



From SGML to XML

Case Study of A Product Service Guide Development

Chung-Ping Liang

CALS Program Director
Institute for Information Industry
Taiwan, ROC



Discussion Outline

- ◆ Introduction
- ◆ Current Service Guide Development Situation in Taiwan PC Industry
- ◆ SGML, XML and HTML
- ◆ The SGML Pilot Project
 - Document Analysis & DTD Production
 - System Configuration
- ◆ XML for Document Publishing on the Web
- ◆ Conclusion



Introduction

CALS 的意義：

CALS/資訊運籌管理，係指資訊整合與交換；藉由作業程序的改造及資訊技術與標準的運用，建立全球共通商業系統，以無紙化作業環境，將業務上所必要的資訊加以電子化、標準化，並運用資料庫與網路系統，使得所有資訊得以達到交換與共用為目標之策略。

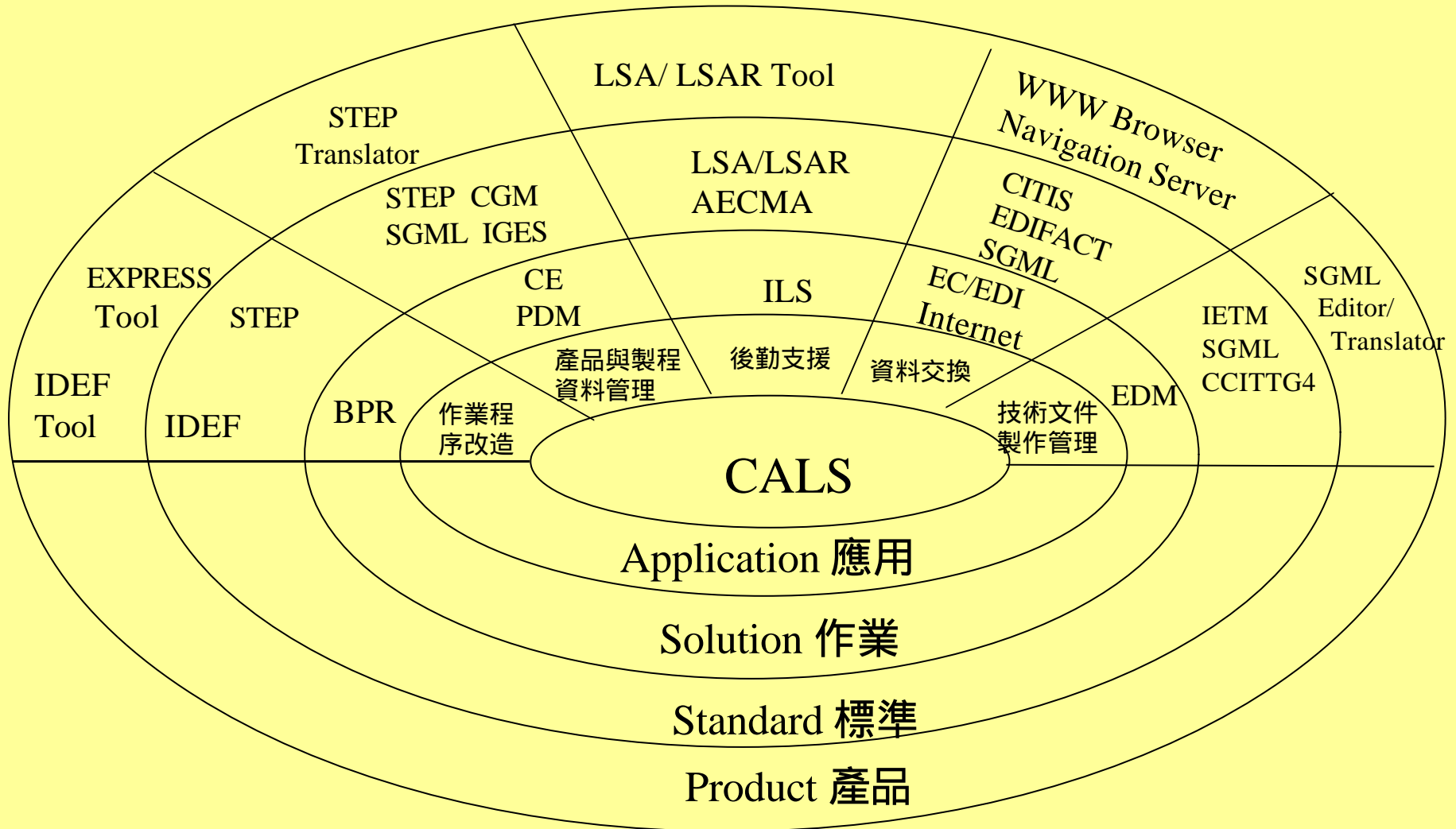
CALS最主要的精神在於：

Data Create Once, Use Many Times.



Introduction (cont.)

CALS 之應用, 技術, 標準及資訊產品關係





Introduction (cont.)

- ◆ MOEA of Taiwan initiated a CALS pilot program and intended to develop CALS environment to maintain Taiwan competitive edge in today's market place
