



WHITE PAPER

The ABC's of Structured Content

CONFIDENTIAL



1-717-764-9720 | contact@vasont.com | www.vasont.com

As your organization expands product lines in different countries, the volume of technical documentation grows. Now, you must produce more versions of documentation in more languages, and in more deliverables (mobile, Web, print, etc.). You contemplate managing this growing content base with a content management system. A smart move! Content management systems save organizations lots of money and time in the editorial, translation, and publishing process.

So, you call up a content management vendor and ask for information. They begin to ask a flurry of questions that boggle your mind:

"Have you structured your content?" Huh?

"Is your content in XML?" What's XML?

"XML is a markup language for content." A what?

"A markup language is a set of tags that identify the elements." Oh...yes, my content is in InDesign, so it's got tags identifying heads and lists and such!

Finally, we're on the same page! Aren't we? *Well, no.* This conversation happens time and time again with technical writers and technical communications managers. They hear about the tools that can help them, but don't grasp the methodology that's behind it. So, let's review the basics in simple terms.

- A. *Content reuse is at the heart of content management.*
- B. *To maximize the reuse of your content, it must be structured.*
- C. *XML is the current standard for structuring content.*

So, what is XML and how does it differ from composition tags? My content is tagged as 1-heads, 2-heads, numbered lists, extracts, and so forth in my composition files isn't that structure? *No.*

FORMATTING TAGS VS. XML: WHAT'S THE DIFFERENCE?

Composition tags (also called formatting tags or style sheets) tell a specific publishing system *how to format* the content; i.e., "make this text 10/12 Times Bold." Formats give content the styling, or the "look and feel."

Don't try running those formatting tags on another publishing system...they aren't transferable. Each system has its own proprietary language. You must create a set of formatting tags for each publishing system you choose to use (i.e., FrameMaker, PubLink, InDesign, Dreamweaver).

Structured content has a hierarchical order and meaningful relationships. XML tags define this order and describe what the content is and the role it plays to many different platforms of software and hardware; i.e., "this is a Section and there must be one or more Sections in a manual." Unlike formatting tags, it does not describe what the content will look like upon output. XML provides a hierarchical structure to the content and enforces specific rules to keep the content orderly. Because XML is a standard markup language, it is able to be read and processed by many different publishing systems to create printed manuals, websites, mobile applications, and more.

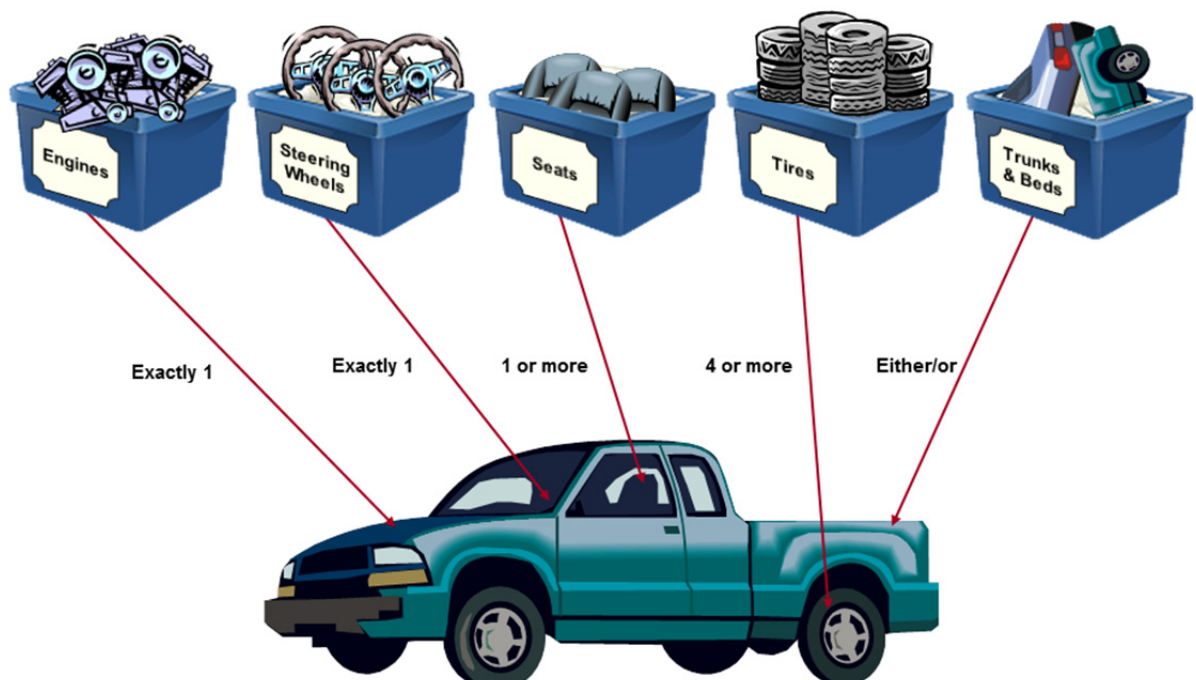
XML and formatting tags work hand-in-hand to publish content. XML ensures content is assembled properly; then, formatting tags apply the style to the content based on a designer's specifications.

