Oracle® Universal Content Management
Product Overview
10*g* Release 3 (10.1.3.3.0)

March 2007

ORACLE

Product Overview, 10*g* Release 3 (10.1.3.3.0)

Contributing Authors: Sandra Christiansen

Table of Contents

Chapter 4: Web Content Management

Chapter 5: Records Management

Index

Chapter

1

INTRODUCTION

OVERVIEW

Oracle® Universal Content Management 10gR3 is made up of a variety of features, which are described in this document. These features can be further divided in the following manner:

- Document Management (page 2-1)
- Digital Asset Management (page 3-1)
- Web Content Management (page 4-1)
- Records Management (page 5-1)

This section contains the following topics:

ABOUT THIS GUIDE

This guide describes the feature set in Oracle® Universal Content Management .

AUDIENCE

This guide is intended for system integrators and administrators.

CONVENTIONS

The following conventions are used throughout this guide:

- The notation *<Install_Dir>/* is used to refer to the location on your system where the content server instance is installed.
- Forward slashes (/) are used to separate the directory levels in a path name. A forward slash will always appear after the end of a directory name.
- Notes, technical tips, important notices, and cautions use these conventions:

Symbols	Description
	This is a note. It is used to bring special attention to information.
	This is a technical tip. It is used to identify information that can be used to make your tasks easier.
	This is an important notice. It is used to identify a required step or required information.
	This is a caution. It is used to identify information that might cause loss of data or serious system problems.

Chapter

2

DOCUMENT MANAGEMENT

OVERVIEW

Document management features are the base of Oracle® Universal Content Management .

This section contains these topics:

CONTENT SERVER

Content Server is the foundation for a variety of Oracle content management products. It provides a flexible, secure, centralized, web-based repository that manages all phases of the content life cycle: from creation and approval to publishing, searching, expiration, and archival or disposition. Every contributor throughout the organization can easily contribute content from native desktop applications, efficiently manage business content via rich library services, and securely access that content anywhere using a web browser.

All content, regardless of content type, is stored in the web repository or database for management, reuse and access. While stored in the repository, all types of content—ranging from e-mail, discussions, documents, reports, spreadsheets and records to images, multimedia or other digital formats—receive the same set of fundamental core services.

Components

A number of components, providing advanced functionality are provided with Content Server. These components may be rolled into the core, available to choose upon installation, or available in an “extras” directory on the CD.

Desktop

Desktop Integration Suite provides a set of embedded applications that help you seamlessly integrate your desktop experience with Content Server. More specifically, it provides convenient access to the content server from Microsoft Windows Explorer, desktop applications like Microsoft Word and Excel, and e-mail clients like Microsoft Outlook and Lotus Notes.

As a result, you can easily manage files in the content server and share files with users directly from your desktop instead of logging onto the content server and using a web browser.

Dynamic Converter

Dynamic Converter is transformation technology and on-demand publishing solution for critical business documents. With Dynamic Converter, you can easily convert any business document into a Web page for a specified audience without use of the application used to create that document. The benefits are immediate—information can be exchanged freely without the bottleneck of proprietary applications.

When a Web browser first requests a document, a set of rules are applied to determine how that document should appear as a Web page. These rules can be defined in a template, a core component of Dynamic Converter.

Dynamic Converter offers a number of benefits to the user:

- Business documents can be easily viewed in a Web browser.
- Native applications (such as Adobe Acrobat, Microsoft Word, etc.) are not required.
- Multiple renditions of a document are available for different devices (Web browsers, wireless devices, etc.)
- Templates are interchangeable with Content Publisher.
- Numerous business document types, including legacy formats, are supported.

Content Integration Suite

Content Integration Suite (CIS) enables communication with Content Server and Image Server and is deployable on a number of J2EE application servers, in addition to working in non-J2EE environments. CIS provides an interface to the fine-grained services of the servers through an object-oriented services API called the Universal Content and Process Management API (referred to as the UCPM API).

This product includes an SCS Command Layer and information about extending the SCS Command Layer. The SCS Command Layer is a set of Command Objects, implemented as Java Beans, that are built to communicate to Content Server through the content server's IDC Command interface. The SCS Command Layer is based on the J2EE “Command Design Pattern,” which means that all business logic is wrapped into a business logic bean (called a “command”) and sent to an invoker object, often implemented as a stateless session bean in a J2EE environment, for execution.

