

# Bridging the Real Silos in Content Management with Flexible Content Objects

*Leveraging the Right Content Management Technologies to  
Bring Content and Process Together*

*Bill Trippe  
March 30, 2003*



# Executive Summary

As departmental and point content management solutions give way to an emphasis on solutions for the larger enterprise, CIOs and other technology managers need to focus on the core value proposition and potential ROI from content management technology. The thesis of this paper is that the real value in content management technology will be realized in its use in optimizing business processes and applications. This is especially true as organizations extend more applications to the Web and to other devices. As applications extend to an increasingly dispersed enterprise, content must be flexible enough to go with it.

To be effective, content management technology needs to work across the entire lifecycle of content, which we see as the stages of creation, management, integration, transaction, and distribution. As a result, content must move freely through these stages and be associated with applications and logic at each stage of the lifecycle. This requirement suggests a need for systems to manage content “objects” that have a rich expression of both content and metadata. Moreover, it suggests that content “objects” themselves may be single, physical objects or more abstract objects—compound documents and aggregations come to mind. In both cases, content management technology must be flexible in its ability to manage both objects and the associated metadata.

Organizations that adopt an object approach to their information architecture will be well positioned to solve the myriad applications and business processes that require tightly integrated content. They will be able to act quickly on simpler integration problems and will have the underlying infrastructure to begin tackling the most complex integration issues.

Content management technology must evolve to support such an approach. Indeed, the long-term approach we see is one of Enterprise Information Integration (EII), where both content and processes are supported by an infrastructure that manages both complex content objects and complex code.

We conclude by looking at Vignette’s latest content management offering, Vignette V7. Vignette V7 has an impressive set of new features and architectural changes from previous Vignette solutions that position it well to create the kinds of integrated content and process that we think will provide the most ROI for organizations investing in content management technology.

# From Creator to Consumer: The Content Lifecycle

Behind all the details of content management technology looms a single, overriding requirement—delivering the right information to a user in the context of an action. This could be catalog content to someone trying to make an informed buying decision, engineering information to a customer support troubleshooter, or clinical information to a doctor writing a new prescription. In all these cases and in thousands more, content management exists to optimize *communication* between the person creating the information and the person consuming it. More specifically, content management technology works best when it optimizes communication in the context of some action or process. In other words, communication is not an end in and of itself. Instead, communication drives process, and process, in turn, drives business.

The end result of improved communication is a business decision. When the consumer has been provided with sufficient content, he purchases the product (or not). When the customer support troubleshooter has reviewed the engineering documentation, she can solve the customer's problem and close the trouble ticket. And when the doctor is satisfied he has enough information to prescribe the right medicine, the prescription is written and the patient's illness comes under management. With more optimized communication, businesses get more done—more decisions are made more efficiently, leading to a more flexible, productive organization.

Viewed this way, content management technology must (1) support collaboration and (2) integrate with other key applications and business processes.

- *Collaboration* means the content is available in the right form for steps such as creation, review, and consumption; it also means that access control is flexible enough to support an extended group of users who each may have multiple roles with content under management.
- *Integration* means much the same thing, except in machine form. That is, the format for collaboration may or may not be the right format for machine consumption. Similarly, the levels of access and control need to be tuned for various applications.

