

XML(XHTML1.1+管理)データのWeb入力・データ作成システム

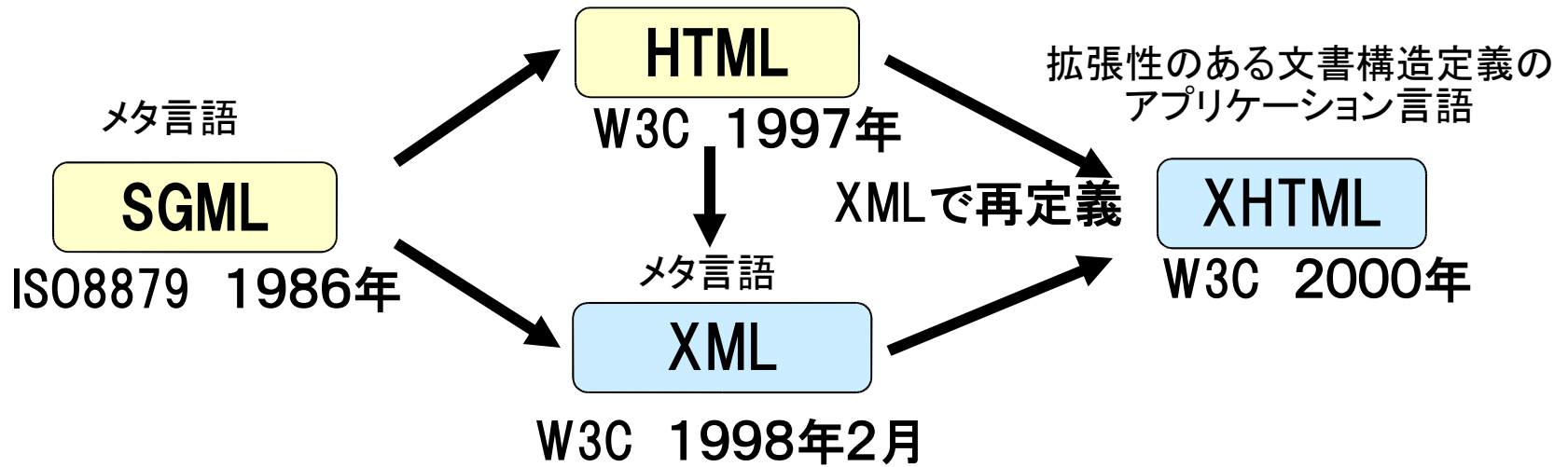
Web-Document::Maestro (XHTML1.1)



株式会社 エーティーエル システムズ

XML と XHTML の関係

ブラウザで表示可能なアプリケーション言語



HTML 4.0

Strict

Transitional

Frameset

現状の主流

XHTML 1.0

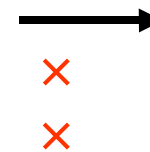
Strict

Transitional

Frameset

XHTML 1.1

Module化

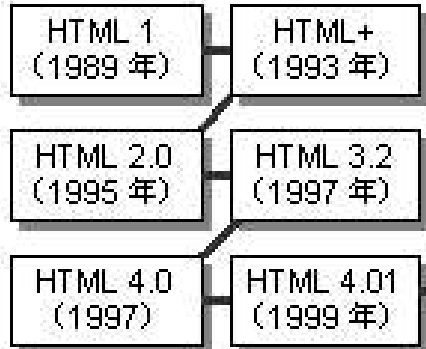


HTMLは表示が中心であり、データを厳格に決めるStrict型は使用する意味が無かった。ほぼ全てが、Transitional型。Framesetは、枠組み定義。

HTMLの各文書型に対応してXML化。再構成。Webブラウザで見ることが可能でかつXMLへの橋渡し。

Strict型を再構成。
データとスタイルの分離。
CSSは必須

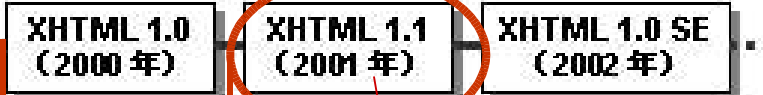
【マークアップ言語】



構造情報と表示スタイル情報を分離

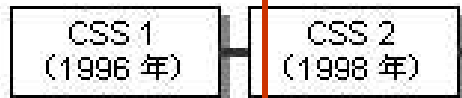
1999年・W3C
WCAG1.0

2004年
JIS規格策定
JIS X8341-3



完全にXML化

【スタイルシート言語】



高齢者・障害者等配慮設計指針
情報通信における機器
ソフトウェア及びサービス
第3部：ウェブコンテンツ



1998年・米国
リハビリテーション法508条
アクセシビリティ基準作成開始

2000年・総務省
14項目の指針

© CYBER@GARDEN

実施：2001年

◆ 文書は、意味を表現する構造と表示スタイル情報で構成される

- 文書構造情報は、全ての文書で同一
 - ◇ 見出し、段落、リスト、テーブル… 強調、ルビ…
- 表示情報は、デバイス(印刷、携帯、放送…)で異なる
 - ◇ 表示情報: フォント、罫線、色情報…

