One of the challenges facing anyone considering a content strategy [1], whether on the scale of a single web offering or a global enterprise, is sustainability. It is only with intelligent content [2] that it becomes possible to talk about a sustainable enterprise content strategy. Automation can be used to minimize the time, effort and money needed to apply a good content strategy. However, automation doesn’t just happen. Content must be consciously designed to support it. An intelligent content strategy establishes a coherent plan under which content will be designed, developed and deployed so as to achieve maximum benefit to the customer and the organization while minimizing the cost to the organization.

What Is Intelligent Content?

Historically, content has been managed as documents. Metadata is applied to documents to facilitate document search and retrieval for both users and for the content creators. Unfortunately, applying metadata to a completed document means that it can only adequately describe the content at a very superficial level; it cannot identify the many types of content within the document. The searcher must still examine the complete document and extract the information they were looking for.

This limitation explains why there has been a steady increase in interest being directed toward open content standards [3] and specifically the Extensible Markup Language (XML) [4]. If we design and prepare content in a way that is completely portable and open, then a wide range of applications can be used to automate common content tasks such as formatting. If we make the content intelligent by tagging and structuring it, designing and preparing it for discovery and reuse, we can be freed from managing it within the “black boxes” of completed documents.

We can move forward to actually managing the content itself once we take the step of making it intelligent. Intelligent content is content that is structurally rich and semantically categorized, and is therefore automatically discoverable, reusable, reconfigurable and adaptable.

Let’s look at this definition of intelligent content in a little more detail.

Structurally Rich

The structure of a marketing brochure might contain a positioning statement, value proposition, features and benefits. Structure makes it possible to manipulate it. For example, we can automatically determine how to publish it to multiple channels (print, web, mobile) or we can filter out some content (for example, tables may not work as well in the mobile environment). We can perform searches or narrow our search to the particular type of information we are interested in (for example, we can look for all occurrences of a word in the context of a specific element such as positioning statement).

Semantically Categorized

The word semantic means “meaning.” Semantically categorized content [5] is content that has been tagged with metadata to identify the kind of content within it. For example, you might tag your content with industry,
role or audience, and product, allowing you to automatically build customized information sets based on audience or industry. As content is pushed to wikis, integrated through mashups [6] or pipes [7], it becomes even more important to ensure that our content is semantically tagged. Without semantic metadata it is very difficult to automatically, let alone manually, find the content we need.

**Easily Discoverable**

If the content has semantic tags and is structurally rich, it is a whole lot easier to find exactly what we are looking for. And when it is structurally rich, and assuming our content is in XML, we can use XQuery [8], a standard that supports queries of XML data—not just XML files, but anything that can appear as XML, including databases. We can use XQuery to query the structure of the content to find specific information. Then when we add semantic tagging to the content, we have a great deal of information that will allow us to zero in on exactly the content we are looking for (that is, content mining).

**Efficiently Reusable**

Reusable content [9], content that is created once and used many times, reduces the time to create, manage and publish and reduces translation costs. We can create modular structured content that can either be easily retrieved for manual reuse or automatically retrieved for automated reuse.

**Dynamically Reconfigurable**

In structured content the words and the look and feel of the content are not embedded in the content. That independence makes it very powerful. Knowing the structure of the content, we can output it to multiple channels reconfiguring it to best meet the needs of the channel, or we can automatically mix and match content to provide us with the information customers need [10]. We can even transform content (reconfigure it) from one structure to another, but only if we know what the structure is in the first place.

**Completely Adaptable**

We frequently create our content for a particular need or audience, but content can be adapted (used in a different way), often without our knowledge, to meet a new need. Think of mashups: We don’t know how our content is being aggregated, but we know that it can be because we have structured and tagged it intelligently.

**Who Is Using Intelligent Content?**

A number of industries are making use of intelligent content. Companies whose product is content, such as publishing and media companies, have begun to adopt intelligent content as a methodology for moving away from their traditional print to a truly multichannel (print, web, mobile, eBook) and often personalized content offering. Companies who produce huge volumes of content, such as life sciences and financial companies, use intelligent content to optimize access and retrieval. The high technology and aerospace industries have been developing intelligent content for a number of years. Government is starting to use intelligent content to manage and deliver legislative content.

