Enterprise Content Management (ECM) Strategy

Structured Authoring

August 11, 2004

What is Structured Authoring?

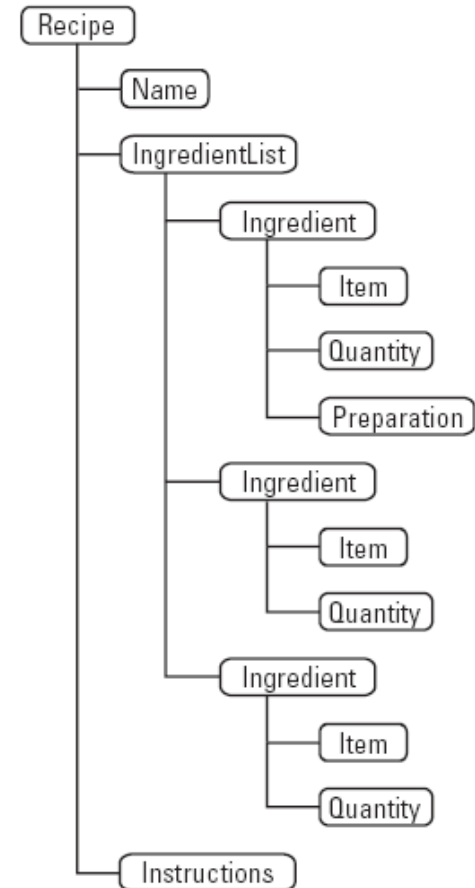**Structured Authoring is the process of creating content that is machine parsable.**

What is machine parsable?

machine: an automated process

parsable: the ability to resolve (as a sentence) into component parts of speech and describe them based on a set of rules
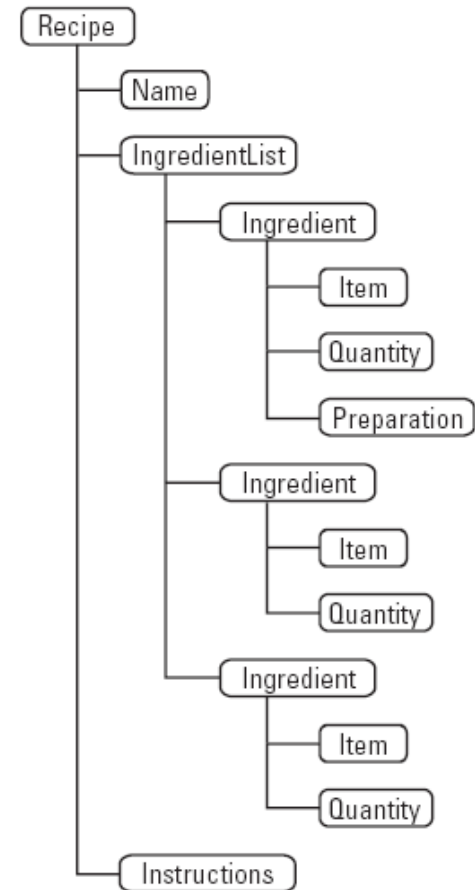
How is this done?

- **STYLE GUIDE -** Technical Writers use this document for advise on formatting document components such as bulleted lists, copyrights, tables and page layouts.

- **STRUCTURE DEFINITION DOCUMENT** - A file in the form of a Document Type Definition (DTD) or XML Schema is used in structured authoring to store similar rules. The authoring tool validates that the content conforms to the the Structure Definition Document through an automated process.

- **ELEMENTS** - A structured document has its content organized into elements. These elements are marked by "tags". It is these tags that allow a document to be parsed or read by a machine which in this case is an automated process.

- **XML** - Extensible Markup Language is a specification that describes the use and properties of elements.

- **RECIPE EXAMPLE** - Elements are usually organized into hierarchical trees. In the example at right a recipe has ingredients, quantities and maybe individual preparation methods.

Recipe
- Name
- IngredientList
  - Ingredient
    - Item
    - Quantity
    - Preparation
  - Ingredient
    - Item
    - Quantity
  - Ingredient
    - Item
    - Quantity
- Instructions

How is this done? (continued)

- **Rules -** The Structure Definition Document in this example would require each recipe must have a name, ingredient list and instructions. It also might require them to appear in that order for every recipe. Each Ingredient must have an item and quantity nut the preparation could be optional.

- **Reuse** - Any particular ingredient may be used as part of multiple recipes. Suppose there became available a substitute ingredient that because of health issues is preferable for a particular version of the recipe such as a low-carb version. Structured Authoring would make the automatic replacement of a particular ingredient possible.

- **Concurrent Authoring** - Because the recipe has been broken into smaller parts different authors could work on separate parts at the same time.

- **SUMMARY** - In unstructured authoring the parts of a recipe would be determined through Fonts, Indentation or Headings. In Structured Authoring this is determined through the use of elements represented by tags.

  Appearance is added **separately** based upon the output.

Recipe
— Name
— IngredientList
  — Ingredient
    — Item
    — Quantity
    — Preparation
  — Ingredient
    — Item
    — Quantity
  — Ingredient
    — Item
    — Quantity
— Instructions

Content Reuse further Explained

- **CONTENT REUSE - Is one of the major benefits of Structured Authoring**

- **INFORMATION MODELING** - Before creating content the technical writer would consider how this content would be used. Information Modeling is the analysis required to maximize reuse for any content created. The following example involve a fictional company called Reo Auto Company and is extracted from a document entitled **Enterprise Content Management: The Next Frontier (**http://www.webreference.com/internet/enterprise/index.html)

A new car has been announced and the Marketing Department must now create ancillary documentation. Each information product requires different content and design.

The table at right shows a portion of the information modeling that would take place.