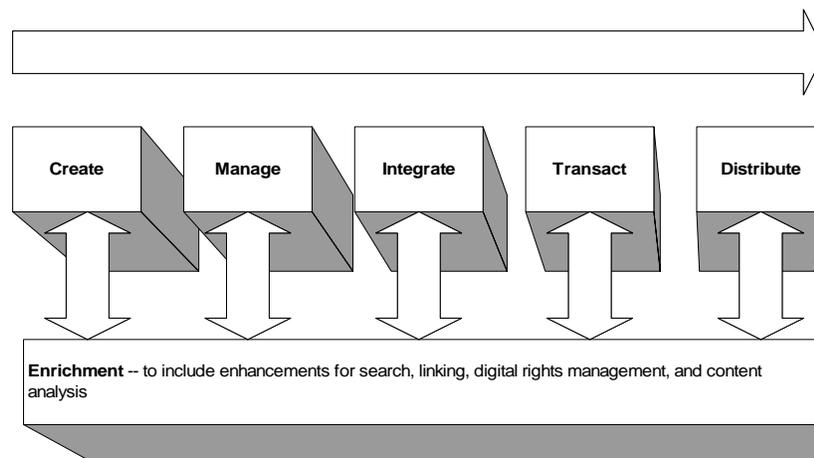
In other words, when you harmonize collaboration with integration, the end result is that business users can manage content where it lives, allowing them to bring more information to bear more quickly and efficiently, and thus to move the organization forward toward its business objectives. In this way, content management technology serves the core needs of the business.

Our recent research report on the content management marketplace described five stages in content management. Individual technologies typically support more than one stage, and many support several. It's worth looking at the stages or functions of the content management lifecycle, and to highlight how each function needs to support collaboration and integration.

## The Five Stages of Content Management

Content management is a broad problem, with many vendors offering solutions ranging from enterprise class to departmental to single-purpose to tools for personal productivity. In 2002, *The Gilbane Report* and colleagues from MIT's Sloan School of Management researched over 400 vendors and 135 users of content management technology in an attempt to develop a framework for understanding the marketplace. The research clearly identified five stages of content management—from creation to distribution. These five stages are:

- *Create*. The creation stage allows users to create original content, to transform content into other formats, and to create supporting data, often referred to as metadata.
- *Manage*. The management stage enables users to manage and control content in a repository, to provide library services such as check-in/check-out, and versioning, and to control read and write access to content.
- *Integrate*. The integration stage allows users to access content in more than one source format or repository; it also allows organizations to tie content to applications and business processes more easily. Organizations may have documents under the control of one system and multimedia elements under the control of another. Content management integration enables users to have a common view of multiple repositories; it also allows other applications a common view or interface with multiple repositories. More significantly, content management integration can help bring together not just the content itself but the business process and applications that support this content.
- *Transact*. The transaction stage enables organizations to monetize or otherwise realize value from content. For a commercial publisher, this could be a simple example such as the component technology that allows them to sell their content over the Web. For other kinds of businesses, this includes technologies such as integrated email campaigns that lead to specific interactions. Examples of this include the customer support portal that emails a group of customers, directing them to a download with new information about a product in the field.
- *Distribute*. The distribution stage enables users to distribute content in many forms, and to many internal and external touchpoints.



*Content enrichment* is itself an umbrella category, spanning technologies such as linguistic tools for enhancing search and retrieval and tools for translation and localization. It spans the five stages because content enrichment happens at many points in the content lifecycle. Indeed, enrichment includes activities that may be appropriate at virtually any stage of the lifecycle:

- Association-type activities are useful at many points in the lifecycle; for instance, cross-referencing and sequencing are value-add activities that can happen at creation, after content has been ingested for management, at a point of integration, or at the point of distribution.
- Digital rights management (similar to association) is a set of value-added activities that spans the entire lifecycle. Some DRM activities can occur at creation. Others might occur upon ingestion and still others might be dynamic, allowing management of rights "on the fly" *after* distribution.
- Content analysis and tracking tools provide content owners with ongoing feedback of content after it has been distributed. Analysis of usage patterns can lead to content being updated, enhanced, replaced, or purged. Used correctly, such analytical tools can give organizations excellent visibility into the value of the content—who is using it, how it is being used, and so on—leading to improving practices and opportunities for continuous improvement.

As content management technology has matured, so too has the notion of the "content object." The object under management is both the content object itself and its associated metadata. The change, over time, has been in how complex, granular, and multivariate the content object has become. In a traditional document management system, the document itself was usually the object, and it was managed along with its associated metadata. With the advent of Web content management systems, the object was sometimes document length (*e.g.*, an article for a news Web site), but just as often, these systems also managed other, finer-grain objects—such as the title for the article, byline, associated artwork, and so forth.