- ◆ PC industry was chosen be as a pilot industry to implement CALS applications on their Product Data Management and service Guide production to enhance their logistics support capability
- ◆ In comply with CALS standard, an SGML-Based Document Management system was developed on PC Service Guide Production



Current PC Service Guide Development Situation

Current Situation

. DC

- MRS
- Model no. definition
- EES
- Schematics

. EE

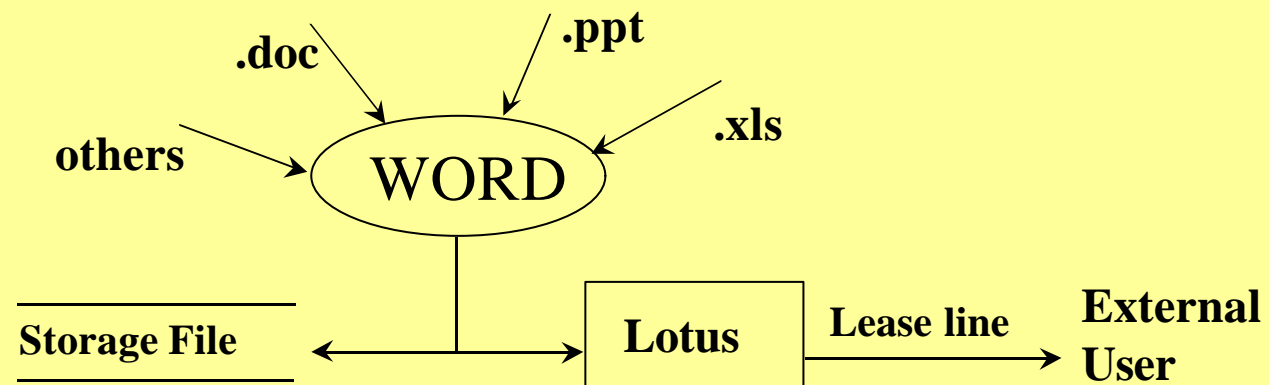
- Schematics
- major chips data sheets
- Jumper settings
- Block diagrams

. ME

- Disassembly procedure presentation
- Exploded diagram

. SW

- POST check point list
- BIOS error message list
- Spare parts list





SGML, XML and HTML

- ◆ SGML Environment
- ◆ SGML/XML Families
- ◆ SGML, XML vs. HTML



SGML Environment

產品分類：

- (1) 資料管理工具
- (2) SGML Converter
- (3) SGML Editor
- (4) SGML Parser
- (5) DTD Editor
- (6) SGML Publisher
- (7) SGML Web Browser



SGML/XML Families

◆ SGML Family

SGML

HyTime

DSSSL

◆ XML Family

XML(Extensible Markup Language)

