

Oracle Outside In Technology



Outside In Technology enhances hundreds of software applications from the world's leading technology companies. Outside In Technology provides software developers with a comprehensive solution to extract, normalize, convert and view the contents of over 600 unstructured file formats. From standard business documents to specialty formats and legacy files, Oracle Outside In Technology turns unstructured files into accessible information.

THE MOST COMPLETE SOLUTION FOR UNSTRUCTURED INFORMATION

- Extract content from over 600 unstructured file formats
- Normalize to output of choice—HTML, XML, Text for further processing and analysis
- Convert to display output of choice—HTML, HTML5, PDF, Image, Viewer

SOFTWARE DEVELOPMENT KITS

- Web View Export (*New – August 2014*)
- Clean Content
- File ID
- Content Access
- Search Export
- HTML Export
- XML Export
- PDF Export
- Image Export
- Viewer

Outside In Software Development Kits

Oracle Outside In Technology is organized into ten software development kits (SDKs). Each SDK provides a particular solution for extracting, converting and viewing files, but they are all highly flexible and interoperable. Developers can quickly implement any combination of Outside In SDKs to provide exactly the right functionality in their application while minimizing integration effort and code footprint. The SDKs offer a wide range of options to give developers programmatic control of their workflow and output. Thorough documentation and sample applications with source code are included to further accelerate implementation.

Outside In Web View Export

Outside In Web View Export is a software developer's kit (SDK) that enables an application to produce high quality HTML5 renditions of documents created by standard business software. Unlike other document-to-HTML products, the output that is produced by Web View Export mimicks the look and feel of the native applications in which the documents are created. HTML5 and CSS3 allow for a high level of fidelity, and a robust Javascript API allows the Web View Export output to be interactive and customizable. This "web view" requires an HTML5-capable browser, but nothing else. It offers C, Java and .NET APIs and is available on 64-bit Windows and Linux.



MARKETS SERVED

- Content Management
- Compliance
- Customer Relationship Management
- Data Leak Prevention
- Digital Asset Management
- Electronic Discovery
- Forensics
- Email
- Knowledge Management
- Legal
- Portal
- Product Lifecycle Management
- Publishing
- Records Management
- Search
- Security
- Social Collaboration
- Text Analytics

Outside In Clean Content

Clean Content addresses particularly challenging issues in native file processing. Focusing specifically on widely used formats (Microsoft Office and PDF), it extracts all text, properties, hidden information and system data embedded in native files. Clean Content can also programmatically modify native files, enabling features such as scrubbing, property modification and document assembly. Clean Content is a native Java technology that offers Java, C/C++ and .NET APIs.

Outside In File ID

File ID identifies over 600 file types without using unreliable file extensions or MIME types. The technology is particularly useful at the start of a workflow when dealing with unknown data and when the format of a file needs to be quickly and accurately identified. Available as a C-API.

Outside In Content Access

Content Access extracts text and metadata from over 600 file types and interactively provides the extracted data to the host application in memory as the input file is processed. It also provides a mechanism to extract embeddings and archive objects (ZIPs, PSTs, etc.) and save them as standalone files. The technology is widely used in search, data forensics, and security applications. Available as a C-API..

Outside In Search Export

Search Export provides the text, metadata, and pagination information of the supported file types in XML, HTML, or text specifically designed for search and forensic applications. It also provides a mechanism to extract embeddings and archive objects (ZIPs, PSTs, etc.) and save them as standalone files. Search Export offer multiple output formats including XML, light-weight HTML and text. Available with C, Java and .NET APIs.

Outside In HTML Export

HTML Export converts the contents of over 600 file formats into HTML, rendering embedded graphics as a choice of GIF, JPEG, or PNG. It uses sophisticated template architecture, as well as a rich set of options, to allow a high degree of customization of the HTML output. HTML Export can use cascading style sheets for more complex formatting, or alternatively can be scaled down to use least-common-denominator versions of HTML for the broadest compatibility with browsers on various platforms or devices. Its ability to provide browser access to hundreds of file types without plug-ins or other proprietary applications makes it one of the most widely used SDKs of the Oracle Outside In Technology suite. Available with C, Java and .NET APIs.

Outside In XML Export

XML Export converts and normalizes the content of supported file types into XML that explicitly describes all the elements of the document's content, structure, properties, and formatting. The resultant XML can be consumed by an application or further

transformed by the developer via XSLT into a developer-defined schema. Oracle Outside In XML Export is appropriate for use in any application that can benefit from normalizing documents into robust XML. Available with C, Java and .NET APIs.

Outside In PDF Export

PDF Export provides cross-platform, application-independent PDF conversion. It uses the Outside In proprietary rendering engine, which eliminates any need for native applications resulting in simplified work flow and reduced cost. PDF Export can embed and sub-set fonts and can create a single PDF output file from multiple native input files. When used in combination with Content Access or Search Export, annotations or redactions can be added to the PDF output programmatically. PDF Export can generate fully compliant PDF/A-1a and A-2a (Archive) PDF, which is a significant added value for applications in the records management, archive and compliance markets. Available with C, Java and .NET APIs.

Outside In Image Export

Image Export converts the supported file types into TIFF, JPEG, JPEG 2000, BMP, GIF, or PNG images. It offers numerous options, including the ability to size the image output from thumbnail to full size and to control image resolution. When used in combination with Content Access or Search Export, annotations can be added to the output programmatically. Image Export is used extensively in e-discovery and in other applications that require high fidelity, static images of business documents. Available with C, Java and .NET APIs.

Outside In Viewer

Outside In Viewer displays high-fidelity representations of files without using the files' native applications. It provides cross-platform, client-side file viewing, printing, and copy/paste functionality of over 600 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.

CONNECT WITH US



Hardware and Software, Engineered to Work Together

Copyright © 2014, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group. 0814



Oracle is committed to developing practices and products that help protect the environment