

Taking Control of Your Content: The Oracle Content Management Software Development Kit

An Oracle White Paper
June, 2002

Taking Control of Your Content: The Oracle Content Management Software Development Kit

EXECUTIVE OVERVIEW

Today's businesses must manage large amounts of data stored in a wide range of formats. Taken as a whole, this unstructured content is a valuable resource in today's information economy. Businesses need to make that information accessible in a timely, secure, and reliable way.

This white paper describes how your organization can use the Oracle Content Management Software Development Kit (Content Management SDK) to take control of your content and:

- Store any type of content in a single repository.
- Access that content using familiar interfaces and protocols.
- Customize the system to manage the life cycle of your content.
- Automate business processes.
- Organize the content for quick retrieval.
- Share the content with those who need it..

MANAGE CONTENT IN A SINGLE REPOSITORY

For years, conventional wisdom held that it was cheaper and more efficient to store content in many, small file servers rather than one enormous centralized mainframe database. At first, this model was successful. So successful, in fact, that individual databases, file servers, mail servers, and web servers appeared to sprout up through spontaneous generation, their numbers growing exponentially throughout the 80s and 90s. Today, there's a wealth of available data, but it's isolated in individual islands of information, scattered throughout the enterprise.

Progressive companies now realize that the information isn't useful unless it's available where you are, when you need it. The trend now is toward centralization, using fast central servers supported by middle-tier machines that handle the global traffic as people request and receive the information they need to remain competitive in the world arena. An Oracle database is the obvious choice for

It's been estimated that 90% of the electronic data in the world is stored in unstructured form, such as word processing documents, spreadsheets, and electronic mail.

storing large amounts of data, due to its scalability, security, and reliability, but traditionally the database has been used to store structured data rather than unstructured content in a variety of formats.

The Content Management SDK allows you to store all of your data, from graphics and word processing documents, to audio and video clips, to presentations and email, all in a single repository. You can use the Oracle 9i database to store not only relational data, but also any kind of content. No longer will you have to guess where your information is hiding - you know that it's safe and accessible in the Oracle database. You no longer have to be concerned about the timeliness of the information you're using - everyone is sharing the same content, real time, wherever they happen to be.

ACCESS YOUR CONTENT FROM FAMILIAR PROTOCOLS

A barrier to implementation of any system is the willingness, or lack thereof, of users in your organization to embrace change. The common lament is "I don't have time to learn a new system, I need to get my job done!"

If your server is accessible via the Internet, your users will be able to log in and access their files from any Internet browser, anywhere in the world.

With no customization, the Content Management SDK is ready at installation for access through all of the common protocols in use today, including:

- Windows (SMB)
- NTFS
- NFS
- FTP
- HTTP (WebDAV)
- IMAP4
- SMTP
- AFP

No special client is needed to access files in the database. Any Windows, UNIX, or Apple compatible computer can access content in the database immediately after it's set up and connected to a network. Users can immediately begin taking advantage of advanced file and print services, through familiar clients without learning a new interface or changing the way they work. For example, word processing tools can be used to directly open, edit, and save content in the database as if it were on a shared file server.

CUSTOMIZE YOUR CONTENT MANAGEMENT INTERFACE

As professionals in the age of the Internet, most of us have become resigned to the style of work that file systems have imposed on us. We adjust our work habits, forms and processes to fit with the idiosyncrasies of twenty-five year old

The Content Management SDK is
100% pure Java. You implement
your custom interface using
industry-standard technologies:
Java, JSPs, Servlets, HTML and
XML.

technologies. Computers were meant to be our servants, but at times it feels as if they tell *us* how to do our jobs.

The Content Management SDK allows you unprecedented control over how data is stored, shared, and processed to meet your company's needs. Now, you can decide how the information should be organized and presented based on how you do business.

The Content Management SDK includes a Java API that allows you to automate business processes and build custom web user interfaces. You can design an input strategy that matches your data in its raw form, accepting legacy data from existing systems, mimicking paper forms, or creating a new interface for capturing knowledge from your users. Your content can be displayed and edited in its original form, or rendered in a completely different format (for example, information might be entered as raw text, displayed online as HTML and rendered for data transfer as a PDF file). The server can be programmed to send of electronic mail notifications when content has changed to alert personnel that review is required.