CONTENT CATEGORIZER

Content Categorizer provides organizations the capability to use one or more taxonomies within Content Server. In addition to its out-of-the-box categorization tools and functionality, Content Categorizer provides an open API for third-party categorization engines. With this open architecture, users can take advantage of the rule sets and taxonomies provided by third-party categorization tools. As a result, organizations can choose the categorization engine that best fits their business needs. For example,

organizations can use their existing vertical industry taxonomy to organize their managed content into specific categories and subcategories.

Content Categorizer enables administrators and content contributors to automatically, uniformly and intelligently categorize content as it is checked into the Content Server. Perfect for loading large amounts of existing content into Content Server, Content Categorizer can be used in batch-mode—freeing administrators of the responsibility of assigning metadata to each individual content item. End-users, on the other hand, will appreciate Content Categorizer for its ability to suggest appropriate metadata as they check in new pieces of content.

To suggest a category or specific value for each of the Content Server metadata fields, Content Categorizer uses a set of rules to analyze content items. Some automatic categorization rules include: direct correlation using file properties or text references; score computations based on word matching; sentence or paragraph summaries automatically pulled from files; or certain language pattern recognition. Rule sets can also be used from other third party categorization engines.

Content Tracker

Content Tracker allows website administrators to use key metrics to define reports for analyzing site traffic and usage of Content Server. Content Tracker's tight integration with Universal Content Management makes it possible to run reports based on specific user profile information, groups of users, or on any set of content that can be defined by a query or group of metadata values. This flexibility allows for extremely granular site and content analysis, giving web teams crucial information to more strategically manage and improve upon their site.

Content Tracker uses data gathered in the web server log files, web filter log files, and Content Server's database tables to generate information regarding the content items accessed. This information could include metadata, user profile data as well as information from the users themselves. Content Tracker then populates this information into database tables in any RDBMS database such as Oracle or Microsoft SQL Server. Once this database is populated, customized reports can be generated according to the needs of the user.

Inbound Refinery

Inbound Refinery is an add-on module to Content Server that manages all file conversions at the input side of Content Server (hence “Inbound”), and also provides thumbnailing functionality. Files are converted upon check-in of the content into Content Server.

Inbound Refinery includes Outside In Image Export, which can be used for the following:

- To create thumbnails of files checked into Content Server. Thumbnails are small preview images of content. Outside In Image Export can also be used to create thumbnails of PDF files generated by PDF Converter.
- To convert files checked into Content Server to multi-page TIFF files as the primary web-viewable rendition.

In addition to the conversions that Inbound Refinery can perform using Outside In Image Export, several conversion add-ons are available for purchase and use with Inbound Refinery. The additional types of files that Inbound Refinery can convert, and the result of each conversion, depend on the conversion add-ons that are installed on the Inbound Refinery computer.

Basic Refinery Process

When a file is checked into Content Server, a copy of the native file is stored in the native file repository (the ‘*vault*’ directory). The native file is the format in which the file was originally created (for example, Microsoft Word).

If the file format is set up to be converted, it is placed in a queue for further processing. At set intervals, Inbound Refinery checks the queue. If a file is present, Inbound Refinery calls the appropriate conversion add-on to perform the actual conversion. The exact conversion process depends on how Inbound Refinery is set up. In some situations, conversion may be done entirely in the background, with no noticeable interaction. In others, a file may be opened in its native application and printed to a PostScript file, which is subsequently translated into a different format (for example, a PDF file). In that case, windows are automatically opened and closed on the Inbound Refinery machine. The converted file (for example, a web-viewable PDF file) is then copied to the web-viewable file repository (the ‘*weblayout*’ directory). Users can then view the file through their web browser.

If the file format is not set up to be converted (or if the conversion fails), no web-viewable file is created and a copy of the native file is placed in the *weblayout* directory. This means

that the file is passed through to the library in its native format. Users must then have the native application installed on their computer to view the file.

PDF Converter

PDF Converter enables the automatic publishing of native content items to web-viewable PDF (Portable Document Format) files. A PDF rendition of the native format is immediately generated upon check-in of new content into the content server. This PDF rendition allows web viewing of that content item without requiring users to install native applications. PDF Converter converts over 35 files formats to PDF, such as Adobe Framemaker, Illustrator, InDesign, PageMaker, and Photoshop as well as Hangul, JustSystems Ichitaro, Lotus Smartsuite, Microsoft Office, Microsoft Visio, OpenOffice, and Sun StarOffice.

PDF Converter optimizes non-optimized PDF files and also processes links—for example, Microsoft Word hyperlinks, ‘mailto’ links, and table-of-contents links.