個別の表示情報を含むデータは変換できない

文書コンテンツを他用途に利用可能になる

◆ XMLは、本来表示を考えないデータ ⇒ XHTML

- XSLで、目的の構造・表示に合ったデータに変換
- WebではCSSとXSLを使うことが可能

Webの場合は、プログラムで表示を変更可能

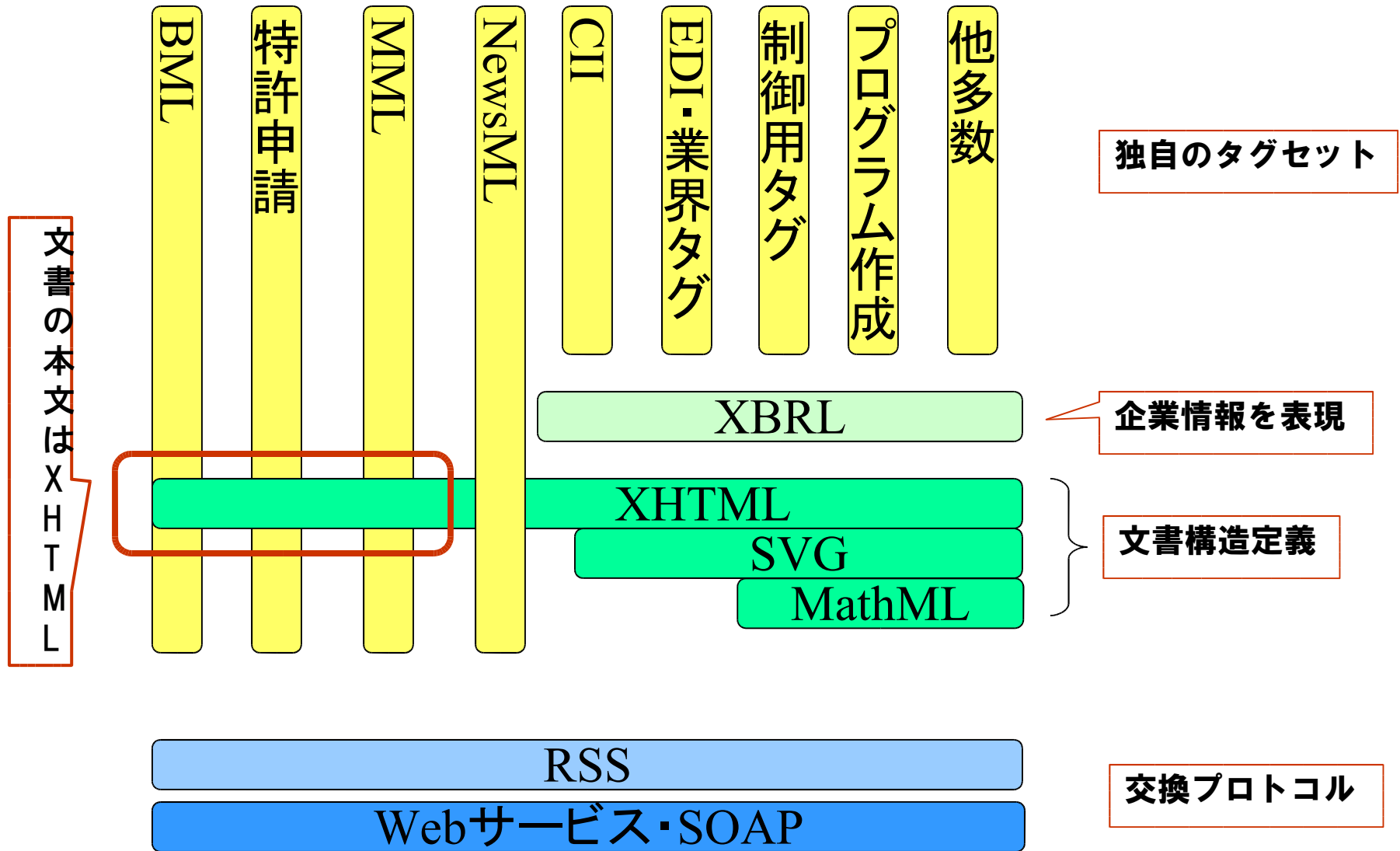
◆ XHTML2.0

- 画像、動画、音声等は、それぞれのObjectとして統合

マルチメディアとして統合化

文書コンテンツを統合化

XMLファミリー関連図

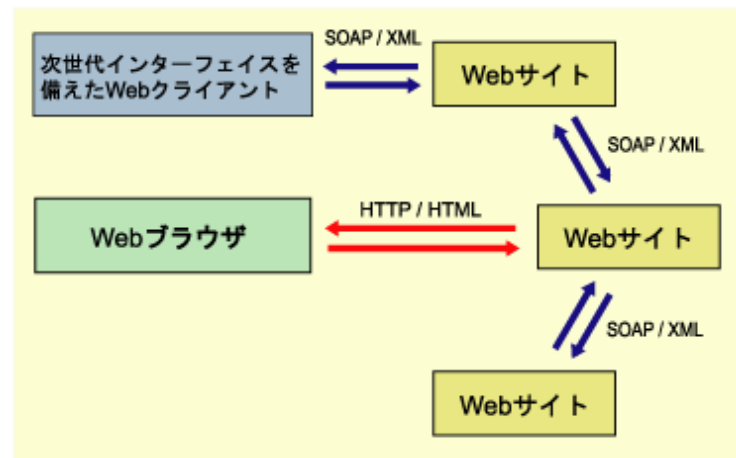


XMLのテクノロジーを外部とのインターフェイスとして全面的に採用し、人間を介することなくWebサービスが他のWebサービスとやり取りすることを可能にする。

•Webサービス (XML Web Service)

XML形式でデータ交換を行い、データ・アクセス・プロトコルとしてはSOAP (Simple Object Access Protocol) を使用する。下位プロトコルとしてHTTPなどを使用し、簡単なXMLベースのメッセージをやり取りすることで、RPC (Remote Procedure Call) のようにリモート・マシン上のサービス・ルーチンを呼び出したり、オブジェクト (データ) にアクセスしたりできるようにする。

利用例



•RSS(RDF Site Summary)

サイトの概要をメタデータとして簡潔に記述するXMLフォーマット。サイト間でデータ更新などの情報を交換。

•NewsML

ニュース管理の標準フォーマット。XMLをもとにしており、テキスト、写真、音声、動画などのマルチメディア素材を配信するのに適しています。

•ロゼッタネット (RosettaNet)

企業間電子商取引のためのインターフェースや用語の標準化を行っている団体。基盤となる言語としてXMLを採用し、企業間商取引を効率化して世界規模でネットワークを構築することを目指している。

◆コンテンツを管理するとは

- コンテンツの管理とコンテンツ制作のワークフローの管理
 - ◇ 文書データ(内容)作成 作成者 時間 版管理・履歴
 - ◇ 文書作成者に関する情報 起案→決済→施行→登録→保管→保存→廃棄
- コンテンツの構造化 枠組み(テンプレート)の必要性
 - ◇ 要約 見出し キーワード

◆コンテンツの多用途利用 作成後の利用のニーズ分析

- ワンソース・マルチユース
- 検索エンジンの高機能化

◆テーブルやリストは理解し易いためのタグ

- 管理タグは<見出し>程度 ?
 - ◇ XHTMLの<body>以下のタグの意味
- 特許申請のXMLやNewsML、BMLなどが手本
 - ◇ 本文はXHTML1.1を利用

XMLコンテンツであると、多用途利用が可能

ContentXML (ATL独自仕様のXML)

