



## IDCMS Blue IBM Component Content Management System for DITA



*Tonya Holt, Mike Iantosca, Ellen Patterson, & Sophie McMonagle, IBM Corporation*

*IDCMS Blue is a new Darwin Information Typing Architecture (DITA)-optimized component content management system. The system described in this article is developed and used internally by IBM Information Development teams worldwide to manage the development of client-facing technical documentation and integrated user assistance. In this article, project team lead Mike Iantosca, Information Architect, Sophie McMonagle, and Infrastructure Lead and IDCMS Blue advocate, Ellen Patterson, describe the benefits of the new system and the experiences of two teams that are now using it to produce their product documentation.*

Imagine if you were responsible for more than a million technical topics. Imagine if you were responsible for publishing more than a billion words annually and translated those words into dozens of national languages. How would you manage the information? Would you insist that your team adopt a uniform standard such as DITA for structured content? Would you be able to deliver a content management system that is intuitive and easy to use? And how would you customize the content management system to meet corporate and brand-specific processes, ensure conformance to standards, and enable solutions-oriented information?

The Corporate Information Development (ID) Tools product development team at IBM is developing an end-to-end lifecycle management system that IBM ID teams are using to author, manage, reuse, and deliver product documentation for the company's vast product line. This strategic content management system is known internally as Information Development Content Management System (IDCMS) Blue. This system, the result of an ongoing multi-year development initiative, serves as the core of an ID authoring tools ecosystem to enable the reuse and sharing of technical documents between technical writers across the enterprise; it also automates many ID processes and integrates with development systems.

Now in full production, IDCMS Blue will eventually encompass the total information lifecycle that extends from planning and design through authoring, quality management, translation, and distribution. IDCMS Blue is based on IBM FileNet P8. IDCMS Blue has been customized to meet the needs of the IBM ID authoring community and is intended to be a repository for documentation projects. ID teams can store and

manage DITA-based HTML/XHTML, and SGML documents. These document files include not only the marked-up text files, but also artwork and other collateral required to generate the final deliverable.

### DITA-AWARE COMPONENT CONTENT MANAGEMENT

Not all systems are optimized to support DITA documents. Unlike source control management systems that ignore the contents of DITA topics, IDCMS Blue includes a DITA document classifier that is content-aware. The document classifier organizes DITA documents by topic type and represents relationships between DITA maps and DITA topics, including relationships between topics such as DITA content references. As a result, a user can quickly determine in which DITA maps a topic is referenced and reused.

The document classifier is fully extensible and can be tailored to support any DITA specialization. On document add or check-in, the classifier optionally extracts metadata from the mark-up within DITA topics and makes that metadata available for many uses, including intelligent search and reuse.

IBM initially considered bursting DITA content into objects smaller than a topic but quickly discovered that doing so wasn't practical for end users in a large scale production environment. "At first glance, being able to drill down to view and manage individual DITA elements in the CMS seemed attractive" says Mike Iantosca, project team lead, "then we discovered that DITA itself provides all the capabilities needed to manage and reuse content at the sub-topic level, leading to greater ease of use, better document portability, and far higher system performance."

"One of the best aspects to this approach in managing DITA this way is that a content manager doesn't co-opt control of DITA relationships and reuse—our content doesn't become dependent on a proprietary system; DITA remains the trusted source—always.

### MULTIPLE INTERFACES

IDCMS Blue has several interfaces that enable authors and applications to access various functions in the solution:

- ◆ **A web-based client** provides a full-functioned interface provided with drag-and-drop ease of use. For DITA, a

hierarchical view of an entire DITA collection (DITA Maps, their child DITA topics and child DITA Maps) is highly desirable.

- ◆ Support for multiple DITA **content editors** from various providers include connectors and interfaces that seamlessly interact with documents in FileNet P8 storage. The editor integration supports key functions for check-out, check-in, cancel check-out, download, and checking status. Multiple editor support is important to serve various authoring skill levels.
- ◆ A **command-line interface** enables teams to create scripts for batch operations to download, check out, and check in large numbers of files.
- ◆ A **rich connectivity Java API** developed by the IDCMS Blue team enables a wide variety of applications and systems to seamlessly integrate with the repository.