AN ANALOGY: BUILDING VEHICLES

A manufacturing environment consists of an assembly line with parts that, when combined in the correct sequence, make a workable, quality product. For example, to build a vehicle, lots of parts are pulled from parts bins and assembled based on specific rules:

- ▾ One vehicle—a sedan—has an engine, a bench seat, two bucket seats, four tires, a steering wheel, and a trunk.
- ▾ Another vehicle—a truck—has an engine, a bench seat, six tires, a steering wheel, and a bed.

Even though these vehicles are different, they share some parts and some basic rules that allow them to be assembled properly: they must have exactly one steering wheel, exactly one engine, one or more seats, and four or more tires. It must also have either a trunk or a bed. Without these specifications for the construction of the vehicles, they will not operate properly.



Now that the mechanics of the vehicles are assembled, it's time to give them some aesthetics that make them look nice and feel comfortable for the customers. Specifications indicate:

- ▾ The sedan should be painted red and the interior should be tan leather.
- ▾ The truck should be painted blue and the interior should be gray cloth.

Table 1 shows a summary of the requirements to assemble and finish the vehicles.

TABLE 1: REQUIREMENTS TO ASSEMBLE AND FINISH VEHICLES

	REQUIREMENTS	CAR	TRUCK
PARTS:			
Engine	Exactly 1	1	1
Steering Wheel	Exactly 1	1	1
Seats	1 or more	2 bucket, 1 bench	1 bench
Tires	4 or more	4	6
Trunk, Hatchback, or Bed	Choose 1	Trunk	Bed
AESTHETICS:			
Exterior Paint	n/a	Red	Blue
Interior Color	n/a	Tan	Gray
Interior Fabric	n/a	Leather	Cloth