The real requirement is even more complex than either the document or Web content management approaches suggest. Content objects can be compound objects in and of

themselves and can also be aggregations or views of a set of other content objects. In today's organizations, compound documents can source objects from multiple locations (*e.g.*, the parts catalog with three-dimensional drawings that are stored in a digital asset library). Content management technology should account for and manage such objects.

The reality, of course, is that organizations have much content, but have less than 100% of this content under management, and have much less than 100% of their managed content as flexible, composite objects. The average organization faces a complex, heterogeneous set of content types, managed under different systems, with varying levels and types of metadata to help describe and control the content. Yet, at the same time, these same organizations face a need to bring content into more and more business processes. As organizations continue to move applications and business processes to the Web, they need to tie well-managed, easily distributed content to these applications.

Put another way, organizations face a need to manage these content objects in ways that tightly integrate them with real-time business processes. So we like the term “real-world business object” to describe:

- the requirement for content to work closely with business process
- and the recognition of the messy reality of how difficult such integration has proven to be.

Such real-world business objects must first be successfully managed through the five stages of content management we have described—create, manage, integrate, transact, and distribute. They must also be successfully enriched—indexed for search and retrieval, translated for internationalization, links added, and such. One measure of the *comprehensiveness* of a content management solution is how successfully and deeply various content management objects can be managed across all the stages of content management. Yet successful management of the content itself is a necessary but not sufficient step toward the kind of *information management* we have in mind. The real goal is to manage these real-world business objects in such a way that they can be readily integrated with other applications and processes.

Why is such integration important? Organizations face an imperative to extend applications and content to the Internet and beyond. Gartner calls it the real-time enterprise; others talk about communicating and interacting through multiple touchpoints. Regardless of the terms, the problem is a practical one. If organizations are going to extend applications and process to the Web and other devices, they need to bring content along to support the applications. Their content management technology must support such a widely distributed infrastructure.

# Creating a Platform for Enterprise Information Integration

What will better, more integrated content management and process provide? It will provide improved, more efficient, and more transparent information technology that allows organizations to make more business processes automatic and more users self-servicing. These organizations can range from the very large organization processing thousands of maintenance procedures per day, to the small membership organization that wants their members and customers to find, order, and download information on their own.

So if content management technology works best when it supports communication in the context of some action or process, what must content management do to support such communication? At a high level, content management must synchronize content with all required applications. At a more tactical level, content management systems must provide content objects in forms suitable for sharing with a variety of applications. One system may require a structured format, the next a less structured one, and so forth.

Most organizations have heterogeneous platforms of operating systems, repositories, applications, and content types. In such an environment, the task of synchronizing content—and delivering it in all its required forms—requires a flexible creation, management, and delivery system. Such a system must be able to manage content at an atomic level and at various higher levels of abstraction, including complex content objects and aggregations of content. In order to facilitate delivery, the system must also have an efficient and easy to use process for assembling and deploying content and metadata. Significantly, it also must have open, services-based programming interfaces that make it easy for other applications to request content for consumption and to return content for the management system itself to consume.

---

## Integration in Practice

Consider, for example, the manufacturer of electronic supplies who wants to allow their customers to look up products, order them online, and check on the status of shipments. This company has grown through both growth and acquisition, so its internal systems have strained to keep pace with the growth. A legacy parts database still controls technical specifications and pricing on core products, and secondary systems control other parts, inventory, and shipping information. The primary goal, as the manufacturer sees it, is to use a content management system to create a customer-facing Web site that presents all available products, with content drawn from the core parts database and a new, but growing repository of catalog descriptions and photos. A secondary goal would be to use this same content to begin producing print catalogs.

In this example, the content management system must synchronize structured, semi-structured, and unstructured content into a single unified presentation on the Web. This doesn't *necessarily* mean that the system must control all sources of content, or require that all these data sources be converted to a single source. Indeed, the legacy systems work fine, from the manufacturer's standpoint, and they are in no hurry to change them. Instead, the content management technology in this case must have the capability to integrate these

various data sources, assemble the content into a coherent presentation, and distribute it to the Web and elsewhere.