### A WORLD-CLASS INSTALLATION

A single, central server complex supports all of IBM's DITA content worldwide, making the entire collection of DITA topics searchable and available for reuse to every author across the organization's global enterprise. The ID server infrastructure team built a world-class complex with impressive performance. The production complex can locally ingest the company's entire inventory of DITA topics, and then some, in less than 20 minutes and provides excellent response time across all geographies worldwide. In addition to meeting IBM's stringent Information Technology certification requirements, the company required the system to provide autonomic and failover features to ensure high availability and reliability.

### IDCMS BLUE IN ACTION!

Many Information Development teams have already moved their DITA content to IDCMS Blue and are currently using the content repository to manage their documentation. Experiences from two early adopter teams are provided.

#### CICS Transaction Server

The CICS® Transaction Server team has moved all of its information from the internal Configuration Management Version Control (CMVC) source control repository to IDCMS Blue. The project was sizable, comprised of nearly 38,000 topics in over 50 Eclipse information center plug-ins. Information developers access these topics on a daily basis, with an automated build process that transforms and composes documents for client delivery.

The migration process was relatively quick, even with the team fixing errors that were discovered in original source. In terms of actual resources, this work required approximately two-person weeks of effort for the migration team.

With IDCMS Blue, you can see relationships between files; for example, a DITA map lists all of the files it has as

children, with content, cross-references, or graphics also shown as relationships. IDCMS Blue understands what sort of files these are, and if, for example, you have a cross-reference that IDCMS Blue cannot find, the system creates a stub document as a placeholder to tell you that it is expecting content that is missing, which helps identify broken links and orphaned documents.

Sophie McMonagle, CICS Information Architect, on moving to IDCMS Blue: "The experience was swift and smooth and good planning helped make sure that we could control the transition. The team quickly picked up the new way of working, and we like the current search facility in the web-based interface. We are looking forward to future enhancements that will allow the team to search on tag types, metadata, and any content within a topic."

#### Information Management

The Information Management team also experienced success with IDCMS Blue. The project was small, with DITA topics, DITAMAPS, and graphics. The topics were created as DITA types, and the information was componentized, with separate plug-ins for key components such as installing and administering.

The experience on the production-level server has been successful. Among the highlights:

- ◆ A Lightweight Directory Access Protocol (LDAP) security allows easy control access to plug-ins and files for various purposes, such as modifying or viewing.
- ◆ The performance of the production server is extremely fast. The team members are located across the country, and no one experiences performance problems.
- ◆ A favorites container provides a useful way to access regularly updated documents quickly. IDCMS Blue uses multiple directories, so having some documents as favorites saves time.
- ◆ The team especially likes the close association with the ID authoring tools suite known internally as the Information Development Workbench (IDWB). Writers can check out and check in one file or all files in a DITA map using integrated editor menu options.
- ◆ The system makes it easy to view and update metadata by topic. Within the web client, there are excellent visual cues as to which topics contain other topics (compound documents). The editor connector also provides useful visual cues to remind the writer which document is currently checked out.
- ◆ Users can search across the object store to locate documents and folders based on the values of their properties.
- ◆ There are various command-line functions developed on top of the system that make the use of automation in extracting, building, and transforming a painless process.
- ◆ There is nothing to install by the end user. IDCMS is a web-based content management system, which is one of the team's favorite benefits.

According to IDCMS Blue advocate, Ellen Patterson, IDCMS is the right content management system for the type of documentation that the IBM Information Management team is responsible for delivering. Its ease of use with the current authoring tools enables a writer to be immediately productive.

### A PATH TO VALUE

IDCMS Blue is an integrated and robust content management and workflow system. The system exemplifies what can be accomplished with a robust, DITA-aware component content management system. IBM plans to automate many processes using automated workflow and the best-of-breed tools for each task. An enterprise information planning system built on top of IDCMS

Blue will drive much of that process automation such as single-click collaborative team review, information quality analysis, and reporting—complete with visual quality management dashboards and continuous national language pre-flight and translation. “The list of what can be automated is extensive and the business value enormous” says Mike Iantosca, “in the first few years of operation we are committing business benefits in the tens of millions of dollars.”

The Information Development organization plans to use IDCMS Blue to take advantage of document intelligence, address pain points both new and old, and help teams produce more consumable products with faster time-to-value for IBM clients. 