

Oracle Outside In Technology: Robust and flexible software components for electronic discovery solutions.

Oracle Outside In Technology is used by

- 8 of the top 10 EDD Service Providers
- 8 of the top 10 EDD Software Providers
- 8 of 10 leading Records Management Providers
- 5 of 5 leading Content Management Providers

The rapidly evolving electronic discovery market offers software developers significant opportunity. Innovative solutions abound and new ones enter the market at a rapid pace. Successful electronic discovery solutions share two characteristics: they adapt quickly to address new market requirements, and they overwhelmingly choose Oracle Outside In Technology for native document processing.

The Opportunity

The electronic discovery market is one of the most dynamic technology markets of this decade. With extremely large volumes of electronic information subject to discovery and the high stakes involved, it is not surprising that technology has evolved rapidly to provide electronic discovery solutions.

The numbers show that the market has quickly adopted these solutions. In 2002 the electronic discovery market was estimated at US\$270m; in 2007 it was estimated at nearly US\$3 billion. Market churn is significant with new vendors entering and, in some cases, raising the bar for those that achieved early leadership positions. Most recently, large platform players have begun the inevitable absorption of vertical solutions. It is an extremely competitive market where success requires ongoing innovation and rapid adaptation to new market requirements.

Oracle Outside In Technology Enables Market Leaders

Oracle Outside In Technology is not an electronic discovery product. It is an enabling technology for processing native files upon which electronic discovery applications and services are built. In fact, eight of the top 10 electronic discovery service providers and eight of the top 10 electronic discovery software providers use Oracle Outside In Technology in their solutions. These providers recognize that their competitive advantage lies not in the highly specialized business of dissecting native files, but rather in developing new and innovative products and services that address the overall electronic discovery workflow. Oracle Outside In Technology provides these leaders with critical native file processing capabilities that they have leveraged to build market-leading electronic discovery solutions.

Oracle Outside In Technology Supports Evolving Requirements

The updated Federal Rules of Civil Procedure, particularly Rules 26 and 34 that address electronically stored information (ESI), have created a stir in the market. Case law continues to refine their interpretation and application in litigation. Vendors in the electronic discovery space need to be highly flexible to respond quickly to these changing market requirements. As the leading technology provider to electronic discovery solutions, Oracle Outside In Technology faces the same challenges. Oracle addresses this demand for flexibility with the benefits outlined below.

Broad Access for the Best Coverage

Oracle Outside In Technology supports nearly 500 native file formats, from the first commercial word processing programs of the 1980's through the latest office productivity suites, PDF, email, instant messaging, and multimedia files. The product is updated frequently to support emerging formats. In fact, in the last year alone, Oracle added or updated support for over 25 formats.

Thorough Extraction for the Best Analysis

Virtually all electronic discovery projects require the extraction of metadata and text from native files. Metadata processing in particular has become a critical issue in electronic discovery, whether used at the onset to profile a document collection, to assess the scope of the discovery, or to uncover hidden information that may provide the proverbial smoking gun. Oracle Outside In Technology's extraction is extremely thorough, providing text (including the text of embedded objects), over 260 pre-defined metadata fields (including nearly 100 from email), limitless custom properties, and over 40 potentially hidden and sensitive metadata elements such as tracked changes, obsolete content, user names, routing slips, comments, author history, speaker notes, hidden slides, and hidden spreadsheet cells.

Numerous Conversion Options for the Most Flexibility

Whether a format for machine consumption (such as creating an index or load file), a visual format for human review, or a format for final production, Oracle Outside In Technology offers numerous output options including text, XML, HTML, TIFF, PDF, and native rendering via a viewer. In all cases, documents are preserved in their native format and can be processed without the potential for metadata spoilage that may occur when using native applications.

Highly Interoperable for the Most Value-Added

Oracle Outside In Technology is a family of interoperable software development kits (SDKs) that allows you to build powerful solutions by combining their functions. For example, text offset information generated during extraction can be used to create hit-highlights or redactions in the rendered output. Thorough documentation and sample code speed implementation, allowing developers to quickly build robust and unique solutions.

Outside In Technology in the Electronic Discovery Workflow

Electronic discovery providers need to add value and offer options at all stages in an electronic discovery workflow, enabling them to be a single source provider and to accommodate their customers' particular needs and contingencies. The Electronic Discovery Reference Model (EDRM) provides an excellent framework to discuss Oracle Outside In Technology's support for—and options available at—critical electronic discovery workflow stages.