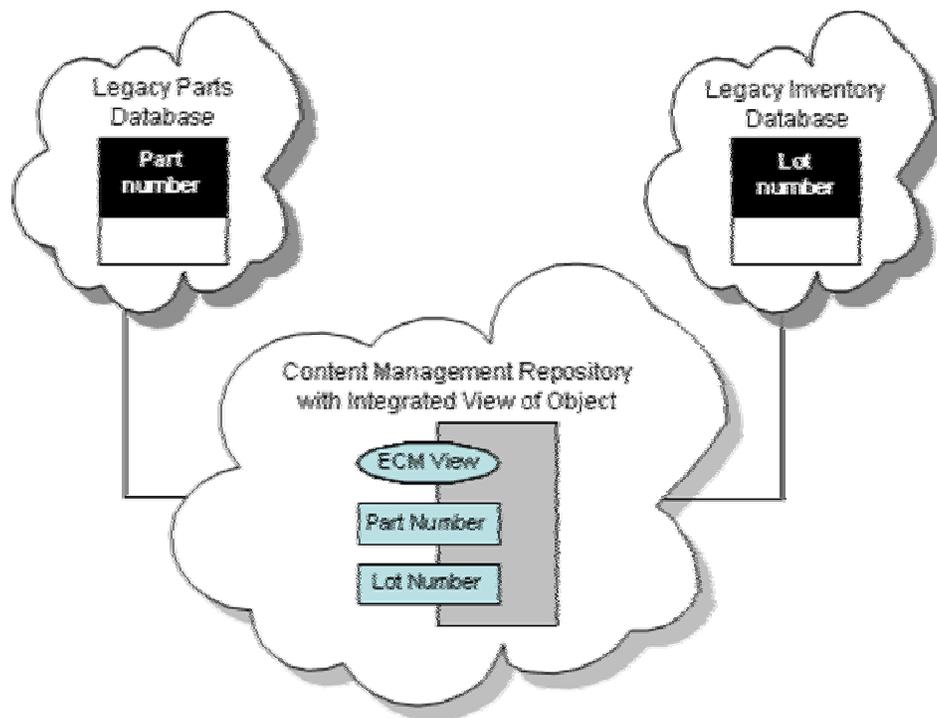
The potential benefit in this example is that the organization is able to extend the life and value of their technology investments. The right content management technology will save this organization the significant cost and time associated with migrating all the various data sources to a completely new system.

The above is a relatively complex example, and one not faced by every organization. But every organization faces a laundry list of small- and medium-scale content integration problems that could be solved with a more flexible and complete content management system. Consider the examples from the beginning of this paper: the needs for catalog content, engineering documentation, and medical research and records. Organizations are awash in content and in potential applications where the content can be utilized.

Given the number of systems involved, the granularity of the data, and need for the real-time synchronization, we see a growing movement toward *Enterprise Information Integration (EII)*. EII is best understood against a backdrop for the content dimension of Enterprise Content Management (ECM) and Enterprise Application Integration (EAI). While ECM (commonly seen as the combination of document management, Web content management, and digital asset management) has been addressing the broad problem of content management in and of itself, it has not directly addressed the problem of integrating content with applications or separated repositories. Similarly, EAI has focused on the problem set programs face in tying together disparate applications through infrastructure buses, but EAI has not focused intensively on the problems developers encounter in more tightly binding content (with metadata) with process.

Enter the real-world business object, with its mix of metadata and content. From a content management perspective, the real-world business object is easily managed. If it is based on open standards such as XML, the object and associated metadata can be easily produced and consumed by most content management technologies. (More systems are storing XML natively, and even those that don't store it natively have transformation tools for going from their internal formats to XML and back again.) EAI tools also easily manage the real-world business object. XML is already predominant in enabling process-to-process communication in EAI. If the content management system can produce an XML object, shouldn't the EAI tools easily handle it?