**Benefits of Intelligent Content**

There are many benefits of intelligent content. The following are among the things we can do with intelligent content:

- find it more easily
- deliver it
- customize it
- personalize it
- automatically deliver it to multiple channels
- simultaneously release content in multiple languages.

And...

- reduce costs
- speed up delivery time
- optimize resources
- do more with the same resources
- increase customer satisfaction.
Case Study: Intelligent Marketing Content

The business problem. A large global telecommunications company had over 150 products aimed at large business telecommunications infrastructure. Marketing was key, but the department had been cut to the bone with the downturn in the economy. A small department of five had to create and maintain all marketing materials. The print design and creation were contracted out to a creative agency. They were responsible for more than 25 different information products including brochures, case studies, data sheets, product overviews, product comparisons, whitepapers, tweets, posts and sales training materials.

A number of pain points existed:

- They were short staffed and unable to keep up with the workload.
- The company planned on releasing more products in a shorter period of time than ever before.
- Content was localized into nine different languages, a costly and time consuming process.
- A core document was typically created and then distributed via email to multiple recipients. Content was modified for each channel, region and audience. Changes were sent via email, but there was no guarantee that everyone who needed it got it or that the revised messaging was incorporated.
- Content was written and rewritten over and over rather than reused.
- The cost of creative services was growing exponentially, and they often had to make the decision to not produce print materials for some products because they couldn’t afford it.

Goals and objectives. The following were their goals and objectives:

- Do more with the same resources.
- Develop standard core information products so it is easy to rapidly create new content rather than redesigning each time.
- Develop a repeatable reuse strategy to reduce the workload and reduce the cost of translation.
- Make it possible to easily reskin content for multiple sites giving it a product look-and-feel while retaining common structures.
- Reduce the cost of translation.
- Reduce the cost of creative services.

The solution. We determined that the Darwin Information Typing Architecture (DITA) [11] was appropriate for the content development. With DITA – an XML-based, end-to-end architecture for authoring, producing and delivering technical information – we could create structured, modular, reusable content [12] that could be automatically adapted to each of the desired outputs (print, web, mobile). It also provided a strong support for translation. While a component content management system was desirable, it was not in the budget.

A core-messaging document was created. Each of the messages within the core document was saved as a separate component, making it possible to rapidly update a single component as necessary. Content was distributed through workflow. Every action on content was tracked (recipients, version, changes and translation). At every point in the lifecycle content was controlled.

Creative services provided traditional well-styled publishing tool templates, and an XSLT [13] (XML stylesheet) was designed to map the DITA to importable XML recognized by the publishing tools. Approved content was automatically pushed through the templates (no designer was required). Final layout tweaks were sometimes necessary, but as automation was optimized this intervention occurred less often.

Project success. The team was able to develop and publish content in a marketing campaign 25% faster than they could before. They reduced their creative costs by 60% and their translation costs by 25%.

Challenges. Marketing was adverse to structure, feeling that it limited their creativity and made all messaging bland and uniform. In addition, XML scared them. We made sure that the XML was under-the-covers by selecting a friendly authoring tool. As far as they were concerned they were working in their familiar authoring environment. However, they had to use defined styles rather than hand-formatting the content in order to publish automatically.
to multiple output types (a Word/printed file, a presentation or on the web). They quickly realized the styles didn’t reduce their creativity but rather helped them save time and minimize mistakes. We also showed them how to create variants on the standard message for specific customer positioning while using the core messaging as a source.