XHTML1.1
対応ファイル

携帯電話
Htmlファイル

PDFファイル

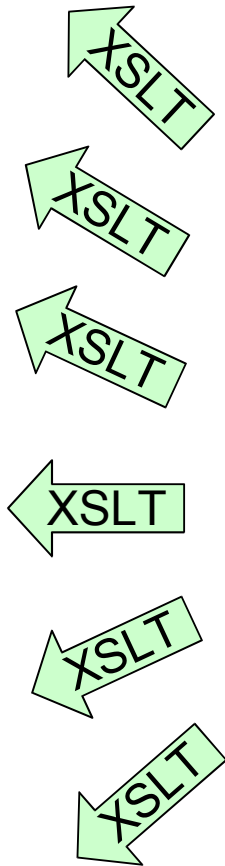
SGML

放送

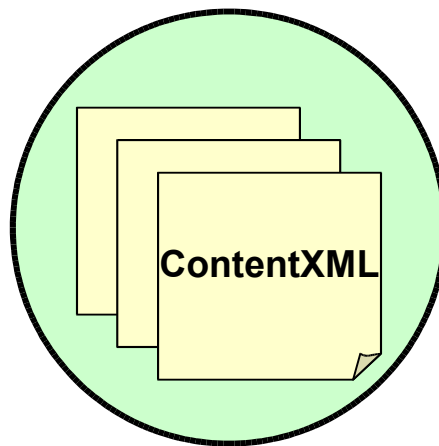
出版

電化製品

PDA



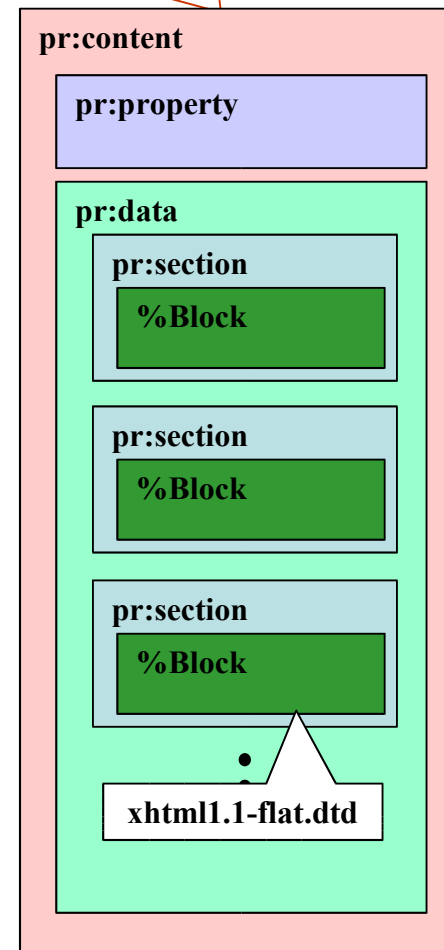
スタイル情報を外し
文書構造のみを定義
したXML



**XMLデータが作成でき
れば他用途利用が可能**

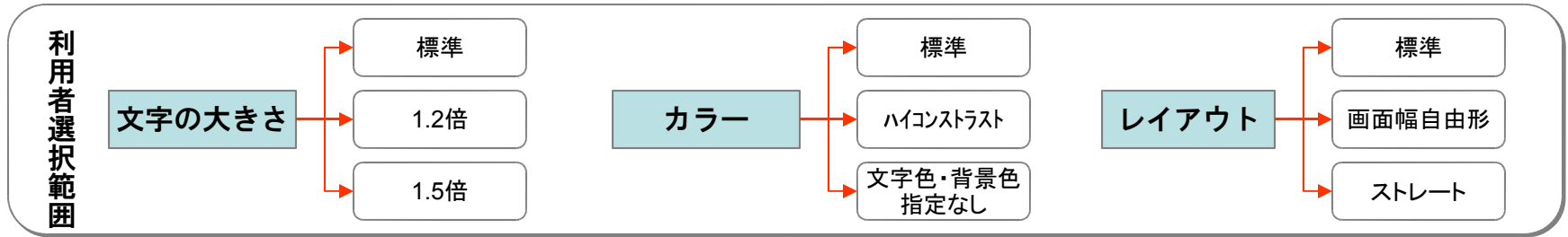
特許申請中

**出来る限り細かいデータ
を作成することが重要**

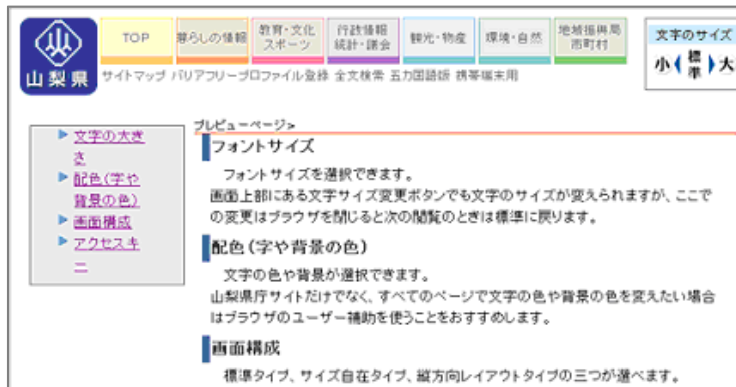


ContentXML

利用者が自分に最適な表現を選択できるシステムにより利用者の負担が少ない情報発信を可能にします。JIS X8341-3にも準拠しています。



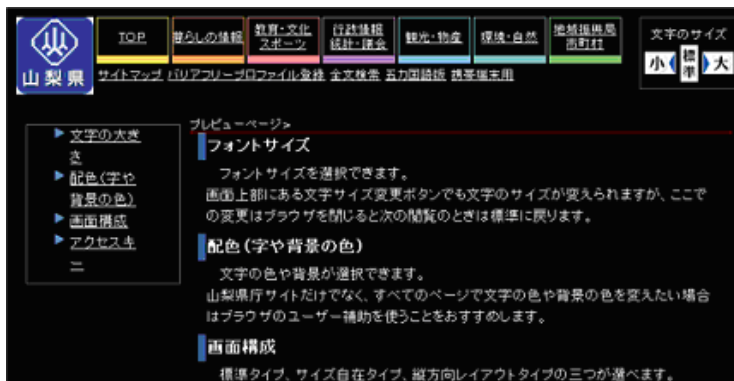
■フォント標準:カラー標準:レイアウト標準



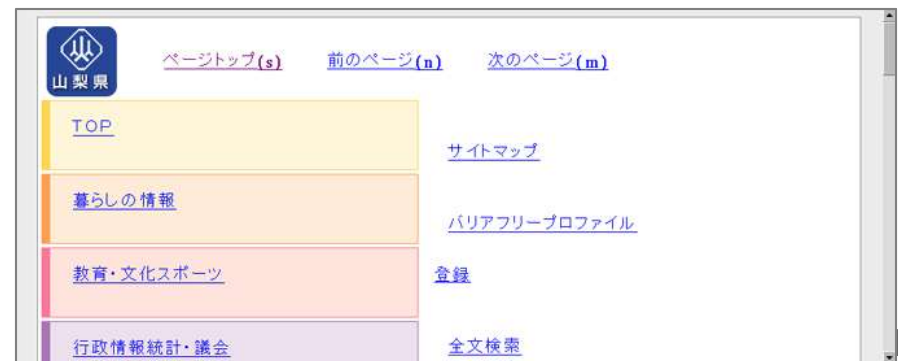
■フォント1.2倍:カラー標準:レイアウト標準



■フォント標準:カラーハイコントラスト:レイアウト標準



■フォント標準:カラー標準:ストレート



企業にとってもデータは同じで、利用者ニーズに合わせてスタイルを変化させる意味・効果

Business mode



Human mode



5.1 規格及び仕様

XHTML等の国際規格の構文に合致すること
⇒バリデート(妥当性確認)

a) ウェブコンテンツは、関連する技術の規格及び使用に則り、かつそれらの文法に従って作成しなければならない。

5.2 構造及び表示スタイル

文書構造とスタイルを分離 多用な表示
⇒XHTML1.1+CSS

a)・・・文書の構造を規定しなければならない。

b)・・・表示スタイルは、文書の構造と分離し、・・・スタイルシートを用いて記述することが望ましい。

5.3 操作及び入力

文書データにキーボード用データを付加
⇒XSLT

a)・・・キーボードによって全ての操作が可能でなければならない。

f)・・・表示スタイル及び表記に一貫性のある基本操作部分を提供することが望ましい。

h)・・・読み飛ばせるようにする

5.6 文字

スタイルは、全てCSSで定義する

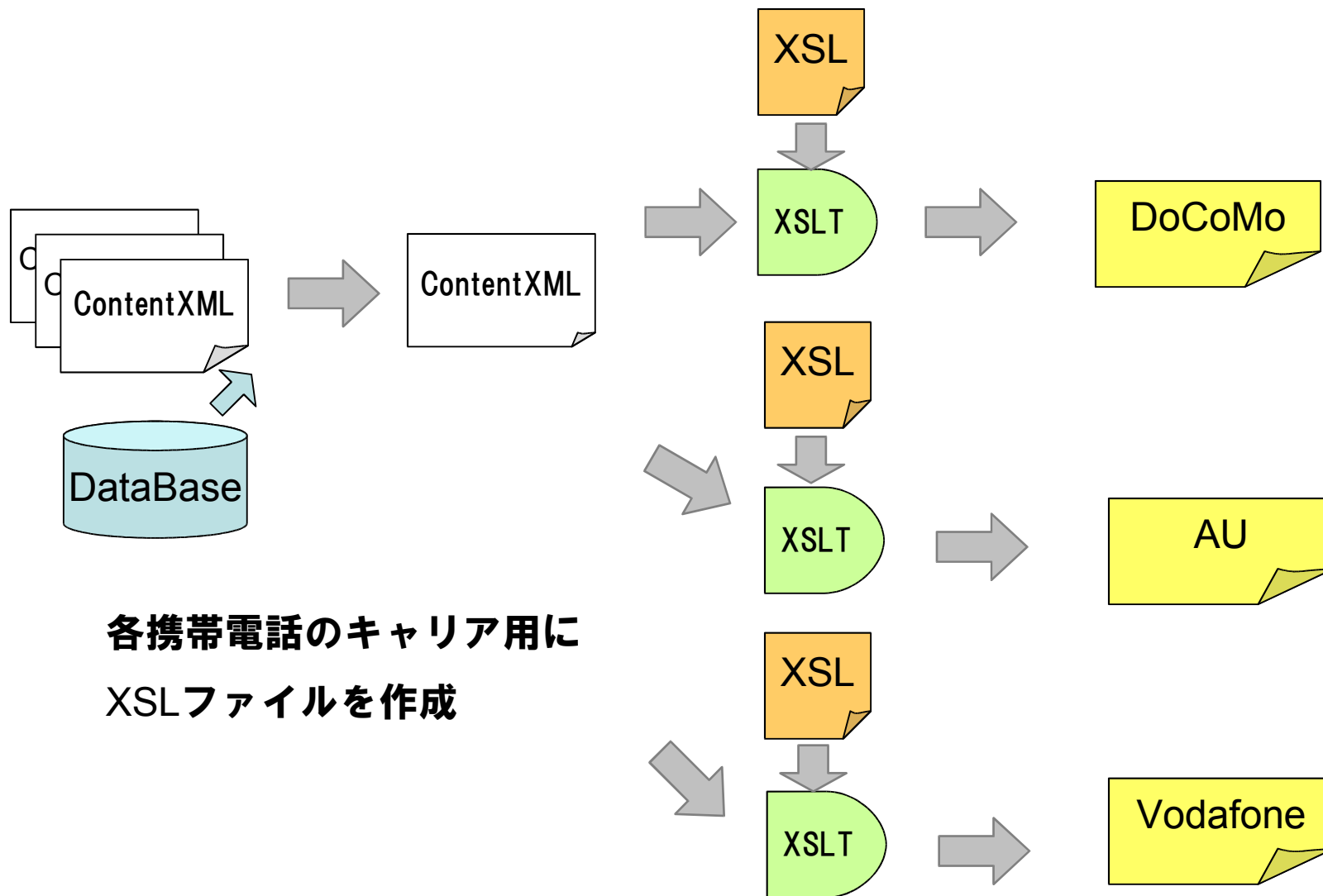
a) 文字のサイズ及びフォントは、必要に応じ利用者が変更できるようにしてはならない。