You can use the following PDF-related products and add-ons to further extend the PDF files generated by PDF Converter:

XML Converter

XML Converter gives XML-based access to information that is trapped in unstructured business content. With XML Converter, content contributed to Content Server is converted to XML at the time of check-in. XML Converter converts over 225 document types and supports the leading word processing formats, such as Microsoft Word, Lotus WordPro, and Corel WordPerfect. It also includes support for popular spreadsheet, presentation, and graphic formats.

When a new content item is checked into the content server, XML Converter converts content to either a SearchML or FlexionDoc format. FlexionDoc is very verbose and captures all information, including attributes such as styles in a Microsoft Word document. From there, administrators have the ability to check in different XSL files that would then convert the SearchML or FlexionDoc document to any XML format. Additionally, administrators have the option to use a DTD (Document Type Definition) to validate the XML generated from the XSL transformation. If there is an error in conversion, all the relevant documents (for example, the original SearchML or FlexionDoc file, the XML file generated after XSL transformation, and the error report) are all checked in and can be sent through a workflow for the developer/administrator to debug.

Since the XML file is stored and managed within the web-based repository, it can be accessed from any location using a web browser—making it available for other enterprise applications, data exchange, re-use and further conversion into additional formats. XML Converter provides an out-of-the box XML solution with enterprise level performance, while ensuring compatibility with the W3C standards specifications.

For dynamic rendering of XML files to HTML, we recommends Dynamic Converter (page 2-2).

Tiff Converter

Tiff Converter enables organizations to check TIFF (Tagged Image File Format) files into Content Server and then web-publish these as multiple-page PDF files.

Tiff Converter uses either CVISION CVista PdfCompressor or Adobe Acrobat Capture to convert single-page TIFF files, multiple-page TIFF files, or zip files containing multiple TIFF files (TIFZ, TIZ or ZIP file extensions) to a single PDF file. Additionally, during the TIFF to PDF conversion, Optical Character Recognition (OCR) is performed—enabling users to perform full-text searches of managed TIFF files in Content Server.

Tiff Converter is very useful in scanning applications. It automatically converts TIFF images to PDF format upon check-in to Content Server for easy viewing of legacy content. The conversion of the released TIFF image can occur on the client side or server side and uses the Adobe Acrobat Capture technology. With TIFF Converter, users can easily view and access managed legacy content (scanned documents) via a browser with the Adobe Reader.

Sharepoint Integration

The Sharepoint Integration enables users to interact with Content Server using Microsoft's SharePoint technology by using web parts. Two primary use cases exist for the combination of Content Server Web Parts and SharePoint:

- Using SharePoint to collaborate on documents then 'promoting' them to Content Server for storage or use on the enterprise system. This allows users to manage content in Sharepoint, including workflows, but use Content Server for storage and retention.

- Using a SharePoint front end but storing the content behind SharePoint in a Universal Content Management (UCM) system. This allows users who are familiar with SharePoint to use that interface for many tasks.

This release of Content Server Web Parts allows a SharePoint server to run using Content Server WebParts to provide a Content Server-backed SharePoint server. The web parts also provide for the use of both a SharePoint content repository and a repository. Content items can be moved between the two systems. This allows consumers who are familiar with SharePoint to interact with the Content Server through a known interface, thus making it easier to contribute and search for content.

VERITY INTEGRATION

With Oracle® Universal Content Management , you have a Verity Integration add-in module for Content Server.The Verity VDK6 add-on provides an alternative content search and retrieval solution when installed with Content Server. Verity provides metadata and full-text indexing and search capabilities, which means every word in a file is indexed, not only its metadata, and all the information can be searched. To use the Verity Integration component, you must install and configure the component. See the *Verity Integration Guide* for more information.

Chapter

3

DIGITAL ASSET MANAGEMENT

OVERVIEW

This section covers the following topics:

IMAGE MANAGER

Oracle Image Manager enables users to quickly find, group, and download images of various sizes and resolutions. For example, an organization's logo may need to be available in a variety of sizes for advertisements, web pages, and presentation. At check-in, the image is automatically converted into the defined formats and sizes. Users can then search for the image using standard metadata, group renditions into a content basket, and download a single compressed file of the image renditions they need.