XLL(Extensible Linking Language)

XSL(Extensible Stylesheet Language)



SGML, XML vs. HTML

- ◆ SGML is the Standard Generalized Markup Language
- ◆ XML is the subset of SGML and is designed to make it easy to use SGML on the Web
- ◆ HTML is the HyperText Markup Language, a specific application of SGML used in the World Wide Web
- ◆ HTML is not flexible enough to define tags
- ◆ XML has strong hypertext linking capability



PC Industry SGML-Based Document Management System Pilot Project

◆ Objective:

By establishing the SGML-Based document management environment within enterprises to reduce the development time and to solve version control and data consistency problems

◆ Approach:

- Document Analysis
- Produce DTD
- Define System H/W & S/W requirement
- Output Specification for chosen system
- Develop this SGML-Based technical document management system



Document Analysis and DTD Production

◆ Document Analysis :

1. Find all possible elements and their attributes
2. Find the document structure
3. Find the relationship between the elements and the document structure



Document Analysis and DTD Production(cont.)

◆ Document Analysis Example : (Product Service Guide)

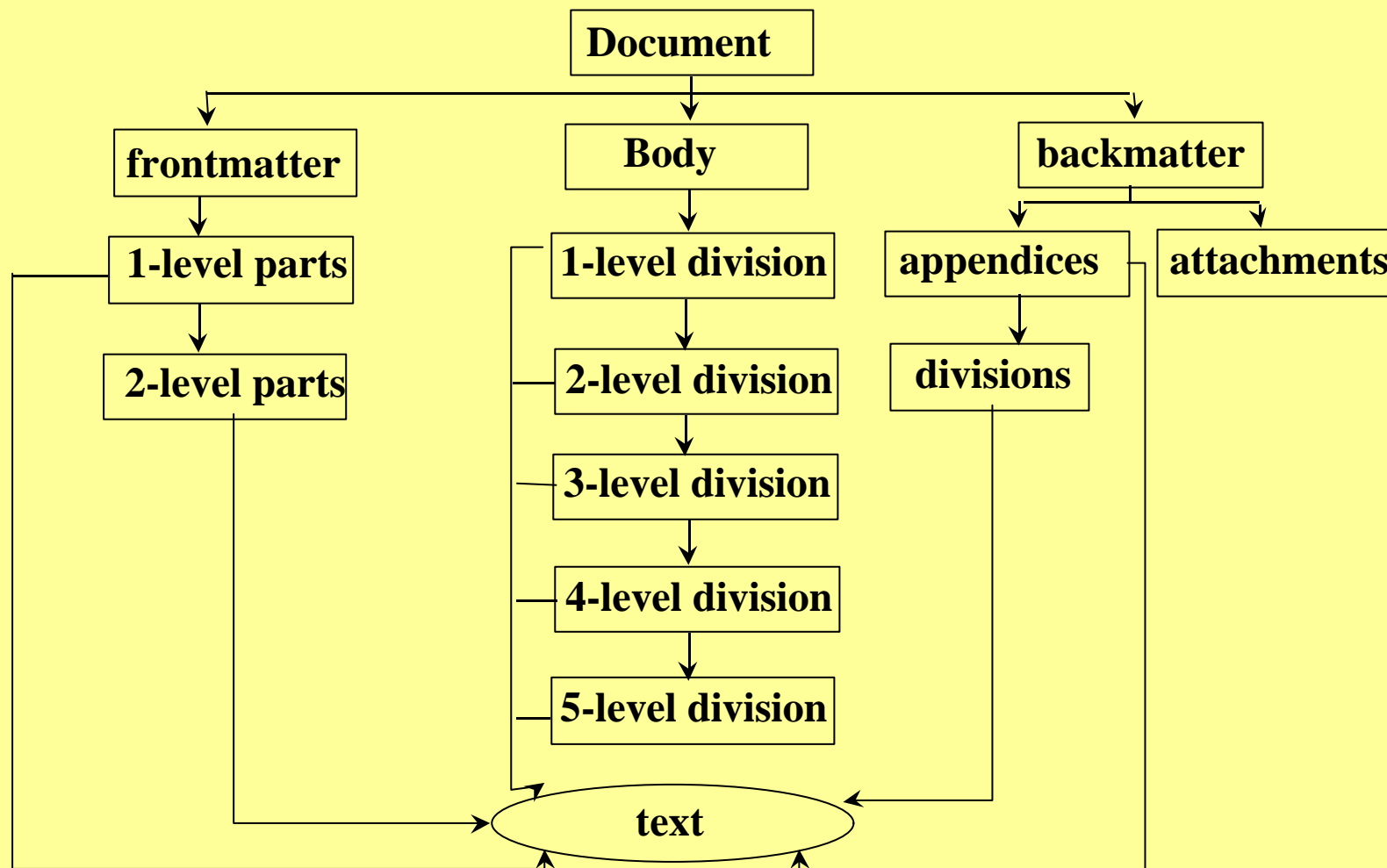
1. Find all possible elements and their attributes