例: "em""%"を使用してスタイルシートで指定する。

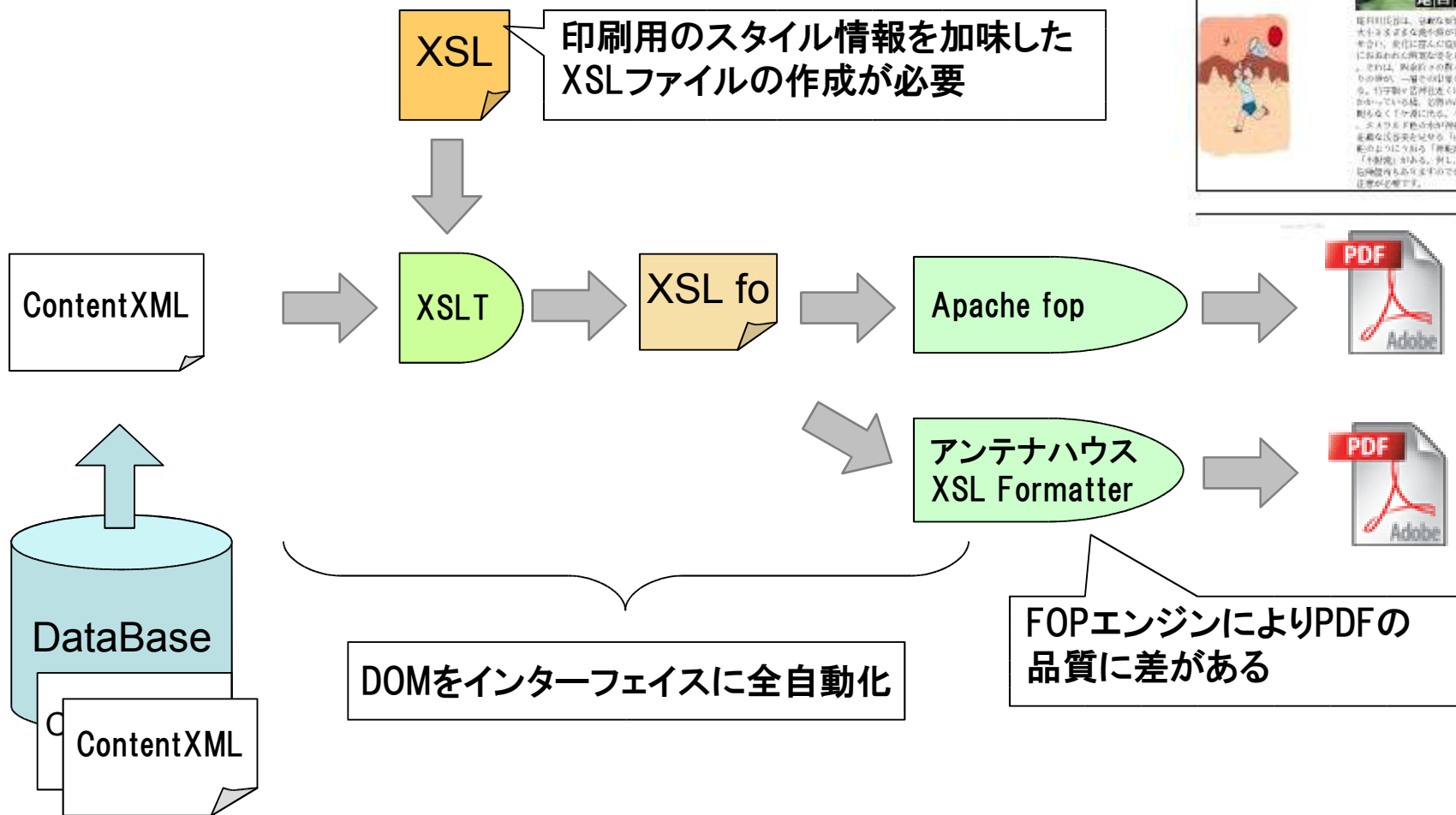
・ユーザ志向
・紙では不可能

多様な読者に合わせて表示スタイルを変更する

携帯電話用のデータを生成方式



文書構造データより直接PDFを生成する方式



XSL 印刷用のスタイル情報を加味した XSLファイルの作成が必要

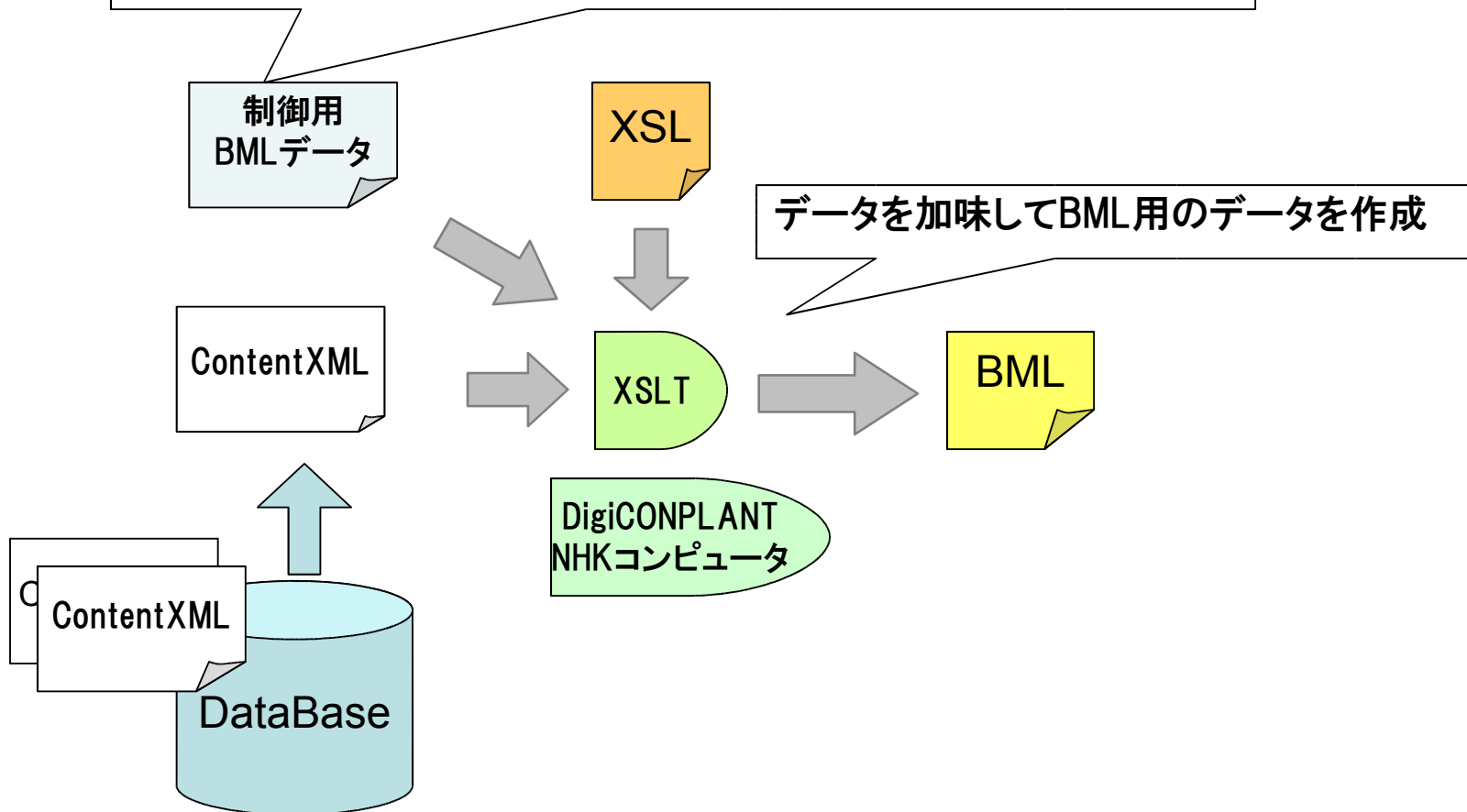
DOMをインターフェイスに全自動化

FOPエンジンによりPDFの品質に差がある



XMLベースのBSデジタル放送向けデータの作成の方式

画面のどこに何をいつ表示するかという空間/時間的提示の制御、
関連する番組やインターネットのURLを示すリンク、
タイトルや本文などを示す構造の表現、
操作ボタンなどのユーザインターフェースなど