Renditions

Image Manager installs six pre-defined rendition sets for use with Image Alchemy:

- DefaultGraphicSet - Required Default Rendition Set
- CorporateImage - Common Corporate Format Requirements
- ProductCatalog - CMYK HiRes and Proof Images for Print

- DigitalPhoto - Process Digital Photos
- WebImages - Renditions for web applications and web images
- Print - Renditions for productivity and layout applications

VIDEO MANAGER

Video Manager enables users to quickly find, group, and download videos of various sizes and resolutions. For example, a company training video may need to be available in a variety of sizes for streaming on an intranet, presenting to an audience, or copying to tape. At check-in, the video is automatically converted into the defined formats and sizes. Users can then search for the video using standard metadata, group renditions into a content basket, and download a single compressed file of the video renditions they need.

Content Basket

Content Basket enables you to quickly find, group, and download multiple content items. For example, you may need to send an organization's logo along with several articles for your newsletter to a printing agency, or e-mail a group of documents to a vendor. With Content Basket, you can select items to add to your content basket from any content information or search results page. Once an item is added to your content basket, you can access the content basket from the My Content Server tray and download a single compressed file of the content items you need. Content Basket works with both Image Manager and Video Manager

Chapter

4

WEB CONTENT MANAGEMENT

OVERVIEW

This section covers the following topics:

CONTENT PUBLISHER

Content Publisher is an extremely powerful tool that provides advanced template-based technology to automatically publish over 225 file formats as well-designed, fully linked websites in HTML, XML, cHTML, or WML format. Content may be managed in a Product_Name content server or stored in a standard file system. Features include single-source publishing, abstraction, personalization, scripting language inclusions, content release scheduling, and automated creation of navigation.

Content Publisher consists of two main components: Site Builder and Site Server. With Site Builder, you create project files to capture all the information necessary to build web publications. You can then either translate and publish manually in Site Builder or use Site Server to automate the process. If you use Content Publisher with Oracle Content Server, you can draw source files from a content server and use those to create the web output.

Prior to publication to a live site, new web content can be published to a staging site and routed through workflows for approval.

Templates and Abstraction

Content Publisher automates the publishing process using two key technologies: templates and abstraction. A translation template takes content and automatically translates it into standard HTML web pages without altering the source files. The abstraction feature automatically takes content that has been created from many different sources, and intelligently determines what the content elements are—such as title, body, text, and headings. Content Publisher then puts this information into the translation templates, converting the data into professional web pages.

Tech Tip: Site Server can be run as a service in Microsoft Windows.

CONNECTION SERVER

The Connection Server system automates the distribution, transformation, and management of digital content over the Internet. The Connection Server system provides an efficient means for communications between the distributors of content, called content providers, and the consumers of content, called subscribers. Content providers and subscribers agree on the information to be distributed, and on its availability and distribution timing. Content providers then use Connection Server to distribute content that they control to their authorized subscribers, in the form of packages called offers.

Content distribution can take place between businesses, or within a single organization. For example, a news agency can distribute content for use by its affiliates; or a bank can distribute application-software upgrades to its branches. Because the Connection Server system transmission is content-insensitive, any kind of digital information can be distributed.

The Connection Server system uses standard industry protocols such as Information and Content Exchange (ICE), Extensible Markup Language (XML), and the Hypertext Transfer Protocol (HTTP). The Connection Server software is built using the Java programming language and is designed as a stand-alone server.

The Connection Server system also provides the following services:

- Automatically assigns users UUIDs (Universal Unique IDentifiers) and passwords to help ensure the secure distribution of content.
- Keeps track of distribution relationships between content providers and subscribers.

SITE STUDIO

Site Studio is the application most often used for designing websites. It offers site developers and designers a built-in methodology for building websites as well as a customizable library to easily reuse custom code and fragments—enabling them to create and deploy robust websites quickly. Additionally, Site Studio enables companies to design and develop websites for dynamic contribution and viewing, as well as publish these sites as static representations built and delivered with HTML code on standard web servers.

Some of the website creation and contribution best practices that are automated and enforced with Site Studio include:

- In-context contribution and updates directly from the website
- Hierarchical website structure and navigation
- Template-based pages comprised of multiple regions
- Separation of content and presentation
- Separation of navigational structure and its presentation
- Secure, regional-level content authoring and editing
- WYSIWYG XML-based contribution forms
- Reusable content and XML-based fragments
- Single-source content management
- Multi-site management

Organizations may also prefer to build websites using Site Studio if they would like to distribute website development to multiple site designers, yet still have a centralized team of developers maintain control over the brand and look-and-feel. Site Studio gives developers the ability to create customizable libraries to provide site designers with reusable drag-and-drop layouts, fragments, navigation, and code that integrates with back-end applications for developing their own unique websites. With Site Studio's customizable library as well as its built-in framework for designing websites, these site designers can create robust and well-architected sites with little or no knowledge of

HTML or other programming languages. Learn more about our approach to multi-site management.