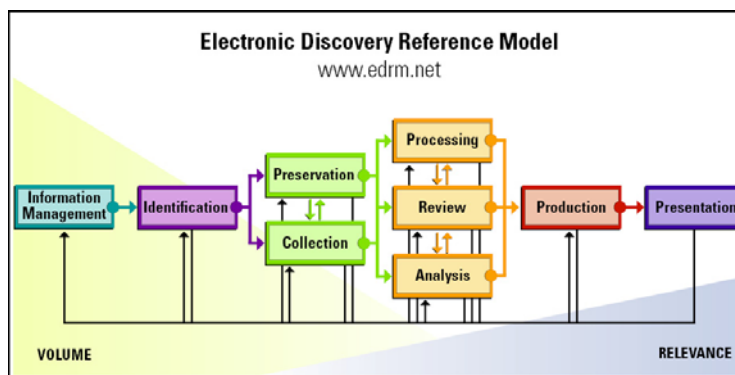


Figure 1. Oracle Outside In Technology supports each stage of electronic discovery.¹

Identification—ESI Scope

Oracle Outside In Technology's ability to extract text and metadata from native files plays a critical role in the identification phase. It can accurately identify file types, extract metadata and full text, and provide on-demand conversion for document sampling. Additionally, the search export engine can be run in "metadata only" mode to provide an efficient means to determine scope of the discovery based on key criteria such as author, date range, document type, and the "from" and "to" fields in emails.

Preservation—Protection From Destruction or Alteration

Oracle Outside In Technology supports the preservation phase in two important aspects. First, potentially relevant metadata is preserved since no native applications are invoked to access files' contents. Using the native applications could potentially raise spoliation issues. Second, document relationships in hierarchical file types—such as email archives—are preserved.

Collection—Acquisition of Data

Oracle Outside In Technology's ability to extract metadata and text from hundreds of native file formats helps ensure a thorough and accurate collection process while maintaining content integrity and preserving its native form. It provides electronic discovery vendors a springboard on which to add significant value in order to differentiate their offering.

Processing—Preparing Files for Review

Oracle Outside In Technology provides vendors with numerous options to process documents to efficiently reduce the overall file set to those that are responsive or worthy of further analysis. In all cases, the documents can remain in their native form in the event that they are ultimately required in the production phase. However, Oracle Outside In Technology can convert the files (either batch or on-demand) into more readily usable formats. The review itself can be done via a viewer, a browser, or Acrobat Reader—all without spoiling the metadata of the native file and without incurring the expense of a permanent conversion.

Review—Determining Relevance and Privilege

The interoperability between the SDKs enables hit-highlighting of the key terms from the identification and collection phases, or on a new set of terms that might indicate privilege. This can significantly expedite the human review process by providing the reviewer with a “short-cut” to key terms or sections within the document.

Analysis—Understanding ESI

Tools for unstructured data analysis are becoming increasingly sophisticated (i.e. clustering, guided navigation, visualization) and application of those technologies to electronic discovery can be a significant differentiator. Regardless of the analysis technology, they all require accurate text and metadata extraction from a broad range of input document types. Oracle Outside In Technology provides this output, as well as access to “hidden” data that might be overlooked or inaccessible via other technologies.

Production—Delivering ESI

The 2006 update to the Federal Rules of Civil Procedure directly addresses forms of production in electronic discovery. However, the Rules do not mandate a form, but rather indicate that the parties should agree on the form and, absent agreement, the Rules provide guidance for production. The bottom-line to electronic discovery vendors is that they should offer several options for production and be able to quickly and cost-effectively produce output in the form(s) most appropriate to the particular situation. Oracle Outside In Technology fully supports this approach with its many possible output formats, all of which can contain, or be linked to, metadata and full text. Further, both Oracle Outside In PDF Export and Oracle Outside In Image Export support document annotations and redactions. Finally, Oracle Outside In Technology does not make any alteration to the original native file so that it can be preserved and produced if required.

The Oracle Outside In Technology Product Family

Oracle Outside In Clean Content

Oracle Outside In Clean Content is tuned specifically for metadata and hidden information analysis. It provides a deep level of access to information in Microsoft Office and PDF documents. Oracle Outside In Clean Content enhances the discovery process by providing additional layers of information, such as comments, redlines, version history, Fast Save data, hidden text, author history, and custom properties, to name a few.

Oracle Outside In File ID

Oracle Outside In File ID identifies over 500 file types without using unreliable file extensions or MIME types. The technology is particularly useful at the start of an electronic discovery workflow when dealing with unknown data and when the format of a file needs to be quickly and accurately identified.

Oracle Outside In Content Access

Oracle Outside In Content Access extracts text and metadata and automatically translates the text and properties from multiple possible encodings into a single encoding specified by the developer. It is optimized for performance, interactively providing data to the host application in memory as the input file is processed. It is widely used in search and data forensic applications.

Oracle Outside In Search Export

Oracle Outside In Search Export provides the text, metadata, and pagination information of the supported file types as XML, HTML, or text specifically designed for search and forensic applications. It offers the developer a number of options, including a choice of four output formats: SearchML, SearchHTML, SearchText, or PageML. Its XML output can be transformed into another XML schema (such as EDRM XML) via XSLT.