XMLデータの作成の新しい手法

- ① 低コストで文書データを「構造化」
- ② サーバによるデータの集中管理

1. 非構造化コンテンツの管理手法

- 企業・官公庁の扱うデータの85%は非構造化データ
- 文書コンテンツの構造化はコストが掛かる
- MS-Office のデータなら誰でも読める

・構造化を放棄

・データの分散

2. 構造化(SGML)コンテンツの管理手法

- システムコストが高い
- 運用・メンテナンスコストが高い

Webシステムのみを構築するための手段

バリアフリー化対応は難しい

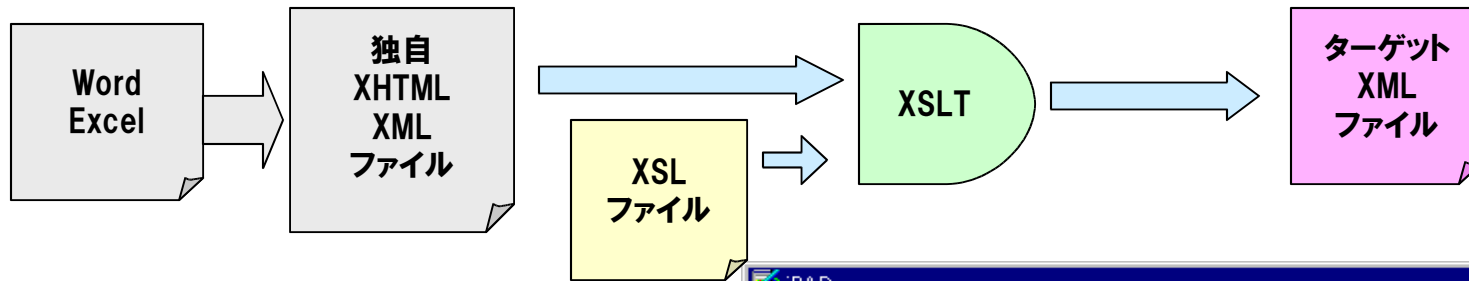
◆ Webサイト構築専用のCMS

- ✓ 一つの文書データを分類・細分化しRDBへ格納
- ✓ テンプレートでHTML4.0、XHTML1.0のファイルを作成

XMLで構造化 ⇒ 効率的に保管、検索、参照、配布

XMLのデータを作成（今までの手法）

Word・Excel+テンプレート

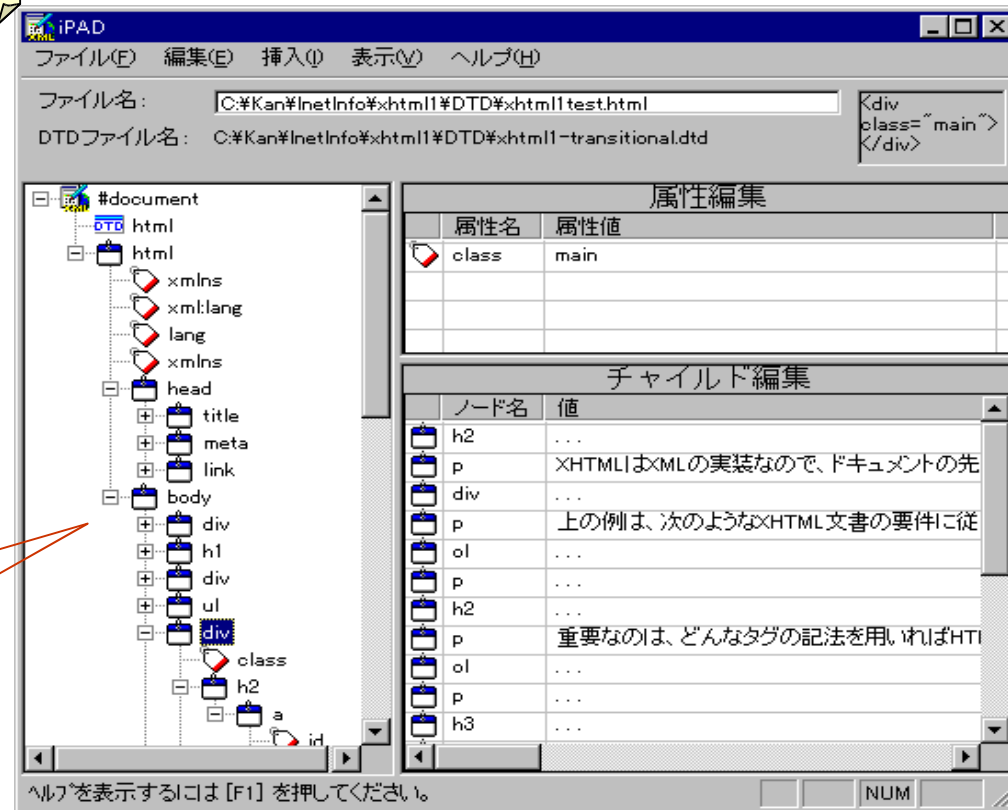


正しいタグに変換できるか入力後確認が必要

タグ・エディター

タグを理解できないと入力できない

XHTMLは100以上のタグがあり、理解するのは不可能

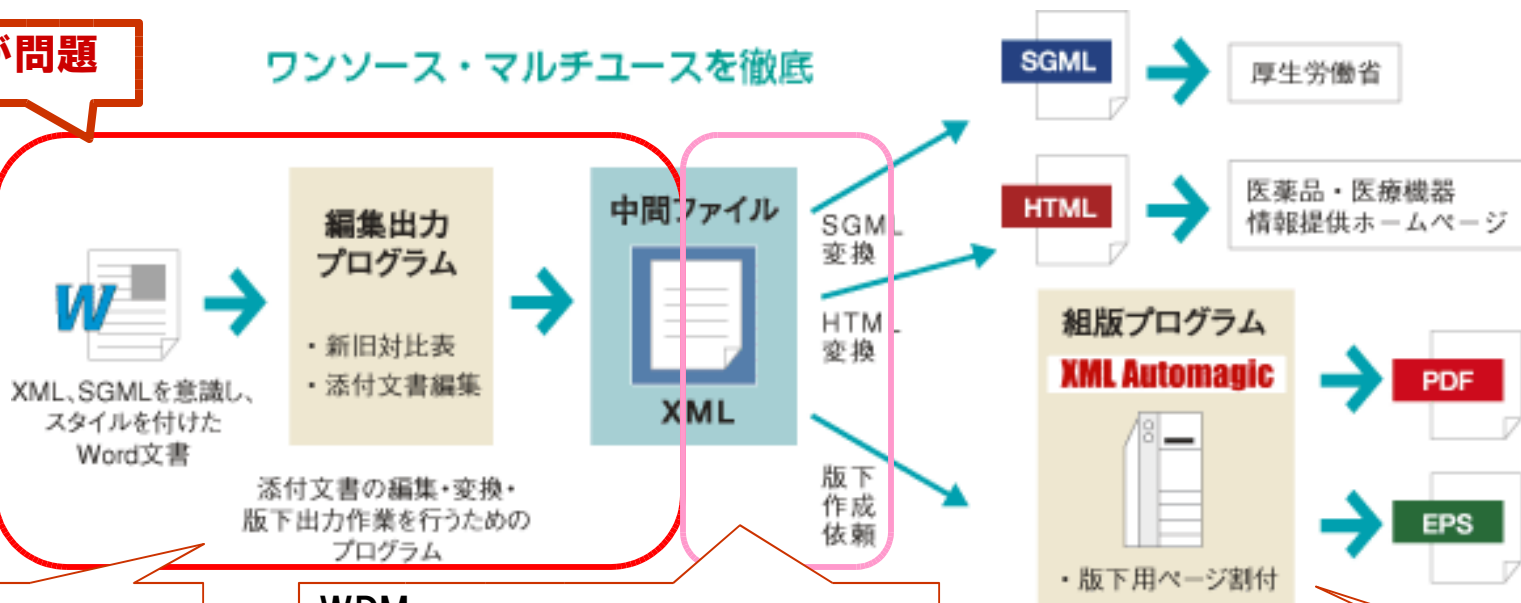


富士ゼロックスのコンテンツ管理システムの概要

XML ドキュメント管理データベースとしてDocumentumを採用すれば、更新時のワークフローによる承認、作成中・承認済みなどの文書のライフサイクル管理、改定時のバージョン履歴管理、アクセス権管理など、ルールに従った業務運用を行うことができ、薬事法改正に伴う医療機器向けISO13485や ISO 9001の取得に向けたドキュメント管理を行う事が可能です。(Webページより)