SITE STUDIO PUBLISHING UTILITY

The Site Studio Publishing Utility enables you to publish a Site Studio web site from a Content Server environment to a pure web server environment that is not running Content server (that is, one running Microsoft IIS, Apache, etc.).

The Publishing Utility creates a static snapshot of a dynamic site by traversing all the links in a web site (visiting all of the linked pages) and downloading a copy of each page and all of the resources (images, flash movies, etc.) on each page. Your entire web site (including the content of queries, layout pages, fragments, contributor data files, and native documents) will be copied and published to the new server.

Content Portal Suite

Content Portlet Suite is built on top of the Content Integration Suite and offers a number of pre-built reference portlets for the BEA WebLogic, IBM WebSphere, Plumtree, and Sun ONE portal servers. Content Portal Suite provides access to content stored in the content server, enabling users to update, search, and view portal content in a way that is efficient and easy to use. Oracle provides the ability to manage the content creation and distribution process through the use of portlets. These portlets can be enabled for different users based upon the user's roles and permissions within the organization. Depending on the permission level, the user may be allowed to browse or search content, contribute a new content item and view the progress of their content through workflow.

Integrating your portal and Product_Name with Content Portal Suite provides an easier way to keep the portal up to date—resulting in greater utilization, lower maintenance costs and a larger your return on your portal investment.

Content Portlet Suite for WebLogic

Content Portlet Suite (CPS) for WebLogic provides a reliable and scalable integration with BEA WebLogic Portlet Server. This suite leverages the Content Integration Suite (CIS) as the foundation layer for integration with WebLogic. The suit provides eight functional reference portlets (Library, Authenticated Library, Basic Search, Authenticated Search, Saved Search, Contribution, Workflow Queue, and Metadata Administration)

which can be used immediately, or as examples of how to implement WebLogic portlets with Content.

Content Portlet Suite for WebSphere

Content Portlet Suite (CPS) for WebSphere provides a reliable and scalable integration with IBM WebSphere Portlet Server. This suite leverages the Content Integration Suite (CIS) as the foundation layer for integration with WebSphere. The suite provides eight functional reference portlets (Library, Authenticated Library, Basic Search, Authenticated Search, Saved Search, Contribution, Workflow Queue, and Metadata Administration) which can be used immediately, or as examples of how to implement WebSphere portlets

Content Portlet Suite for Plumtree

Content Portlet Suite (CPS) for Plumtree provides a reliable and scalable integration with Plumtree Portlet Server. This suite leverages the Content Integration Suite (CIS) as the foundation layer for integration with Plumtree. The suite provides eight functional reference portlets (Library, Authenticated Library, Basic Search, Authenticated Search, Saved Search, Contribution, Workflow Queue, and Metadata Administration) which can be used immediately, or as examples of how to implement Plumtree portlets with CIS.

Content Portlet Suite for Sun ONE

Content Portlet Suite (CPS) for Sun ONE provides a reliable and scalable integration with the Sun ONE Integration Server. This suite leverages the Content Integration Suite (CIS) as the foundation layer for integration with Sun ONE. The suite provides eight functional reference portlets (Library, Authenticated Library, Basic Search, Authenticated Search, Saved Search, Contribution, Workflow Queue, and Metadata Administration) which can be used immediately, or as examples of how to implement Sun ONE portlets with CIS.

Chapter

5

RECORDS MANAGEMENT

OVERVIEW

With Oracle® Universal Content Management , you have a Records Manager adapter for Content Server or Records Manager Corporate edition.This section covers the following topics:

RECORDS MANAGER ADAPTER FOR CONTENT SERVER

The Content Server URM adapter provides the bridge between URM, which manages the retention policies and the Content Server system, which stores the content. The Content Server URM adapter also sends information back to the URM server, so it can maintain an up-to-date catalog of the enterprise's important content. Thus, organizations can apply their retention policies to more content, more consistently, with less administrative effort, and less disruption for users. These same benefits apply to litigation searches and holds.

RECORDS MANAGER: CORPORATE EDITION

Records Manager Corporate Edition effectively manages content items on a retention schedule. The focus of retention management of content items tends to be the ***scheduled elimination of content*** in which the costs of retaining content outweighs the value of keeping it.

Index