Oracle Outside In HTML Export

Oracle Outside In HTML Export converts files into HTML via a sophisticated template architecture, as well as a rich set of options, to allow a high degree of customization of the HTML output. When used in combination with Oracle Outside In Content Access or Oracle Outside In Search Export, annotations (such as search hit-highlights and redactions) can be added to the HTML output programmatically. It provides browser access to hundreds of file types without plug-ins or other proprietary applications making it an ideal solution for web-based document review.

Oracle Outside In XML Export

Oracle Outside In XML Export converts and normalizes the content of supported file types into XML defined by Oracle's FlexionDoc schema. This schema uses XML tags that closely mirror the information in files created in popular business applications. In addition to the document content itself, FlexionDoc includes information related to document properties, structure, and formatting. The resultant XML can be consumed by an application or further transformed by the developer via XSLT into a developer-defined schema or an industry-defined schema such as EDRM XML.

Oracle Outside In PDF Export

Oracle Outside In PDF Export is a cross-platform, application-independent PDF conversion solution. It has its own rendering engine, so it does not require native applications to generate PDF. Its self-contained conversion simplifies workflow, decreases cost, and allows easy deployment on Windows, Linux, and Solaris. As with all Oracle Outside In Technology SDKs, it is interoperable with the others. For example, the text offset information from one of the search enabling SDKs can be used to apply hit-highlighting, annotations and/or redactions in the PDF output. It is especially valuable for processing and producing legacy document types for which the applications are unavailable or difficult to obtain.

Oracle Outside In Image Export

Oracle Outside In Image Export converts the supported file types into TIFF, JPEG, JPEG 2000, BMP, GIF, or PNG images. As with all the Oracle Outside In Technology SDKs, Oracle Outside In Image Export offers numerous options, including the ability to size the image output from thumbnail to full size and control image resolution. When used in combination with Oracle Outside In Content Access or Oracle Outside In Search Export, annotations and/or redactions can be added to the output programmatically. Results can be tailored for either screen or high-dots per inch-(DPI) print applications, and advanced smooth scaling techniques are used to optimize fidelity.

Oracle Outside In Viewer Technology

Oracle Outside In Viewer Technology renders high-fidelity representations of files without using the files' native applications. It enables client-side file viewing, printing, and copy/paste functionality of more than 400 file formats. An annotation API is also included, which allows the developer to highlight, hide, or insert text in a document view without altering the underlying native file. It is also available as an ActiveX control for Microsoft Visual Basic or other ActiveX-compliant development environments.

CONTACT US

For more information on Oracle Outside In Technology, call +1.800.ORACLE1 to speak to an Oracle representative or visit www.outsideinsdk.com.