model name, document name, recycle mark,
part number,document number, copyright, disclaimer,
trademarks list,document description, table of contents,
figure list, table list,division, paragraph, lists, dialog box,
PC screen, PC button, notes macro, caution macro,
warning macro ...



Document Analysis and DTD Production(cont.)

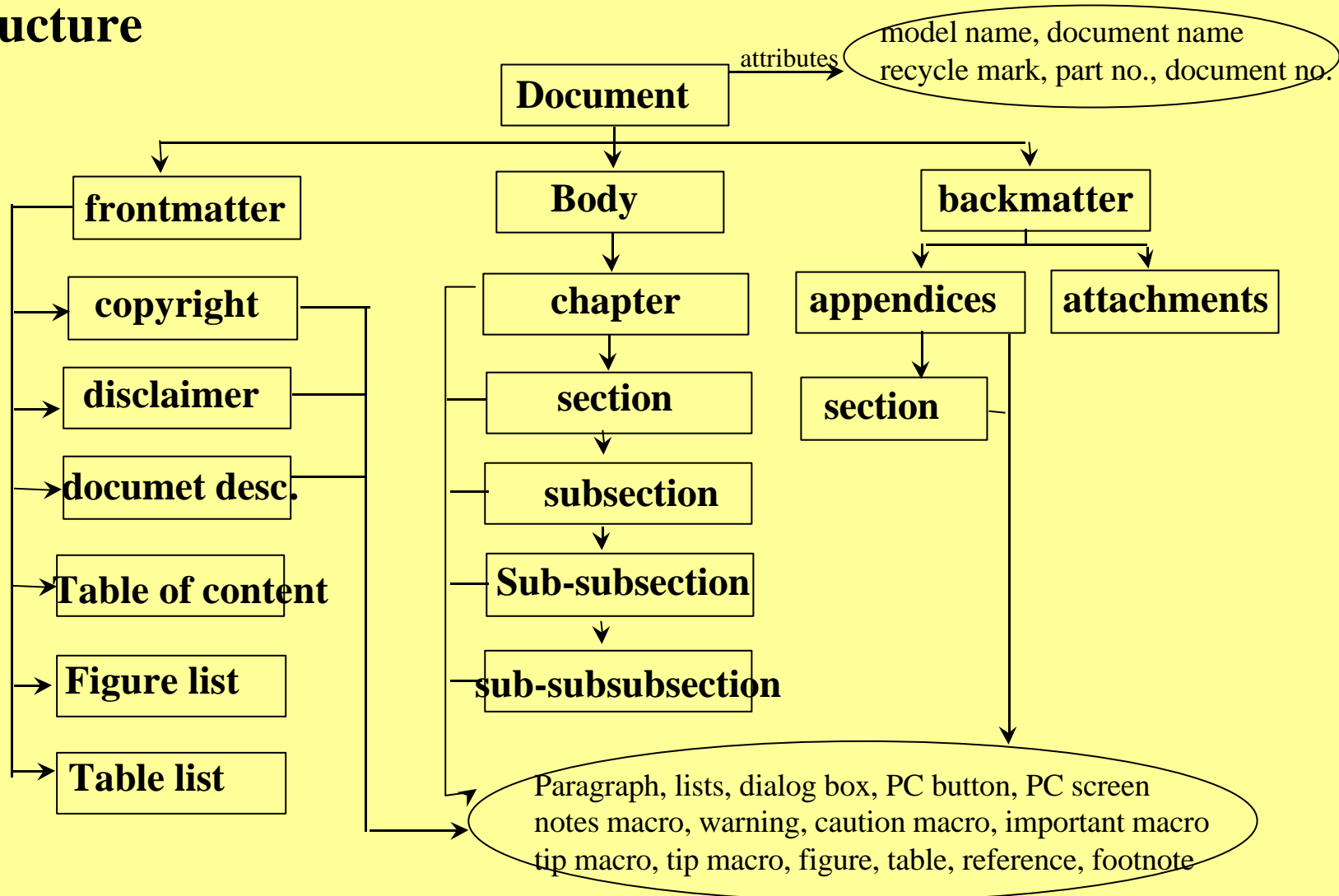
2. Find the document structure





Document Analysis and DTD Production(cont.)

3. Find the relationship between the elements and the document structure





Document Analysis and DTD Production (cont.)

◆ DTD Production:

1. Select from Modular/Book(contiguous) DTD
2. Define all elements with the attributes
3. Build the DTD structure



Document Analysis and DTD Production (cont.)

◆ DTD Production Result : (Product Service Guide)

```
<!ENTITY % Float    “Figure | Table | FootNote” >
<!ENTITY % Title    “PgTitle | DpTitle | ChTitle | ScTitle | AxTitle ” >
<!ENTITY % Elm      “Artwork | Button | DirButton | DgBox” >
<!ENTITY % Copy     “Para | Info | %Elm;” >
<!ENTITY % List     “Ls | TopLs | DefLs ” >
<!ENTITY % Text     “%Copy; | %List; ” >
<!ENTITY % Emph     “Roman | CourierNew | Italics | Bold | BoldItalics | Underline ” >
<!ENTITY % Ref      “TiRef | FigRef | TbRef | FnRef | ItRef ” >
```

```
<!ELEMENT ACERDoc - O (Front?, Body, Back?) + (%Float; |%Ref; | IxEntry)>
<!ELEMENT Front - O (Prolog, TbCnt, (FigList? & TbList?))>
<!ELEMENT Body - O (Chapter)+>
<!ELEMENT Back - O (Appendix*, (Others? & Index?))>
<!ELEMENT Prolog - O (Copyright?, Disclaimer?, Trademark?, DocDesc?)>
<!ELEMENT (Copyright | Disclaimer | Trademark) - O (PgTitle?, (%Text;)*>
```

