

XML技術の現状と 今後の可能性について

土屋俊
(千葉大学)

まとめ

- XMLとは、インターネット時代の構造化文書技術である
- XMLは、企業を含むさまざまな組織における文書管理を格段に向上させる(One source, multiple use) 環境を変えることが大事
- XMLは、すでにパソコンレベルの技術となった
- 日本での普及が遅れているのはなぜか？

XMLとは？

- 構造化文書の(ひとつの)歴史
 - IBMの社内文書(GML)
 - 国際規格へ(SGML、ISO 8879 - 86)
 - ある種の停滞とひそかな発展
 - ビジネスとして展開しかなかった(抽象的かつ厳格)
 - 大量文書保存技術としての展開
 - 印刷会社
 - 電子図書館(TEI)
 - 電子ジャーナル
 - 電子資源管理(メタデータ)

XMLとは？

- インターネット文書の(ひとつの)歴史
 - ハイパーテキストの実現
 - HTML(1989)
 - MOSAIC(1993)
 - 画期的なのは、
 - インターネット上の文書間の相互参照(これは、SGMLでは簡単にはできなかった)
 - SGMLアプリケーションとしての見直し

XMLとは？

- 要するに

SGML + HTML

- - - - - = XML

2

構造化文書 + リンク (+ プレゼンテーション)

XMLとは(goals) ?

- XML shall be straightforwardly usable over the Internet.
- XML shall support a wide variety of applications.
- XML shall be compatible with SGML.
- It shall be easy to write programs which process XML documents. (XLST)
- The number of optional features in XML is to be kept to the absolute minimum, ideally zero.
- XML documents should be human-legible and reasonably clear.
- The XML design should be prepared quickly.
- The design of XML shall be formal and concise.
- XML documents shall be easy to create.
- Terseness in XML markup is of minimal importance.

何ができるのか

- アプリケーションに依存しないデータ共有
 - ワープロ
 - スプレッドシート
 - データベース
- プラットフォームに依存しないデータ共有
 - パソコン
 - ウェブ
 - サーバ
- 多言語対応はUnicodeで (現在はUTF-8)

MSOffice 2003で実装

- WordファイルをXML形式で保存する
 - テキストファイル 16バイト
 - Wordファイル 24000バイト
 - XMLファイル 5000バイト
- ExcelファイルをXML形式で保存する
 - テキストファイル 18バイト
 - XLSファイル 14000バイト
 - XMLファイル 5000バイト
- PowerPointはまだ
- InfoPathを導入

データ共有

- ExcelからWordへ
 - データの共有
- Wordからウェブページ
- AccessからSQL
- ExcelからExcelへ (XLST)

- <http://www.w3.org/XML/>