ここが問題

ワンソース・マルチユースを徹底



WDM

Webブラウザからのデータ入力により大幅コスト削減

WDM

XHTMLのため変換の工数が削減

添付文書コンテンツをDocumentumにて管理
承認ワークフロー(フローの履歴管理)、ユーザーアクセス管理、コンテンツ版管理

FOPと比較

1. GUIによるデータ入力

このツール用に決められたスタイルシートを使用 スタイルは変更可能

- 文書構造タグを意識しないグラフィック・ユーザ・インタフェースを採用
- XHTMLのタグを利用することで、ブラウザで再描画

2. XMLの規格の範囲で独自のXML(ContentXML)を策定

- 「Namespaces in XML(名前空間)」を利用
- XHTMLのタグを内包する独自XMLを定義
- 管理タグと文書構造タグを一体化
- パーサー、XSLTやFOPを利用

XHTMLを使用することで、社会共有資産を利用

3. 文書データをサーバ上のプログラムで管理

- 文書構造の定義とコードの生成 プログラム化
- 入力時に文書構造をチェック(バリデート)

W3C XML Schema は複雑すぎて実装できない

4. データはファイルでもXMLDBでも

- データ一元化

ツリー構造のXMLデータはRDBには合わない

コンテンツの構造化:コンピュータによるデータ管理

Webブラウザ上で文書構造・タグ情報をスタイルに合わせてタグ情報を生成
生成されたタグ付きデータをサーバに送信



構文検証チェック



Webブラウザから入力

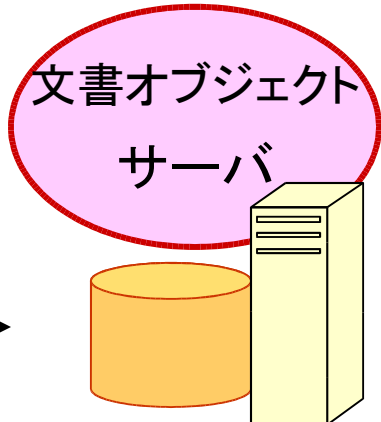
制作されたタグデータを指定されたスタイルシートで再描画



検証・更新

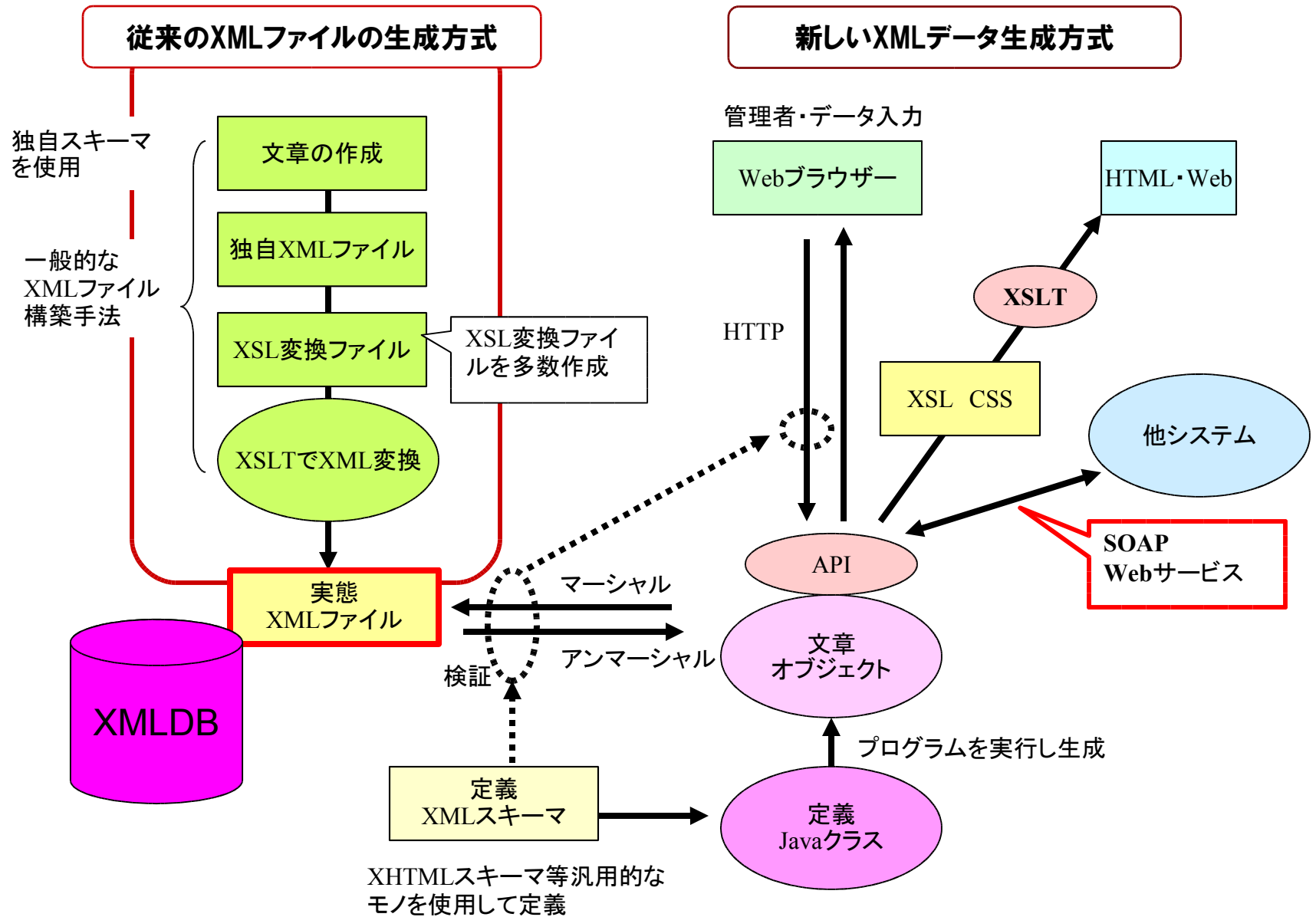


直接タグを入力することも可能

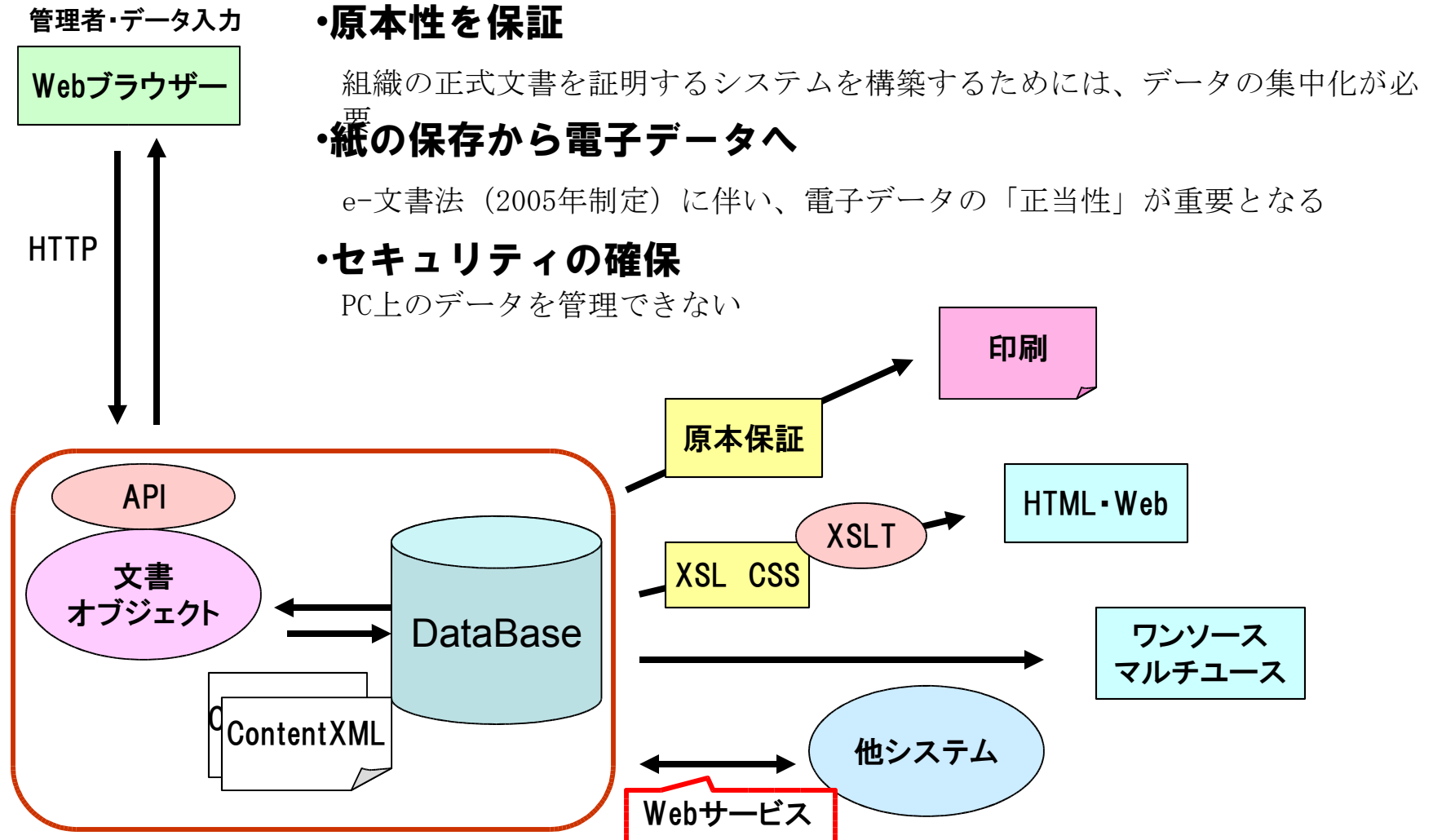


特許申請中

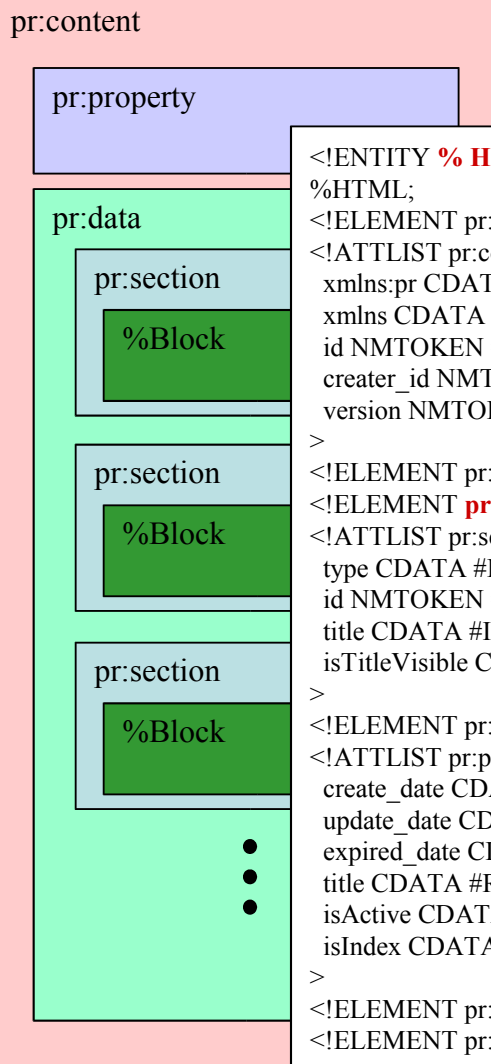
XMLデータの生成方式の比較



サーバで文書データを集中管理



複数のスキーマ(文書構造)を使うことが可能



```
<!ENTITY % HTML SYSTEM "xhtml11-flat.dtd">
%HTML;
<!ELEMENT pr:content (pr:property, pr:data)>
<!ATTLIST pr:content
  xmlns:pr CDATA #FIXED "http://www.atl-systems.co.jp/wdm/pr"
  xmlns CDATA #FIXED "http://www.w3.org/1999/xhtml"
  id NMTOKEN #REQUIRED
  creator_id NMTOKEN #REQUIRED
  version NMTOKEN #IMPLIED
>
<!ELEMENT pr:data (pr:section+)>
<!ELEMENT pr:section (%Block);>
<!ATTLIST pr:section
  type CDATA #IMPLIED
  id NMTOKEN #REQUIRED
  title CDATA #IMPLIED
  isTitleVisible CDATA #IMPLIED
>
<!ELEMENT pr:property (pr:keyword?, pr:description?)>
<!ATTLIST pr:property
  create_date CDATA #REQUIRED
  update_date CDATA #REQUIRED
  expired_date CDATA #IMPLIED
  title CDATA #REQUIRED
  isActive CDATA #IMPLIED
  isIndex CDATA #IMPLIED
>
<!ELEMENT pr:keyword (#PCDATA)>
<!ELEMENT pr:description (#PCDATA)>
```

content

- namespace: <http://www.atl-systems.co.jp/wdm/pr>
- id: 識別子
- creator_id: 作成者識別子
- version: ファイルバージョン

property

- namespace: <http://www.atl-systems.co.jp/wdm/pr>
- create_date: 作成日