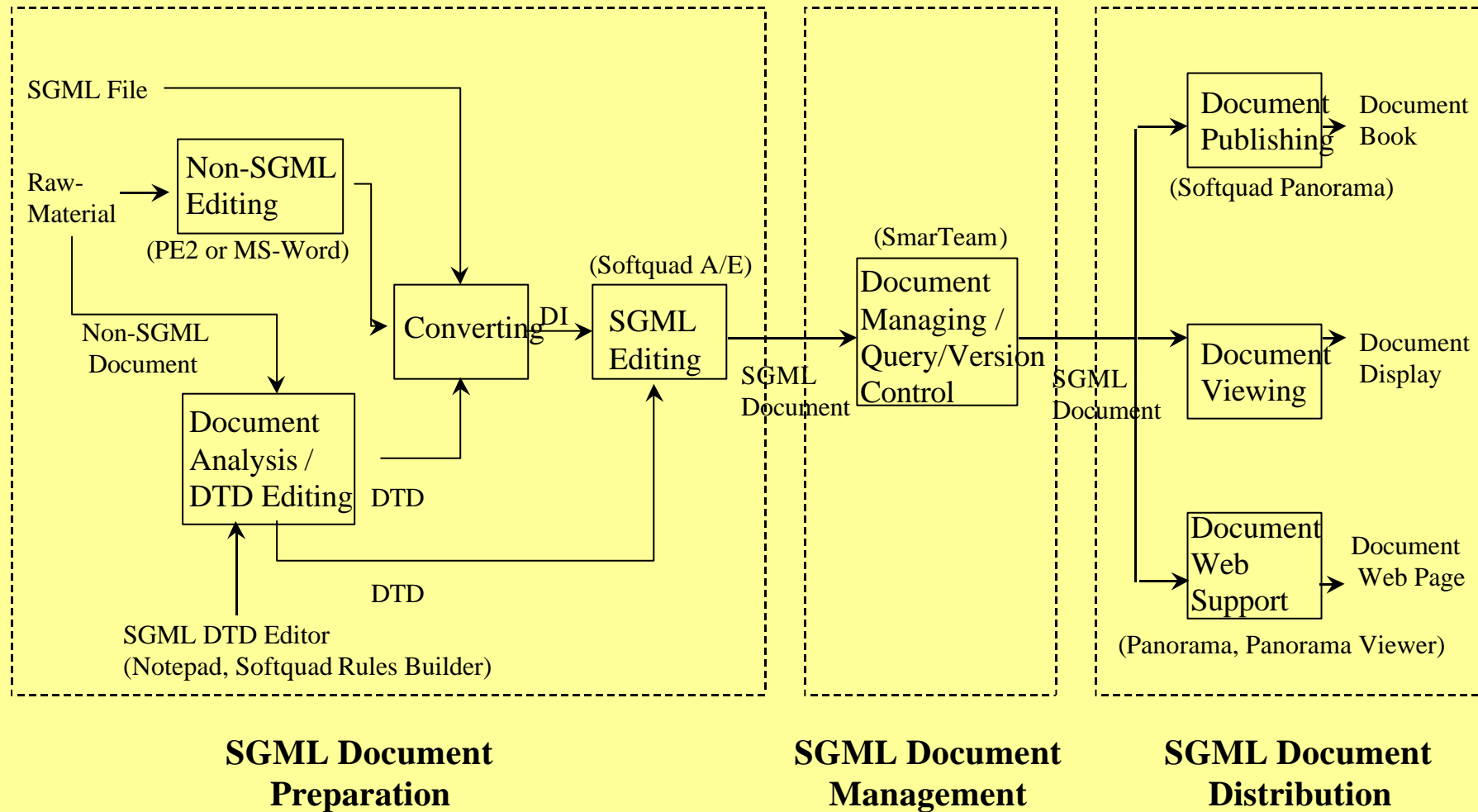
:

```
<!ATTRLIST (Chapter | Appendix)>
      Id      ID      #IMPLIED >
<!ATTRLIST (ChDesc | ApxDesc)
      NewPage      (yes | no)      yes >
```

:

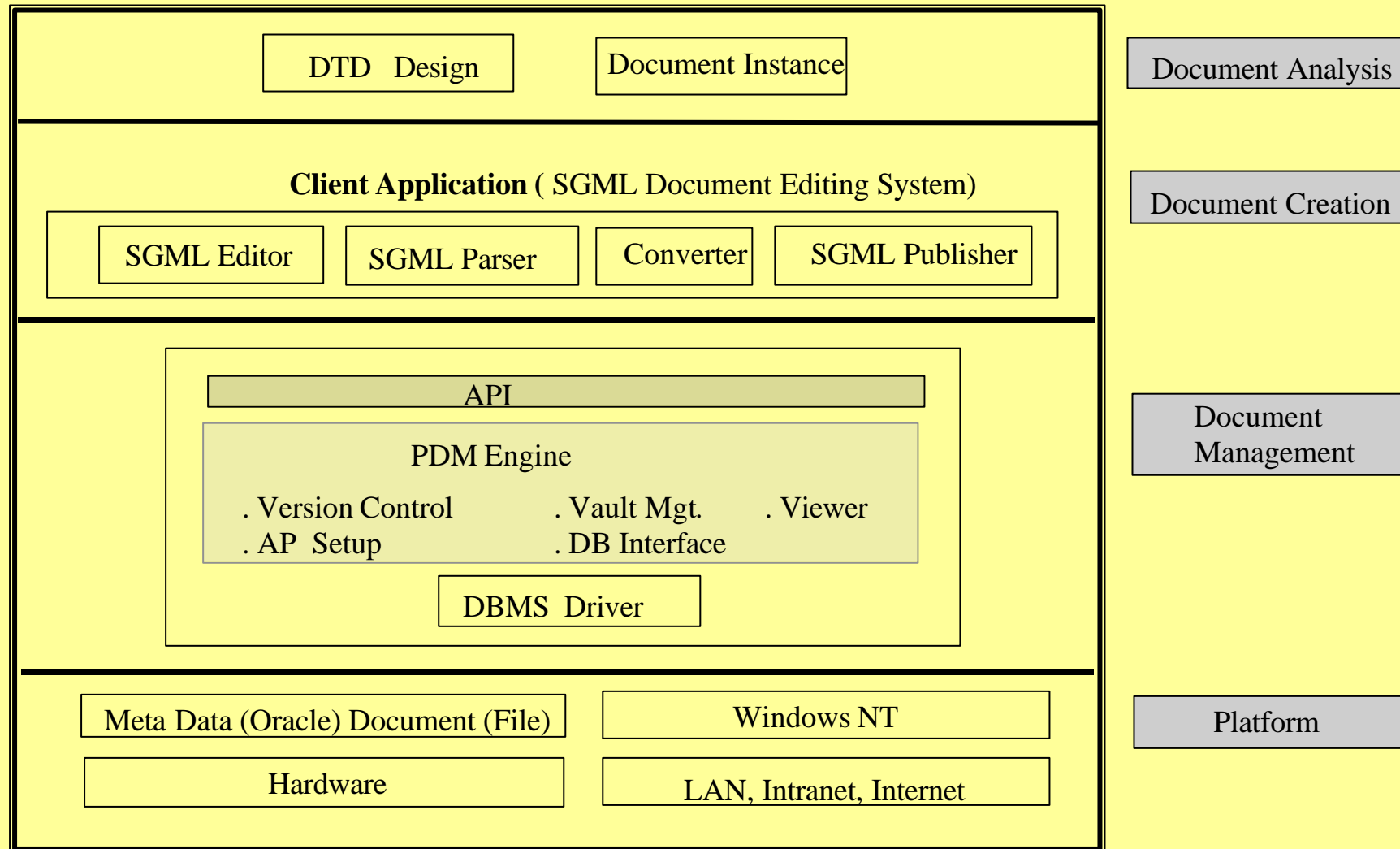


SGML-Based Document Production System Configuration





SGML-Based Document Management Structure





XML for Document Publishing on the Web

◆ Why XML ?

1. Public information (such as change notices) need to be broadcast to the whole enterprise or even to the retail dealer, in this case, the data/document should be delivered in some manner which can meet those receivers' needs (eg. for different style output).
2. To improve the data/document reusability can reduce the business management cost.
3. Web publishing is the trend and it can meet our needs.



