inc.