- update_date: 更新日
- expired_date: 有効期限
- title: 文書タイトル
- isActive: 公開状態であるか否か
- isIndex: インデックス文書であるか否か
- keyword: 文書キーワード
- description: 文書補助説明

data

- namespace: <http://www.atl-systems.co.jp/wdm/pr>
- sectionを格納するためのコンテナ

section

- namespace: <http://www.atl-systems.co.jp/wdm/pr>
- id: localなスコープで扱われる識別子
- title: セクションタイトル
- isTitleVisible: セクションタイトルを表示するか否か
- type: セクションを扱うJavaClass名

%Block

- namespace: <http://www.w3.org/1999/xhtml>
- 既存のDTDを展開し、内容モデルの一部(%Block)をパラメータエンティティとして参照。

XHTML1.1のモジュール化機能を利用

特許申請中

資料:1 サイト構築比較 : 社員各々がデータ入力できる ... 大幅コスト削減

データ入力する人は、タグの意味を知る必要はない

構築後も専門家が必要

構築後は社員がデータ入力

ウェブサイト構築比較(従来システム)		コンペ提案作成←																				コスト										
スケジュール		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	人	月	合計								
仕様策定	サイトコンセプト																					1.00	0.25	0.25								
	サイト機能																					1.50	0.25	0.38								
	サイト構成																					1.50	1.00	1.50								
	アカウント設計																					1.00	0.25	0.25								
コンテンツ 素材収集・整理	DB設計																					1.00	1.00	1.00								
	コンテンツ作成計画																					1.50	0.25	0.38								
	行政情報																					1.50	1.00	1.50								
デザイン	観光情報																					1.50	1.00	1.50								
	トップページ																					1.50	1.25	1.88								
	行政情報ページ																					1.50	1.75	2.63								
	観光情報ページ																					1.50	1.75	2.63								
システム構築	共通ページ																					0.00	0.00	0.00								
	CSS/XSLT																					0.00	0.00	0.00								
コンテンツ 制作	CMS新規構築																					2.00	4.00	8.00								
	静的ページ																					3.00	1.25	3.75								
	動的																					1.00	0.50	0.50								
確認	コンテンツ																					1.00	0.50	0.50								
	管理情報登録																					2.00	1.50	3.00								
	リンクチェック・動作確認																					3.00	1.50	4.50								
	修正																					2.00	1.50	3.00								
《データ入力状況》		行政情報ペー	400	ページ	30分	200		時間															300	時間	43	日	8.6	週間	(5日)	2.1	人月	34.13
		観光情報ペー	200	ページ	30分	100		時間															300	時間	43	日	8.6	週間	(5日)	2.1	人月	34.13