BUILDING DOCUMENTS IN A STRUCTURED ENVIRONMENT: A “MANUFACTURING” PROCESS

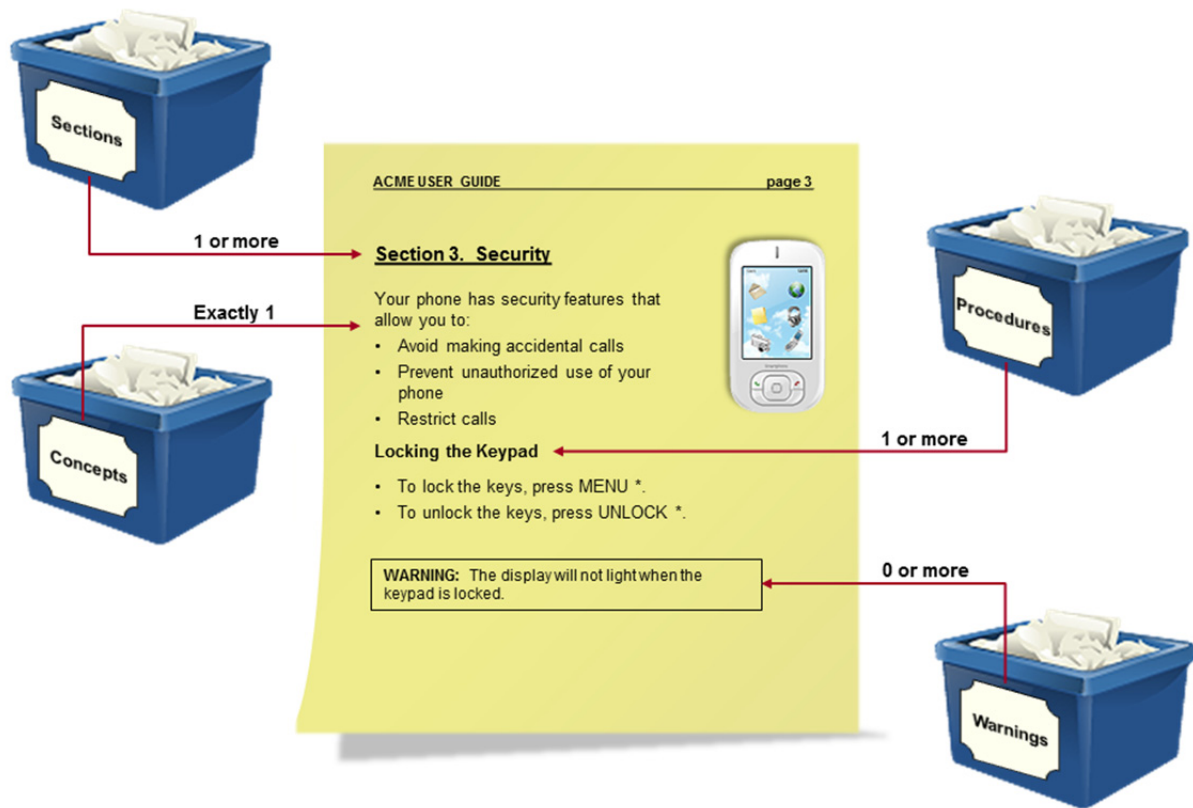
Technical documents are much like vehicles. Just as there are different types of vehicles (e.g., sedans, trucks), there are different types of documents (e.g., user guides, service bulletins, training materials). Like vehicles, the documents contain “parts” or elements (e.g., sections, concepts, procedures, warnings) that can be written independently and generically enough to be reused in many different documents. These elements have requirements within the documents so they are assembled properly (i.e., a manual must have at least one section; a section must have exactly one concept and one or more procedures; warnings are optional).

XML is the tagging scheme that defines these elements and their requirements.

Consider a user guide for a cell phone. XML defines the structure of this document. It is required to have at least one Section. Each Section must have exactly one Concept that explains what the Section is about, but this Concept must contain no more than three main points or items. There must be at least one or more Procedures in the user guide to explain “how to” do something step-by-step, and each Procedure must have at least two steps. (It wouldn’t be much of a Procedure if it only had one step!) The user guide may contain Warnings, but it’s not required.

In the vehicle example, the specifications allow the manufacturer to assemble many vehicles in the same manner in order to efficiently create operational, quality products. Likewise, XML requirements give all documents for our cell phone a consistent structure and hierarchy, regardless of which writer develops the content. When a customer reads your documentation, it is orderly and understandable; in other words, you are delivering a quality document.

These rules become very important in a topic-based writing environment, where many writers may be creating reusable topics that could be used in many different documents. When combined, the document must follow these rules to provide a consistent structure and a single tone of voice. These rules can also be applied to the cell phone’s training materials and other documents as well to maintain consistency across the content base.



When your document is assembled (written, reviewed, and approved), you apply the aesthetics or the “look and feel”—just like applying the paint to a vehicle.

Formatting tags provide the aesthetics for the content.

The formats, or style sheets, created in InDesign, FrameMaker, PubLink, or any other publishing system, are the mechanisms that apply the style to the elements (i.e., font, point size, bold, italic, indent, text width, text color). The publishing step—the step in which the design is applied to the content—doesn’t happen until the content within the document is developed, reviewed and approved. Once the content is approved, the aesthetics can be applied during the publishing process to make it look “pretty.” This process is often automated, eliminating the need for writers to apply the styling.

Despite being the last step in the document assembly process, the design is a critical part of this process. It conveys the tone of the document to the customer. Using different colors, fonts, lines, and other design elements, you can change the feeling a customer gets about your product. For example, a colorful user guide with a playful font may give the impression that your cell phone is marketed towards a younger generation. These design elements are applied using style sheets in the publishing tool you choose to use (Cascading Style Sheets on the Web or tool-specific formatting tags).