Define Custom Content Types

On install, the Content Management SDK feature of the database allows you to store files, folders, and email in the database, and stores standard metadata such as the Name, Owner, Modification Date, etc. You can program the server to automatically extract key data from files when they are stored in the database. Using easy to learn, standard technologies such as JSPs and servlets, your team can build custom web applications for managing content, such as customer relationship management applications, technical publishing applications, and self-service support systems.

One example would be to create a custom content type called *Presentation* to store PowerPoint files. The Presentation content type would include a new attribute, *Audience*, that would be used to store information about the target audience for the presentation (sales, management, etc.). The content type could also have an extended behavior, such as always putting a link to the file in a *Presentations* directory.

Automate Business Processes

Overrides and agents allow you to automate content management tasks or change the standard behavior of *9iFS*.

Overrides are modifications to the Java classes that manage the data stored in the database. You can override methods in the Content Management SDK to add your own behaviors. Overrides modify the behavior of the Content Management SDK during a transaction (synchronous modifications), making the changes as information is uploaded or downloaded from the database.

For example, you can program the server to automatically validate that data has been correctly entered when a new file, folder, or electronic mail message is created in the database. You can build agents that automatically send notifications that important events have occurred in the database, like alerting the owner of a document when it has been revised. You can dispatch requests to initiate workflows that automate complex review and approval cycles. You can integrate the server with legacy systems to import and export information, or trigger tasks based on events that occur in either system. For example, you might program the server to initiate an engineering change order when new information is input that affects the design of a product.

Parsing and Rendering

If you have legacy data in any predictable format, you can construct a Parser that will extract critical information and make it available for rapid search and retrieval.

When information is stored in the database, you can control how it is stored, preprocessing it to extract attributes or bursting it into reusable information assets. For example, a purchase order document could be parsed to extract the individual line items for product, price and quantity. The parsed information could then be entered into the extended attributes of a custom Object Type so that the information could be accessed as structured relational information.

You can also convert content stored in the database to alternative formats that are more convenient for users to view, such as HTML, PDF, and XML. You may want to tailor the layout of a document according to the job function of the person viewing it. Most of the time, files are rendered in the same format they had when inserted to 9iFS (i.e., if you upload a file as a Adobe FrameMaker document, it is rendered to disk as a FrameMaker document). However, the Content Management SDK gives you complete control over how documents are rendered. You can change the format of the document, combine many documents into a single new document, selectively omit certain portions of the content based on the user's privileges, or change the way the information is presented based on the device that is used to access 9iFS (for example, the content could be broken into small plain text strings for an application that is accessed from a cellular telephone).

Parsers and renderers can be implemented as both Agents and Overrides. You can choose the method that would provide the most efficient access to the information while maximizing server performance.

ORGANIZE YOUR CONTENT FOR QUICK RETRIEVAL

The information you store in the database is of no value until you put it to work. You can't put it to work if you can't find all of your relevant content when you need it. This is where the benefits of the Content Management SDK are most easy to see.

9iFS runs on the Oracle9i database, and is able to expose sophisticated database features to help you manage your unstructured content,

Advanced Search Capabilities

The Content Management SDK exposes the features of Oracle Text, leveraging the power of database search and retrieval technology that performs substantially faster than those of a conventional file system.

The Content Management SDK allows you to index any attribute for rapid retrieval, including the custom attributes that you create. You can use this structured information to quickly locate documents.

In addition, the Content Management SDK gives you the option of full text indexing, allowing you to perform lightning fast searches for words and phrases that appear in any of over 150 recognized file formats. No longer do you have to worry about whether the content you need was stored in a spreadsheet, word processing document, email message, or slide presentation: 9iFS allows you to search all of those files at once, and retrieves all of the relevant content in a single operation.

Categories

While it's possible to create a content type with custom attributes, projects often have content stored in a variety of formats. Categories allow you to define attributes and apply the Categories to documents in any format. You can organize different types of information according to how they are used in your business by defining categories for projects, product lines, collateral, etc. For example, you could create the Category *Project*, with the attributes *Client*, *Industry*, and *DueDate*. Using a custom interface, you can populate those values, apply them to any content stored in the database, and use them as search criteria.