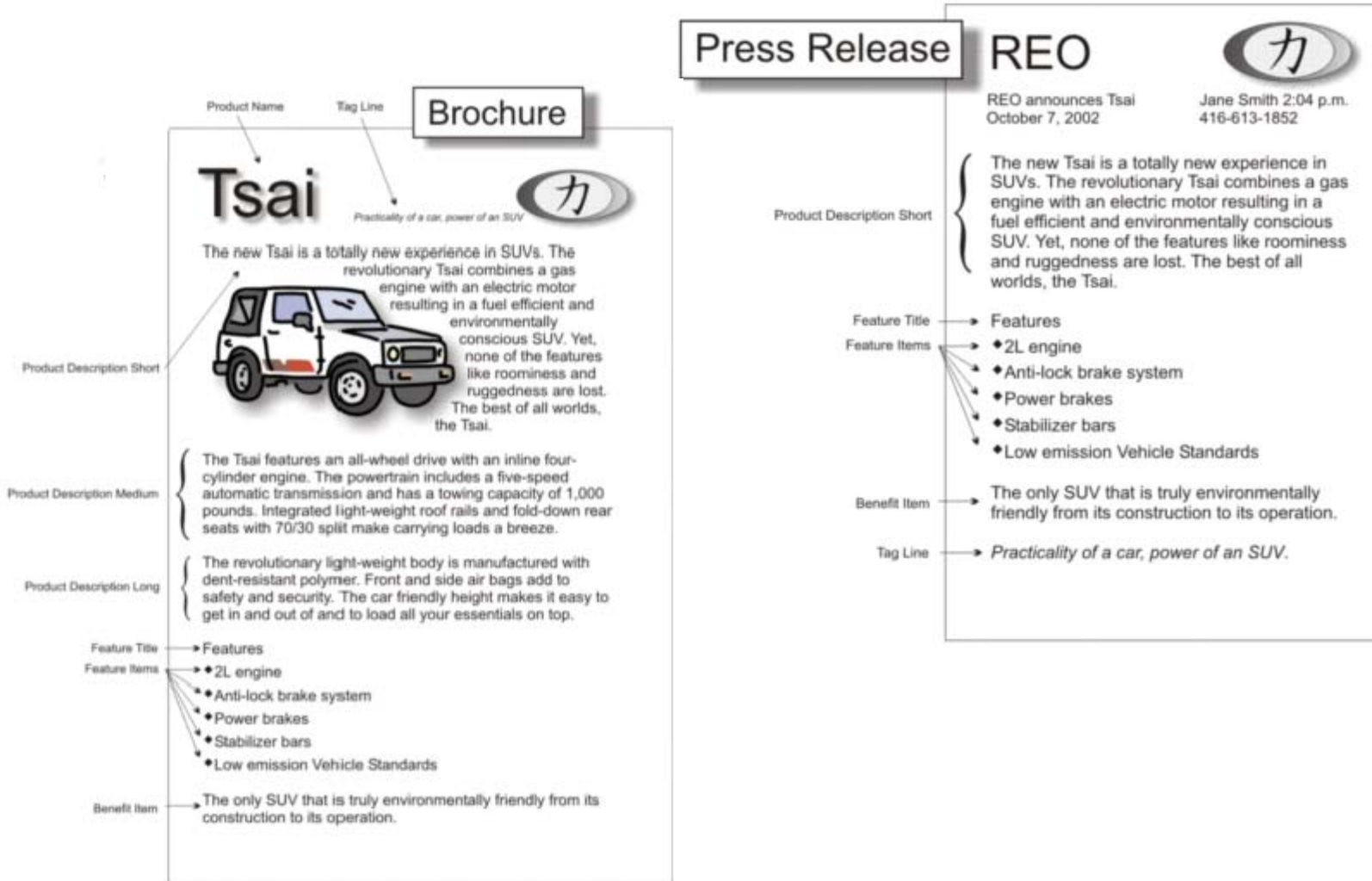In a Structured Authoring environment only one structured document would be created for all four information products

Product Description Element	Show catalog	Brochure	Press release	Web site
Product Name	X	X	X	X
Product Description				
Product Desc. Short	X	X	X	X
Product Desc. Med		X		
Product Desc. Long		X		
Graphic	X	X		X
Features	X	X	X	X [2]
Feature title		X	X	
Feature item	X	X	X	
Benefits		X	X	Xa
Benefit item		X	X	
Tag line		X	X	X

Content Reuse further Explained (continued)

- **METADATA** - Data about data that enables effective retrieval, automatic population of existing elements into placeholders for content reuse (systematic reuse), dynamic content, automatic routing based on workflow status and reporting.

- **ELEMENT METADATA** - The table at right shows a simplified metadata for a usage element.

- **TRANSFOMATION** - An automated process that would include or exclude portions of the document at publishing time based on the intended output.

- **SUMMARY** - All of the components would be created and stored together and used where appropriate.

Element	Metadata
Product Name	All
Product Description	
Product Desc. Short	All
Product Desc. Med	Brochure
Product Desc. Long	Brochure
Graphic	Show catalog Brochure Web site
Features	All
Feature Title	Brochure Press release
Feature item	All
Benefit item	Brochure Press release Web site
Tag line	Brochure Press release Web site

Content Reuse further Explained (continued)

Content Reuse further Explained (continued)

Structured Authoring - Pros & Cons

PROS:

- **RE - USABILITY**

- **IMPROVED DATA INTEGRITY**

- **LOWER DISTRIBUTION COSTS**

- **APPEARANCE EASILY CHANGED**

- **CONCURRENT AUTHORING**

CONS:

- **HIGH STARTUP COSTS**

- **HUGE INITIAL TIME COMMITMENT**

Structured Authoring

Questions & Answers