Consider our catalog example again. Couldn't EAI-style processing integrate the content and data in the various legacy systems with the content as managed in the Web content management system?



Is EII the confluence of ECM and EAI? An extension to both of them? The next big thing? We first see incremental change in this area, and not wholesale development and adoption of whole new platforms. This is true for several reasons:

- The requirement is pressing, and IT organizations are not going to wait for the next big thing.
- The most immediate problem is to get more content under management, and in a way that makes the content available to both human and computer interfaces.
- As more content comes under management, organizations are deploying it in business processes. Even smaller organizations are investing in self-service applications that tie content and applications together.
- Organizations will continue to build on these successes, adding new applications and extending existing ones.

A solution requires modular, interchangeable business objects that can move easily among multiple applications, multiple systems, and multiple organizations. Once these content objects are under management, the right approach to application integration will lead to measurable success.

---

## The Road to There

There is much upside to the efficient management of real-world business objects, and organizations should adopt the approach as core to their information architecture. Our take is

that most organizations are evolving in this manner and have been spending significant time and effort to put some of the key building blocks in place. More content is digitized to begin with, and organizations are doing an increasingly better job of inventorying their content, enhancing it where appropriate, and getting it under management. A reasonable starting point for any organization is a critical mass of content and metadata, and more organizations are arriving at such critical mass.

What is still missing for most organizations, is the infrastructure to support these objects at both the content level and the metadata level. This too, though, is changing. Organizations have invested significantly in content management technology, and continue to do so. Database platforms have added significant features to better manage XML and other content. And specialized repositories for XML continue to show surprising strength and growth. As more of these systems are deployed, more content will come under management and enjoy the benefits of core repository services.

These positive changes will fuel the integration of content and applications, which up until now has evolved through three stages of maturity. We see infrastructure support for EII as the fourth.

1. No integration: Stovepipe applications
2. Wishful integration: Expensive custom development
3. Partial integration: (EAI, ECM, point-to-point)
4. Infrastructure support for Enterprise Information Integration (EII), which we see as a deeper level of integration.

Given the recent focus on *service-based architectures* and Web Services *per se*, it is important to understand the relationship between Web Services and EII. We see Web Services as part of the EII whole. Web Services are the open, self-documenting, discoverable pieces of functionality that can be accessed through simple, high-level programming interfaces. EII is supported by the larger infrastructure that creates and manages both the content objects and the interaction between the content and the applications. Web Services is one way to loosely couple processes and content as part of EII infrastructures.

---

### Conclusion: How Can Systems Best Support EII?

Content management systems can move toward an EII infrastructure by embracing this notion of the real-world business object and supporting such objects at each stage in the content management lifecycle. The best systems will be ones that create, manage, integrate, transact and distribute composite content objects that have a rich expression of content and metadata.

These systems must manage these content objects in ways that expose both the content itself and the metadata to integration with other systems and applications. It should be easy for other applications to produce content that the content management system can consume. And the content management system should be able to produce content and metadata in as many forms as associated systems require.

Along with increased support for content objects, content management systems should also continue to add more features and support for application integration. Content management does not happen in a vacuum. Systems need to have open interfaces so that other applications can quickly integrate. To this end, we see more development of Web Services-style interfaces, and eventually, a deeper level of support for EII.

## Vignette V7

Vignette V7 is an intriguing new version and significant revamp of Vignette's content management platform. Note that Vignette V7 is a broad offering that provides much more than content management; offering six sets of core services for content management, portals, integration, business process, collaboration and analysis. Continuing down a path towards J2EE-centered applications, Vignette V7 hands over the invocation and management of the services components *in total* to J2EE application servers. Such an approach is consistent with IT organizations' direction to standards based computing and shows an evolution from previous versions of Vignette's content management technology which have been at an arm's length from such an approach.