Case Study: Design of a New Aircraft

The business problem. In setting out to design a completely new aircraft, an airplane manufacturer realized that they were faced with both an opportunity and a challenge. The global marketplace for aircraft was changing rapidly, and radically new design concepts were required. This business environment meant that the very latest in design technologies and manufacturing techniques would be needed. The content sources existed in a number of different formats, ranging from proprietary databases, arcane desktop publishing files and even custom data structures with their own unique, dedicated compilers. The sources were shared across many aircraft fleets, encompassing both military and civilian variants. Some were even shared with competitors. They would need to dramatically increase the level of intelligence exhibited by a bewildering volume of content sources in order to succeed.

Goals and objectives. What was needed was an intelligent content strategy that would establish the authoritative source for all content assets and that would set out a sustainable approach to managing these sources so that they could be used by a massive array of consuming applications.

The solution. The intelligent content strategy needed to accommodate what was termed a multidimensional content architecture where content assets would be managed in a way that would simultaneously support many different standards. This goal was accomplished by deploying an extensibility framework based on the DITA.

Once in DITA, the content sources would be pulled into the three-dimensional design modeling environments, into the part selection applications and into the manufacturing control tools. In all of these environments, applications and tools would be operated by different suppliers working in various locations around the world and using software products provided by many different vendors. A sophisticated content-sharing architecture was established where content was dynamically accessed, modified, augmented and monitored across this global network of collaborators. Driving the sophistication of the architecture were considerations such as security, with the entire program operating under strict export controls, and performance, as necessitated by the fact that the design and manufacturing tasks needed to be coordinated on a near real-time basis.

Project success. Leveraging the new level of content intelligence they were able to move ahead with their design innovation goals while at the same time ensuring that the rich design knowledge available within historical repositories could be leveraged. They were able not only to maintain the required levels of control and oversight, but to take them to an even higher level. One of the benefits associated with content intelligence is the ability to apply very precise analytics to every step in the content lifecycle.

The types of aircraft that can be designed and manufactured using an intelligent content strategy are fundamentally superior to anything that has come before. The aircraft being produced are safer, more maintainable and much more economical to operate. And future aircraft design projects will have the benefit of starting from a far more intelligent content repository of historical knowledge and regulatory guidance.

Challenges. Finding the authoritative source for any given element of content was far harder than we expected, and once identified, the authoritative content sources were found to exist in a wide range of proprietary formats. Establishing reliable and cost-effective ways to extract the content sources from these legacy formats and to enrich them with the necessary intelligence proved to be a challenge. A number of technologies and techniques were introduced to overcome these obstacles. Authors and editors were also going to need specialized tools to handle these complex structures efficiently and effectively.

At the end of the project, one of the lead developers working on the solution confessed something to the client: “I have to tell you that many parts of this project were really difficult.” A senior technical representative from the client organization did not hesitate with his answer. “That’s OK, we thought it was impossible.”
Developing the Intelligent Content Strategy

Content models. The information modeling process [14] forces you to consider all information requirements (either for a specific project or within an entire organization) and to assess what information is available to fulfill those requirements. In an intelligent content strategy, the information model reflects the semantic structure of your information both at the information product level (for example, brochure) and at the element level (for example, value proposition).

Reuse strategy. A reuse strategy identifies what types of content will be reused, the level of granularity, how the content will be reused and how to support authors in easily and effectively reusing it. Your strategy will depend upon your goals, your content, your authors and your selected technology.

Taxonomy strategy. The taxonomy strategy enables you to intelligently store and retrieve your content based on a common vocabulary and shared metadata. In addition to traditional metadata for information storage and retrieval, it is important to develop metadata to define the delivery channel (print, web, wireless), the method of filtering the content (product, customer segment/audience, region, product version) and the final information product (brochure, web, eBook).

Creating intelligent processes (workflow). An intelligent content strategy also involves people and intelligent (collaborative) processes. Collaboration ensures that the content elements are consistent and can be reused wherever they’re required. Processes should be redesigned to match the intelligent content strategy and support the way the authors work. Workflow can be used to support these processes.

Implementing your strategy: The role of XML. Everywhere you go, you hear about the use of XML. XML is being used on the web, in rich media and for content. While you don’t have to use XML for your content, XML really helps make your content intelligent. Traditional office documents are simply files, and you have no access to the content because content is unstructured.