Table 2 provides a summary for the requirements for the manuals in our cell phone example.

TABLE 2: REQUIREMENTS TO ASSEMBLE AND FINISH MANUALS

	REQUIREMENTS	USER GUIDE	TRAINING MATERIAL
PARTS:			
Section	1 or more	6	10
Concepts	Exactly 1 with 3 or less items	1	1
Procedures	1 or more containing at least 2 Steps each	3	15
Warnings	Optional	2	0
FORMATTING:			
2 Heads	n/a	12/14 Arial Bold	14/16 Futura Condensed Bold
Text	n/a	10/12 Times	12/14 Garamond
Bulleted Lists	n/a	10/12 Times, 3 pts. after	11/13 Garamond, 6 pts. after

Building a document is similar to the manufacturing process used to build vehicles. There are parts that need to be assembled in the proper order, and then the aesthetics are applied at the end of the “assembly line” (during the publishing process) to complete the deliverable.

ARE THE XML TAGS AND FORMATTING TAGS THE SAME?

No. In the cell phone example, the chunk of text, “Locking the Keypad,” is tagged as a *Procedure* and *Steps* in XML, but might be tagged as *2H* (2 head) and *BL* (bulleted list) in formatting tags. The tags serve different purposes:

- ▼ The XML tag *Procedure* indicates that this chunk of text contains a set of *Steps* telling the user how to do something, and there must be on or more Procedures in a Section.
- ▼ The *2H* tag tells a publishing system to style that text string as 12/14 Arial Bold, flush left; while the *BL* tag tells a publishing system to style its text string as 10/12 Times, flush left, add the bullets and allow 3 points of space between bulleted items.

Remember, XML focuses on hierarchical structure of the content elements while formatting tags focus on the appearance or design of the content. While the tags themselves may be similar, they perform very different functions.

WHY IS XML IMPORTANT?

The XML tags define what the content is and how it relates to other elements. In the cell phone example, notice that in addition to the bulleted list in the Procedure, there is also a bulleted list in the Concept. Because we want both bulleted lists to look the same in the publication, they are tagged as *BL* in the formatting tags. But in the XML, they are different elements with different requirements:

- ▼ In a Concept, a bulleted list can reside within a Section, but it can contain exactly three or less items.
- ▼ In a Procedure, a bulleted list must contain two or more items (Steps).

Working in a structured authoring environment means working with documents in a different way. Content and design are processed as two separate steps. Content is developed, reviewed and approved before the design is applied. This allows writers to *focus on the quality of the content* rather than the design of the document.

Structured authoring can be a challenging change for writers who are used to applying bold, color, and other styling to the content as they develop the content. It requires a shift in the mindset of technical communicators.

SUMMARY

XML performs the following important functions:

- ▼ It defines relationships between content.
- ▼ It provides a hierarchy to content and enforces a specific order.
- ▼ XML makes your content portable and accessible because it is a standard that can be read by many different hardware and software platforms.
- ▼ XML allows your content to be published to many different channels since the XML tags can be automatically transformed into formatting tags.

Structured authoring provides the following benefits:

- ▼ It forces writers to focus on the quality of the content.
- ▼ Content creation is more efficient since writers are not responsible for applying the design.
- ▼ The publishing process is more efficient since it is often automated.
- ▼ Content is more comprehensible for your customers.

XML establishes a consistency to content, and enables it to be managed and reused in a content management system. In organizations with many writers, it is important to maintain consistency and accuracy of the content throughout the writing process so that when content is reused and assembled into documents, it flows well for the reader. XML is the standard markup language that provides this order.

Vasont Systems provides component content management software and XML data services to Fortune 1,000 companies and global organizations from a variety of industries, including manufacturing, technology, publishing, financial, and healthcare. Since 1992, Vasont Systems has helped companies organize their critical business assets, manage productivity, and disseminate information in many languages to multiple media channels while reducing their costs by an average of 63% and shortening overall cycle times by 75% on average. The Company has been named multiple times to the EContent 100 list of "best and brightest digital content companies." Vasont Systems is a member of the [TransPerfect](#) family of companies. For more information, visit www.vasont.com.