

- ・ SPARQL: セマンティック・ウェブと Web 2.0 が出会うところ
- ・ セマンティック・ウェブ、あるいはルーズさを生かした構造
- ・ Web 2.0 の目指すところは結局 Semantic Web + Web Services では

top down & bottom up

- ・ W3C 流 : top down 気味
- ・ その他 : bottom up
- ・ bottom up な技術群
 - ・ RSS, ATOM による metadata description/content delivery
 - ・ ATOM Publishing Protocol
 - ・ REST/SOAP

関連記事

- ・ <http://www.yamdas.org/column/technique/history.html>

雑感

人間ではなく「コンピュータが扱える Web」を目指すものが Semantic Web だろう。

Web コンテンツが増えるにつれ、人間がブラウズするだけではなくコンピュータによる加工が必須になりつつある現在、この目標を掲げることは自然なことだ。

現在広く使われている検索エンジンの技術の基礎は Semantic Web を前提としない。すなわち HTML をパースしてあの手この手で情報を抽象化するという非常に泥臭い作業を行っている。検索エンジンが広まるにつれ、コンテンツ提供者側に、簡単なメタデータを提供させることが可能になってきた。これは提供者側が機械可読なメタデータ情報を付与するという、従来から見れば余計な作業を担うことが、十分な利益として返ってくるようになってきたことが大きな理由だ。その理由というものはもちろん検索エンジンに見つけられやすくなる、ということである。

Semantic Web のような考えを推進するときの一番大きな問題はメタデータ付与に必要な労力である。はっきりとした利益が得られない限り、コンテンツ提供者が進んでメタデータを作成することはない。

このメタデータ問題を解決するための方法は、メタデータの自動生成と、メタデータ作成支援ツールの提供、メタデータ作成によるインセンティブの付与などが考えられる。

blog や wiki などはメタデータを自動生成することによって、コンテンツ作成者にメタデータ作成という作業を意識させないようにしている。

メタデータの有用性をはっきりと分らせてくれたものは RSS, ATOM などのフィード技術だろう。現在はフィードされてきた情報を RSS リーダなどを用いて直接人間が読む使い方が主流だが、いずれフィード情報のメインユーザは人間からコンピュータになるだろう。フィード情報を加工して提供するという使い方が主流になると考えられる。

はてなの Semantic Web の説明にあるように Semantic Web には懐疑的な意見が多かった。これは W3C で行われている活動内容がトップダウン的、抽象的、学術的、完璧主義的すぎてこれはとてもではないが実現は無理だ、という印象を与えたからではないかという気がする。

一方、最近では REST, RSS, ATOM 等の技術（もしくは考え方）が普及してきており、Amazon, Google, Yahoo, Flickr などの大手コンテンツ提供者が Web Services API を提供するようになることによって、Semantic Web の目指すところが十分見えるようになってきた印象がある。

トップダウンおよびボトムアップの複数の流れで発展してきた技術が統合されることによって、ようやく Semantic Web が目指すものの形が見えてきたのかもしれない。

misc

- ・ オントロジ
- ・ Dublin Core
- ・ RDF