DITA. DITA, which has been mentioned above, is being adopted faster than any other XML standard today. It is an open content standard that defines a common content structure that promotes the consistent creation, sharing and reuse of content. DITA is supported by Organization for the Advancement of Structured Information Standards (OASIS) [15]. It was originally developed for technical documentation but it is now being adopted for business documents and pharmaceutical materials. It is also being used for eBooks.

DocBook. DocBook [16] has been around for almost 20 years. It began to lose ground with the advent of DITA, but the eBook revolution has revived it. Like DITA, it was originally developed for the technical documentation industry, but it was also adopted by organizations managing large volumes of content and the journal publishing industry. Business documents can be converted to DocBook relatively easily. The DocBook content can then be converted to EPUB [17], a standard promoted by the International Digital Publishing Forum for reflowable electronic books. DocBook does not support reuse as effectively as DITA, but it does provide a simpler conversion path from traditional business documents to XML.

The power of XML for delivery. When it comes to delivering content, XML gives us a very wide range of options. In fact, part of the rationale for XML was to liberate content owners from being limited to providing only one or two delivery formats. Once content is encoded with XML, its intelligence can be leveraged by automated publishing processes that can be put into place, and continuously refined, so that all of the output formats that the customers need can be produced with the push of a button. With the steady advances in the level of XML awareness in mainstream software applications and infrastructure components, it is becoming increasingly common for delivery processes to simply package XML-encoded content so that these tools can provide minute-by-minute views of the content.

Writing in XML. When the discussion turns to XML for content, there is often the concern about complexity. Certainly in the early days of XML, authors had to work with codes to tag the content, much in the same way early word processors forced the writer to display and use formatting codes as they created content, but author tagging is not necessary any more. XML can be hidden, providing a Word-like interface, or authors can even work in
Word with structured authoring supported by Word styles that are mapped to XML structure. XML does not need to be intimidating.

**Technology.** An effective strategy begins at the design stage, works through the authoring stage, ends at the delivery stage and is continually revisited to ensure it continues to meet the needs of authors, content and customers. When implementing your strategy, you need to assess how authoring, content management and delivery tools will help to support your intelligent content strategy.

**Authoring.** Before content can be managed, manipulated or reused, it must be created. To support an intelligent content strategy, content must be written so that it can be structured and reused according to the content life cycle. When evaluating authoring tools, give serious consideration to whether you should maintain your traditional authoring tools or move to XML.

**Content management systems.** Intelligent content needs an XML-aware system like a component content management system (CCMS) [18]. CCMS manage content at a granular (component) level of content, rather than at the page or document level. Each component represents a single topic, concept or asset (such as an image or table). Components are assembled into multiple content assemblies (content types) and can be viewed as components or as traditional pages or documents. Each component has its own lifecycle (owner, version, approval, use) and can be tracked individually or as part of an assembly.

**Delivery.** Delivery systems have many different capabilities. The content management system may have built-in facilities for delivering content, or you may have to integrate a delivery system with your content management system. Some delivery systems will enable you to deliver to a variety of outputs such as web, HTML, PDF, mobile or eBook while others may be restricted to a single output. Determine your delivery requirements, and see if your content management system will support them. And if you see an opportunity to deliver your content in a new way, you always know that, with your content in XML, you can add a new delivery option at any time. Perhaps you might add a new third-party component to your CMS or perhaps develop a new publishing process yourself to do exactly what you need.

**Closing Thoughts**

With the speed of change occurring in all industries and with the rate with which new devices for content consumption and interaction are proliferating, enterprises must consider seriously how they are going to make their content more intelligent and how they are going to continuously improve the information products they produce. We need only look at the dramatic change in the publishing industry to see how the business can be imperiled by having content locked into old formats and technologies. An intelligent content strategy is adaptable to the changes you will face today and in the future.

---

**Resources Mentioned in the Article**


Continued on next page
### Resources Mentioned in the Article, cont.