Supported Formats

Access ▪ Adobe Acrobat (PDF) ▪ Adobe FrameMaker (MIF) ▪ Adobe FrameMaker graphics (FMV) ▪ Adobe Illustrator ▪ Adobe Photoshop (PSD) ▪ Ami Draw/Ami Draw (SDW) ▪ ANSI Text ▪ ASCII Text ▪ AutoCAD Drawing ▪ AutoCAD Interchange and Native Drawing formats ▪ AutoShade Rendering (RND) ▪ Binary Group 3 Fax ▪ Bitmap (BMP, RLE, ICO, CUR, OS/2 DIB & WARP) ▪ CALS Raster (GP4) ▪ Computer Graphics Metafile (CGM) ▪ Corel Clipart format (CMX) ▪ Corel Draw (CDR with TIFF header) ▪ Corel Draw (CDR) ▪ Corel/Novell Presentations ▪ DataEase ▪ dBASE ▪ dBase ▪ dBase ▪ DEC WPS Plus (DX) ▪ DEC WPS Plus (WPL) ▪ DisplayWrite 2 & 3 (TXT) ▪ DisplayWrite 4 & 5 ▪ EBCDIC ▪ Enable ▪ Encapsulated PostScript (EPS) ▪ Executable (EXE, DLL) ▪ Executable (Windows) NT ▪ EXIF Metadata ▪ First Choice ▪ First Choice ▪ First Choice ▪ FoxBase ▪ Framework ▪ Framework ▪ Framework ▪ Freelance (Windows) ▪ Freelance for OS/2 ▪ GEM Paint (IMG) ▪ Graphics Environment Mgr (GEM) ▪ Graphics Interchange Format (GIF) ▪ GZIP ▪ Hangul ▪ Harvard Graphics (Windows) ▪ Harvard Graphics for DOS ▪ Hewlett Packard Graphics Language (HPGL) ▪ HTML ▪ IBM FFT ▪ IBM Graphics Data Format (GDF) ▪ IBM Picture Interchange Format (PIF) ▪ IBM Revisable Form Text ▪ IBM Writing Assistant ▪ Initial Graphics Exchange Spec (IGES) ▪ JFIF (JPEG not in TIFF format) ▪ JPEG (including EXIF) ▪ JustSystems Ichitaro ▪ JustWrite ▪ Kodak Flash Pix (FPX) ▪ Kodak Photo CD (PCD) ▪ Legacy ▪ Lotus 1-2-3 (DOS & Windows) ▪ Lotus 1-2-3 (OS/2) ▪ Lotus 1-2-3 Charts (DOS & Windows) ▪ Lotus 1-2-3 for SmartSuite ▪ Lotus AMI/AMI Professional ▪ Lotus Manuscript ▪ Lotus PIC ▪ Lotus Snapshot ▪ Lotus Symphony ▪ Lotus Word Pro ▪ LZA Self Extracting Compress ▪ LZH Compress ▪ Macintosh PICT1 & PICT2 ▪ MacPaint (PNTG) ▪ Macromedia Flash (text) ▪ MacWrite II ▪ MASS11 ▪ Micrografx Designer (DRW) ▪ Micrografx Designer (DSF) ▪ Micrografx Draw (DRW) ▪ Microsoft Binder ▪ Microsoft Excel (Mac) ▪ Microsoft Excel (Windows) ▪ Microsoft Excel Charts ▪ Microsoft Multiplan ▪ Microsoft Outlook Express (EML) ▪ Microsoft Outlook Folder (PST) ▪ Microsoft Outlook Message (MSG) ▪ Microsoft Outlook Offline Folder (OST) ▪ Microsoft PowerPoint (Mac) ▪ Microsoft PowerPoint (Windows) ▪ Microsoft Project ▪ Microsoft Rich Text Format (RTF) ▪ Microsoft Word ▪ Microsoft Word ▪ Microsoft Word (Mac) ▪ Microsoft WordPad ▪ Microsoft Works ▪ Microsoft Works ▪ Microsoft Works (DOS) ▪ Microsoft Works (DOS) ▪ Microsoft Works (Mac) ▪ Microsoft Works (Mac) ▪ Microsoft Works (Mac) ▪ Microsoft Works (Windows) ▪ Microsoft Works (Windows) ▪ Microsoft Write ▪ MIME Text MailMosaic Twin ▪ MP3 (ID3 metadata) ▪ MultiMate ▪ Navy DIF ▪ Nota Bene ▪ Novell Perfect Works ▪ Novell Perfect Works ▪ Novell PerfectWorks (Draw) ▪ Novell WordPerfect ▪ Novell WordPerfect ▪ Novell/Novell WordPerfect ▪ Office Writer ▪ OpenOffice Calc ▪ OpenOffice Draw ▪ OpenOffice Impress ▪ OpenOffice Writer ▪ OS/2 PM Metafile (MET) ▪ Paint Shop Pro 6 (PSP) ▪ Paradox (DOS) ▪ Paradox (Windows) ▪ PC Paintbrush (PCX and DCX) ▪ PC-File Letter ▪ PC-File+ Letter ▪ Personal R:BASE ▪ PFS: Professional Plan ▪ PFS:Write ▪ Portable Bitmap (PBM) ▪ Portable Graymap (PGM) ▪ Portable Network Graphics (PNG) ▪ Portable Pixmap (PPM) ▪ Postscript (PS) ▪ Professional Write ▪ Progressive JPEG ▪ Q&A ▪ Q&A Write ▪ Quattro Pro (DOS) ▪ Quattro Pro (Windows) ▪ R:BASE 5000 ▪ R:BASE System V ▪ Reflex ▪ Samna Word ▪ SmartWare II ▪ SmartWare II ▪ SmartWare II ▪ Sprint ▪ Star Office Draw ▪ StarOffice Calc ▪ StarOffice Impress ▪ StarOffice Writer ▪ Sun Raster (SRS) ▪ SuperCalc 5 ▪ Text Mail (MIME)TIFF ▪ TIFF CCITT Group 3 & 4 ▪ Total Word ▪ Truevision TGA (TARGA) ▪ Unicode Text ▪ UNIX Compress/UNIX TAR/UNIX encodevCard ▪ Visio ▪ Visio (preview) ▪ Volkswriter 3 & 4 ▪ VP Planner 3D ▪ Wang PC (IWP) ▪ WBMP ▪ Windows Enhanced Metafile (EMF) ▪ Windows Metafile (WMF) ▪ WML ▪ WordMARC ▪ WordPerfect Graphics (WPG & WPG2) ▪ WordStar ▪ WordStar ▪ WordStar 2000 ▪ XML ▪ X-Windows Bitmap (XBM) ▪ X-Windows Dump (XWD) ▪ X-Windows Pixmap (XPM) ▪ XyWrite ▪ Yahoo! Instant Messenger Archive ▪ ZIP