Multiple Foldering

Of course, most of the time you don't need to search for your content, because you keep it in a folder set aside for that specific purpose. People have their own ideas of how information should be organized, and will choose a method that they feel works best for them. No one's organization is better than anyone else's. This becomes a problem when several users are trying to collaborate on the same files, often for different reasons.

The Content Management SDK allows you to create an interface that allows your users to store the same file in multiple directories. For example, if three people were collaborating on the *XYZ Sales Proposal*, one might store the file in a folder named *XYZ*, the second in a folder named *Sales*, and the third in a folder named *Proposals*. All of the users would be looking at the same file, and any edits would be available to all other users immediately after the changes were saved. But all of them would be able to organize the file in a way that makes it easy for them to find it when they need it.

Hierarchical Storage Management

On the other hand, there is a large body of content that needs to be stored for archival purposes, but isn't used on a regular basis. In this release, the Content Management SDK allows you to control where your content is stored. Seldom accessed files can be moved offline to a cheaper storage medium, while the metadata (categories, attributes, etc.) are maintained in the database. The administrator can easily retrieve the files should they be needed at a future date.

MANAGING THE LIFE CYCLE OF YOUR CONTENT

Before the turn of the century, file servers were geared toward meeting the needs of groups of users at a departmental level. Early on, when workgroups were small, and data was insular and lower in volume, it made some sense. People enjoyed the control they could exercise over their own or their group's dataset.

In the Internet age, the concept of "workgroup" has been turned on its ear. Corporations now must share an enormous amount of information across departmental and organizational lines. Virtual teams can be comprised of workers in locations scattered around the globe. Project plans are written in Bangalore, then reviewed and approved in Buenos Aires while the authors sleep. The goal, then, is to provide reliable access to collaborative materials that can be accessed by any authorized user, anywhere in the world, at any time, without sacrificing the need of individual groups to control who can access which content.

The safety, security, and reliability of Unbreakable Oracle 9i ensure that your content will always be available when it's needed.

Security

You can define levels of security to be applied to files and folders in the database. For each security level, you can have granular control over which users and groups can locate the information, read its content, modify it, and delete it. The security levels can be defined and maintained by administrators centrally, applying security changes across large numbers of documents and folders at once. Users can also control who can access their personal information, keeping it private, sharing it with co-authors, or publishing it for the rest of the company to use.

Versioning

Collaboration is wonderful, except when it results in one person's work colliding with or overwriting another person's contributions. The Content Management SDK allows you to solve this problem by versioning documents.

You can create an interface whereby users can check out documents for editing. While the document is being edited, other users can view the previously saved version, but can't change it. When the edits are complete, the user checks the file back in, and everyone is able to see the changes.

Each time the document is checked in, you can create a new version of the document. This way, users can refer to an earlier iteration of the document or

retrieve information lost in an edit. You can control the number of versions that are retained for each document, according to your business needs.

Workflow

Another way to enhance collaboration is to set up a formal process for reviewing and approving documents uploaded to the server. Using Advanced Queues technology, the Content Management SDK now integrates with Oracle Workflow. The Content Management SDK allows you to trigger a workflow event based on a change in the database, such as a new document upload. Oracle Workflow can also trigger events in database, and wait for a response before the Workflow process continues.

CONCLUSION

To stay competitive in the information age, you need to have safe, reliable, rapid access to your company's valuable content resources. The Content Management SDK provides you with the tools you need to store, organize, retrieve and share all of your content using familiar protocols and clients, and with custom applications tailored to your specific business needs.

You have the information. The Content Management SDK lets you put it to work.



White Paper Title

June, 2002

Author: Dennis Dawson

Oracle Corporation

World Headquarters

500 Oracle Parkway

Redwood Shores, CA 94065

U.S.A.

Worldwide Inquiries:

Phone: +1.650.506.7000

Fax: +1.650.506.7200

www.oracle.com

Oracle is a registered trademark of Oracle Corporation. Various product and service names referenced herein may be trademarks of Oracle Corporation. All other product and service names mentioned may be trademarks of their respective owners.

Copyright © 2001 Oracle Corporation

All rights reserved.