Vignette V7 has moved Vignette forward in an area that is key to customer success and also central to the theme of this paper. That is by providing the ability to model real business content objects up front and quickly, and to enable organizations to bring more content under management much faster than organizations could in the past. Implementers can model content objects in various forms, in various locations and repositories and at varying levels of abstraction. Vignette calls this "managing content where it lives," but it is also "managing content for what it is." The content modeling tools for Vignette V7 allow a content designer or application implementer to associate content objects with any information system, relational database system, or Web Service and to manage discrete content objects such as compound documents and digital assets. Significantly, the content modeling tools also allow the content to be managed as XML described documents. A content type can be defined and associated with an XML schema for validation, but objects can also be defined by their native format and managed "in place" if desired.

Managing content "in place" and in its native form has enormous practical benefits for organizations with distributed and diverse content sources. It saves organizations the time and effort associated with having to transform and re-integrate all of the content sources. (Jupiter, among other research groups, has identified the enormous costs organizations face in content conversion.) Vignette's content integration tools allow organizations to aggregate, integrate, and manage more content sources quickly, accelerating system implementation and user productivity.

Once under management, Vignette V7 application services can manage these business content objects through the entire lifecycle of the content. These increased capabilities in content object modeling will have a beneficial effect on organizations as they bring more content under management more quickly and also manage the content as more flexible and composite objects. This provides an opportunity to manage content object across transactional, structured and un-structured sources.

The increased XML capabilities embedded in Vignette V7 entitles Vignette to their claim of managing real-world business objects. V7 delivers functionality for creating, managing, integrating, and distributing XML and Web Services-based content and providing it to content-centric applications.

Also of interest is Vignette's development in business process management and workflow. Vignette V7 introduces a process workflow engine and an associated Microsoft Visio-based tool for modeling and implementing process workflow. A set of built-in graphical primitives

allows developers and business users to quickly design workflows and associate content and processes with tasks, stages, and business users in various roles in the workflow. The underlying process engine then supports and enforces the desired workflow, capturing operational metrics along the way. Like other workflow engines, the process modeling capabilities of Vignette V7 give organizations a ready means to quickly define, deploy, and standardize business processes. Vignette V7 extends workflow to all kinds of content types and to steps in the application that go beyond content processing itself— basically from the very creation in the original applications (from Word to SAP) all the way to the managed deployment in a .net, J2EE, database or file system, or any other deployment model.

In total, these new features and architectural changes position Vignette V7 well for solving the emerging problem set of managing content objects in a way that better facilitates business process integration. Vignette's explicit goal was to advance Vignette V7 to a point where it not only successfully managed content and associated processes, but also managed them through the content lifecycle. With such object-level management, the task of building composite applications can be done quickly without writing and maintaining large quantities of specialized code for discrete stages in the content lifecycle. For organizations with requirements for rapidly assembling composite applications, Vignette offers the Vignette Application Portal and Vignette Application Builder solutions, which supports the object-level management of content for developers with more complex applications to build.

Given the increasingly open architecture, improved XML handling, content integration capabilities and investment in business process modeling, we see Vignette V7 as an important new platform for the content management market. Vignette V7's focus also underscores our point that ECM is a transitory stage that may well give way to EII infrastructures in the near future.

# Sponsoring Company Information

For more information, please contact:

**Vignette Corporate Headquarters**

1601 South MoPac Expressway  
Austin, TX 78746-5776

p: 888.608.9900 toll free

p: 512.741.4300

f: 512.741.4500

[info@vignette.com](mailto:info@vignette.com)



**Vignette Latin America**

p: +305.789.6603

p: +305.789.6612

[lainfo@vignette.com](mailto:lainfo@vignette.com)

**Vignette Europe/Middle-East/Africa**

p: +44.1628.77.2000

f: +44.1628.77.2266

[euroinfo@vignette.com](mailto:euroinfo@vignette.com)

**Vignette Asia-Pacific**

p: +61.2.9455.5000

f: +61.2.9455.5200

[asiapacinfo@vignette.com](mailto:asiapacinfo@vignette.com)