◆ XML Strength

1. Address problems of SGML
2. For publishing documents on the Web
3. Simple
4. Easy to implement

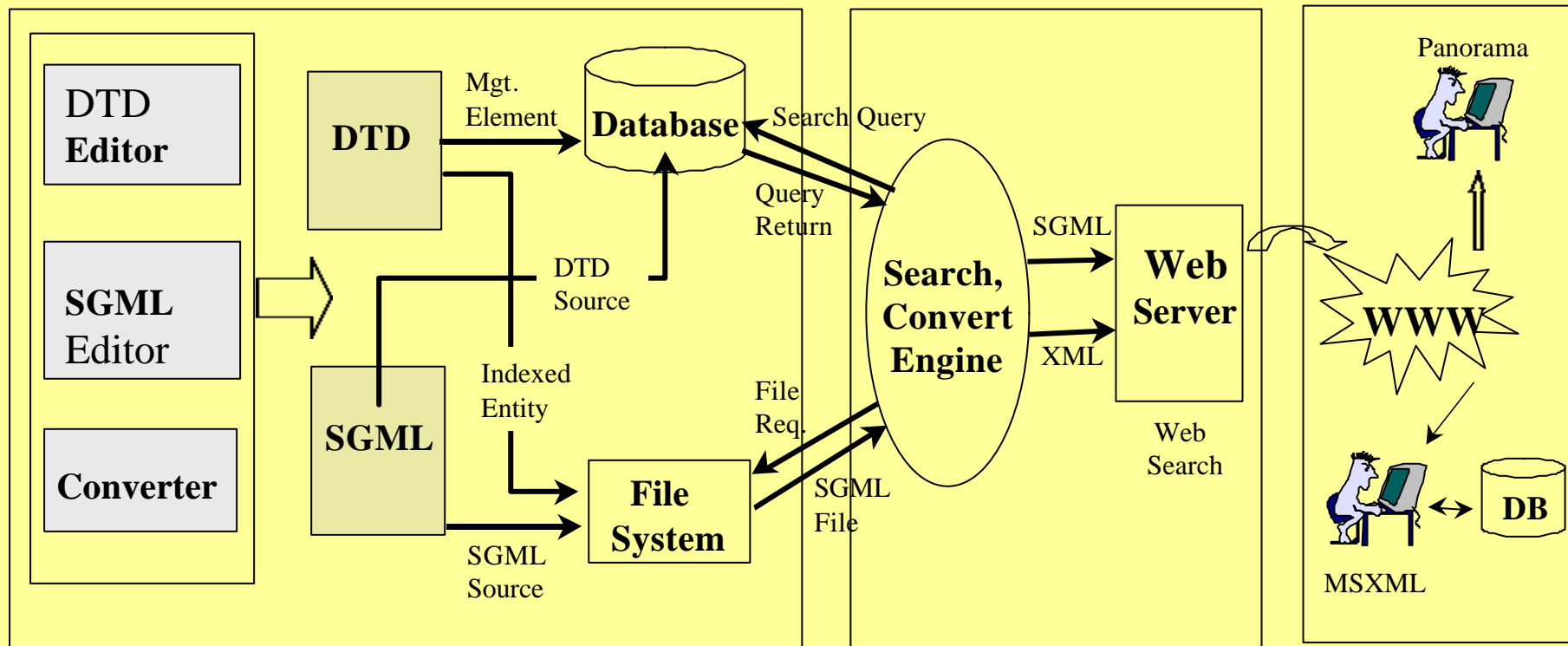


XML for Document Publishing on the Web (cont.)

SGML Data / DTD Preparation/Management

SGML Document Processing

XML Document Distribution





Lessons Learned

1. To decide the granularity of the DTD elements not only depends on the result of the document analysis, the component itself to be managed is another important concerns.
2. It takes great effort to customize a data management repository without SGML features to deal with SGML format data.
3. The enterprise administration belief is the driven force to transfer the document management from conventional method to SGML/XML approach.



Conclusion

- ◆ XML is actually compatible with SGML -- XML documents can be read by any SGML authoring or viewing tool.
- ◆ SGML will probably never gain widespread acceptance on the Internet
- ◆ For high-end ,highly structured publishing applications, SGML will continue to fit the bill.