

Webを支える情報通信技術(2)

Webでのデータの表現形式

今回の目標

- ブラウザは今回どのくらい役に立つだろうか?
 - データ表現の標準が決められていることが、どのように役立っているかを知ろう
- HTMLを見直そう
 - 性格を知ろう
 - 「規格」の読み方を知ろう
 - 人間が書くのではなく、プログラムが生成する場合があることを知ろう。
- XMLとは何かを知ろう
 - 人間に読みやすいことが別に関係ない世界があることを知ろう
 - 具体的な応用を見よう

ブラウザ

- ブラウザのコアな役割は何か？
- 指定されたURLからのデータの取得
+
- 取得したデータの「解釈」

ブラウザ

- 1 . まずブラウザがある
- 2 . ブラウザが「アクセス」
- 3 . 「アクセス」先からデータを取得
- 4 . 取得したデータを「解釈」

「解釈」すべきもの

- ブラウザにとって、「解釈」とは何か?
- ドキュメントの論理的構造の記述
- ドキュメントのレイアウト構造の記述
- 記述言語 - HTML

HTML

- 表示されたページのソースを出せますか?
- ソースを読めますか? (なんとなく or 細部まで詰めて or ...)

もともとHTMLとはなにか

- 1.0 → 2.0 (RFC1866) → 3.2 → 4.0
- 初期は、ブラウザをリリースするところが独自に「規格」を拡張
- SGMLの変種
 - SGML (Standard Generalized Markup Language)

HTML

- ドキュメントの論理構造と、レイアウト構造を記述する
- 現在、レイアウト構造の記述部分が肥大
- 規格の読み方知ってる？

HTML規格を読んでみる

- HTMLの理解の仕方
 - 実地に訓練を積む
 - 規格を読む
- 4.0なら
<http://www.w3.org/TR/html4/>
- 3.2なら
<http://www.w3.org/TR/REC-html32>
- 2.0ならRFC1866

HTML規格を読んでみる

- RFCって何?
- W3Cって何?
- インターネットの中でauthorized documentを探すにはどうしたらいいだろうか?
 - If 書籍の世界なら...
 - If 大学の中なら...

HTML規格を読んでみる

- DTDって読める？
- 規格を読むときにどこから攻めたらいいかわかる？
 - 具体例をそばにおく
 - 「規格」を守っていないものでもたいていのものは正しく表示されるって知ってた？
 - ブラウザは実は優しい
 - ばりばり読む

<http://www.ipsj.or.jp/sig/pro/>

- なんだ、のっけから標準形になっていないじゃないか

HTML規格を読んでみる

- ページをひとつ開いてみよう
- 規格を片手にばりばり読んでみよう
 - <http://www.w3.org/TR/html4/>

HTMLの生成

- 以下をみたすページを作ろう
 - 19 × 19の段までの九九の表
 - table.c, table.xsl
- ページを生成するために
 - 自分でHTMLを書く
 - プログラムに書かせる = オーサリングツール

HTMLの生成

- 実は、HTMLを生成するだけではなく、
- 生成されたHTMLから「ドキュメントの構造」を理解することが、「解釈」に必要
- これをパースという。

HTMLの生成

- HTMLを生成 / パースするためには、規格の理解が不可欠
- HTMLをプログラムが生成するということは、
- HTMLをプログラムが読む場合だってある

XML

- **プログラムが読むためのデータ標準**
 - レイアウトに関する構造の記述いらない
- **例がいろいろ**
 - RSS
 - RDF

ブラウザ、作れる？

- 必要な技術を列挙しよう(ソフトウェアを作れとはとりあえずいわない)
 - あれは？
 - これは？
 - それは？

今回ふれていないこと

- URLの意味は何?
- 「アクセス」の正体は何?
- 次回以降、観察しましょう。