ウェブサイト構築比較(WDM) (単位:時) デザインとデータ入力作業の分離による両作業のスピードアップ/並行作業(←表示と構造の分離)

- WYSIWYGタイプの操作簡単なGUIでデータ入力作業がスピードアップ...制作・運用コスト削減
- ウェブアプリケーションであるため、ブラウザがあればどこからでもデータ入力作業が可能...導入・運用コスト削減
- 同規模予算で、構築「作業」より、機能・コンテンツへの「知恵」や「足」が使える(←廉価パッケージ、使いやすさ)
- CMSとしてパッケージ化されているためわずかなカスタマイズで運用可能
- フォルダを利用者個々につくるためアカウント設計を入力作業前につくる必要があるが、その分ユーザ管理が明確になる
- コンテンツのカテゴリをを進めると、複数のカテゴリーに跨るページを共通のものとして分化・集約することによる省力化

ウェブサイト構築比較(WDM)		コンペ提案作成←																				コスト										
スケジュール		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	人	月	合計								
仕様策定	サイトコンセプト																					1.00	0.25	0.25								
	サイト機能																					1.00	0.25	0.25								
	サイト構成																					1.00	0.75	0.75								
	アカウント設計																					1.00	0.50	0.50								
コンテンツ 素材収集・整理	DB設計																					0.00	0.00	0.00								
	コンテンツ作成計画																					1.00	0.25	0.25								
	行政情報																					1.50	0.75	1.13								
デザイン	観光情報																					1.50	0.75	1.13								
	トップページ																					1.00	0.75	0.75								
	行政情報ページ																					1.00	0.75	0.75								
	観光情報ページ																					1.00	0.75	0.75								
システム構築	共通ページ																					1.00	0.75	0.75								
	CSS/XSLT																					1.00	0.75	0.75								
コンテンツ 制作	CMS/システムカスタマイズ																					1.50	0.50	0.75								
	静的ページ																					1.50	0.50	0.75								
	動的																					0.50	0.25	0.13								
確認	コンテンツ																					0.50	0.25	0.13								
	管理情報登録																					2.00	0.50	1.00								
	リンクチェック・動作確認																					1.00	0.50	0.50								
	修正																					2.00	0.50	1.00								
《データ入力状況》		行政情報ペー	100	ページ	5分	8.3		時間															50	時間	7.1	日	1.4	週間	(5日)	0.4	人月	17%
		観光情報ペー	200	ページ	5分	17		時間															50	時間	7.1	日	1.4	週間	(5日)	0.4	人月	17%
		共通ページ	300	ページ	5分	25		時間															50	時間	7.1	日	1.4	週間	(5日)	0.4	人月	17%

総制作時間 67%短縮
ファイル制作 83%短縮

●制作期間:約1/3に短縮
●制作技術者の分業:デザイナー/データ入力者

南アルプス市ネットワーク保守業務						
LOGIN/OUT	HOME	新規受付	プレビュー	DATA		
報告書(表示)	受付内容編集	作業予定編集	作業指示編集	作業内容編集	今後の予定編集	管理情報編集

必要箇所をPDF出力することも可能

受付		編集
受付日	2003年11月05日(水曜日) 09時30分	
依頼元	松廣(富士通ゼネラル)	
受付者	清水 章弘 / 株式会社 エーティーエルシステムズ	
受付	電話	
受付内容分類	ネットワーク機器	
受付内容		
<p>富士通ゼネラル納品システムである救急統計システムにおいて甲西分遣所のみ11月1日(土)より利用が出来なくなっているとの問い合わせ。</p> <p>(救急統計システム:救急の出動情報をDBに入力し報告書出力するシステムで、情報に高度な一般個人情報が含まれるため消防署間の独立した専用LANとして専用端末で利用されています。このため監視対象外としています)</p>		

ネットワーク保守業務						
HOME	新規受付	プレビュー	DATA			
受付内容編集	作業予定編集	作業指示編集	作業内容編集	今後の予定編集	管理情報編集	
受付内容						
<p>ブロックタイトル 受付内容 表示 非表示 検証・更新</p> <p>選択してください 数字</p> <p>パス: body > p</p> <p>富士通ゼネラル納品システムである救急統計システムにおいて甲西分遣所のみ11月1日(土)より利用が出来ない状態になっているとの問い合わせ。</p> <p>(救急統計システム:救急の出動情報をDBに入力し報告書出力するシステムで、情報に高度な一般個人情報が含まれるため消防署間の独立した専用LANとして専用端末で利用されています。このため監視対象外としています)</p>						

報告書名	南アルプス市ネットワーク保守業務報告書
顧客名	南アルプス市
事業名	南アルプス市ネットワーク保守業務
業務区分	保守管理
管理番号	
受付日	2003年11月05日(水曜日)09時30分
依頼元 部署/指名	富士通ゼネラル / 松廣
受付内容分類	ネットワーク機器
受付内容	<p>富士通ゼネラル納品システムである救急統計システムにおいて甲西分遣所のみ11月1日(土)より利用が出来ない状態になっているとの問い合わせ。</p> <p>(救急統計システム:救急の出動情報をDBに入力し報告書出力するシステムで、情報に高度な一般個人情報が含まれるため消防署間の独立した専用LANとして専用端末で利用されています。このため監視対象外としています)</p>
開始時間	2003年11月04日(火曜日)11時00分
終了時間	2003年11月04日(火曜日)16時45分
作業者	清水章弘 / 株式会社 エーティーエルシステムズ
確認者	南アルプス市役所情報政策課 /

管理タグを保守管理用に、独自XMLを定義した。
ISO9000等の資料に利用
マニュアル作成

◆ 視覚重視 WYSIWYG エディタ方式

- タグを一切意識せず、文書構造のみのデータを作成

◆ Webブラウザによる入力方式により何処からでも入力可能

- 500Kbpsの回線があれば何処からでも入力可能

◆ データ作成時に構文チェック(バリデート)

- コンピュータ処理が可能な正規化されたデータ作成

◆ XMLデータによるワンソースマルチユースの実現

- CSSによりユーザに合わせた多様な表現
- Web、印刷PDF、BML、携帯電話等へXSLTを用いてデータ変換

◆ サーバによるコンテンツデータの管理

- 集中管理によるセキュリティ保証、原本性の保証

◆ CMS機能による役割・権限管理

- データ作成者の個人特定、アクセス履歴管理…… すべてを管理可能

マーケット動向

- ① 文書コンテンツをコンピュータ処理
- ② ワンソース/マルチユースのニーズ

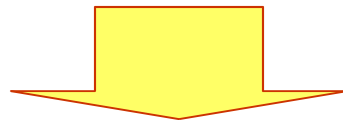
1. 事務仕事の60%は文書作成関係

- 文書を「探す」
- 文書を「作成する」

紙に換えて、Excel, Word
便利な分だけ問題も発生

2. 簡単に構造化文書が作成できなかった

- システムコストが高い。運用、メンテナンスコストが高い。
- MS-Office製品では、構造化はできない。
- 構造と表示情報が分離されていない。



現状：帳票作成システム
今後：専用システム化

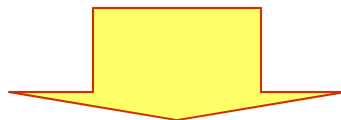
新規市場 : コンテンツのシステム管理化

1. 従来のSGMLで管理していた分野

- 世界規模でのマニュアル管理
- 医薬品の副作用情報など、即時性と一元管理が必要であった分野

2. ワンソース マルチユース (インターネット環境)

- デバイスの多様性。「紙はデバイスの一つ 携帯、BML、他・・・」
- 文書 = 紙 からの決別 「メールを印刷・安心できればしない」
- 作業の効率化の要求 「検索エンジン、コンテンツの再利用」
- 情報開示の重要性 「開示することが組織の自己防衛」
- 一元管理による原本性の保証 「コピー禁止・デジタル化の影」



ネットワークを前提にした情報共有とセキュリティ