

セマンティックWEBとソ シアルインフォメーション フィルタリング

政策・メディア研究科 助教授 北川和裕
政策・メディア研究科 助手 吉田絵里香

本報告のアウトライン

前半

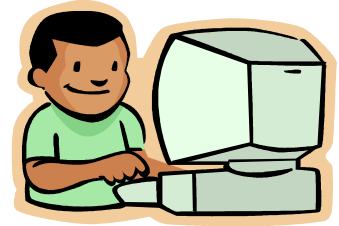
- セマンティックwebはレーティングメタデータを提供
- インフォメーションフィルタリングの必要性

後半

- フィルタリングの技術
- W3C／慶應の活動と今後

コンピュータは人間の言葉を理解しない

- WEB上の情報は
人が読む

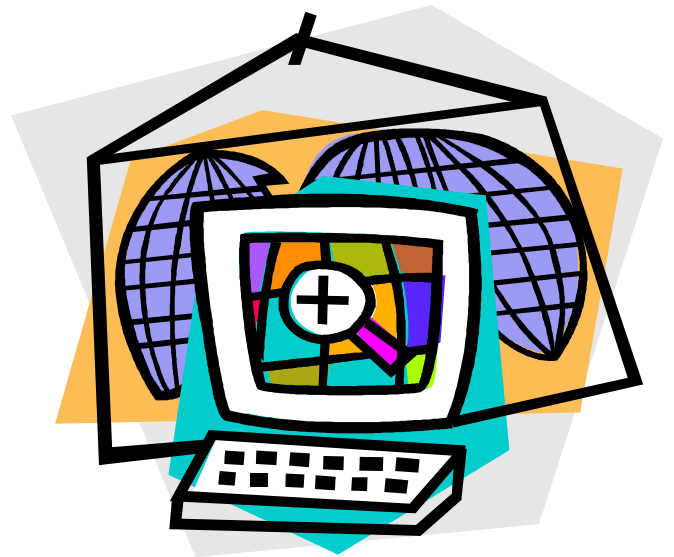


- WEB上の情報検索は
コンピュータ(機械)が行う



セマンティックWebとは

「WEB上の情報を機械的に
処理するための枠組み」



セマンティックWEBとRDF

RDF (Resource Description Framework) とは、

「WEBリソースのためのメタデータ記述形式」のひとつ。



- 機械のための意味情報を、
- 人間が読むHTMLとは別に、
- 機械処理可能なデータ形式で用意

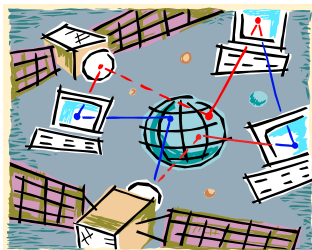
WEBの思想とベネフィット

WEBの思想

- WEBはシンプル

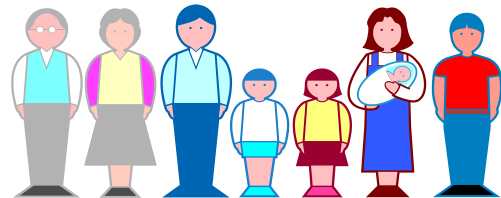


- WEBはオープンシステム



ベネフィット

- 誰でも利用できる



- 情報を共有できる



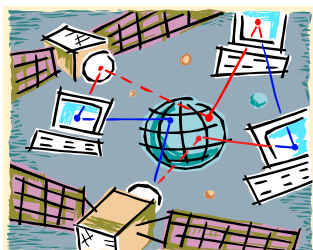
WEBにまつわる社会問題

WEBはシンプル



子供でも簡単に利用できる

WEBはオープンシステム



有害情報や



情報の信頼性の問題をどうするのか？

統計データにみる子供とインターネット



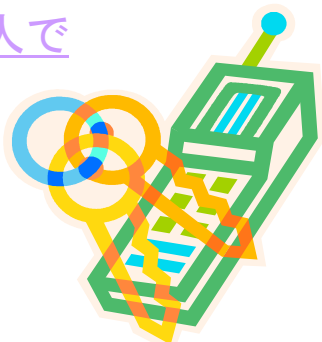
- 勉強時間の減少



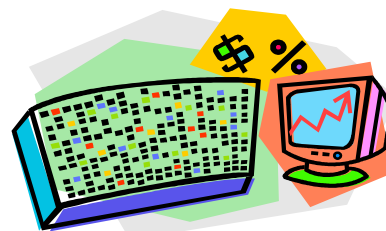
- 親の目の届かないところで



- 子供が1人で



- 親より高いリテラシ





メタデータを活用した ソーシャルインフォメーションフィルタリング

WEB技術を社会において

- 問題なく、
- 快適に利用するために



メタデータの活用例

- PICS
- CC/PP
- DCMES
- RSS



セマンティックWebとソーシャルインフォー メーションフィルタリング

W3C Device Independence
アクティビティリード
慶應義塾大学

北川 和裕

kaz@w3.org

インフォメーションフィルタリング

- 見たいコンテンツにより簡単に到達
- 見たくないコンテンツを見せない
 - レコメンダーシステム
 - サーチエンジンのパーソナライズ
- 代表的なフィルタリング技術
 - コンテントベースフィルタリング
 - ソシアルフィルタリング

W3C／慶應のフィルタリング活動

- 携帯電話を主な対象としたフィルタリング
 - 主な目的: 技術による出会い系サイトの制限
 - 総務省のサポート: インターネット協会、キャリアー、フィルタリング事業者と協業
 - 誰もがWebサイトのラベリングを目指す — フォーマットの共通化
 - 標準化のトラックに至らず
 - EUのQuattroプロジェクトと協業: W3C/ERCIMが参加

現在の状況と今後

- RDF形式の基本フォーマットの決定
 - セマンティックWebで利用されるメタデータの実体の一つ
 - Quattroと基本フレームワークを共通化
- 要求仕様書の出版
- 仕様書の公開
- ユースケースの出版
- アーキテクチャの提案
- W3Cへの提案

フォーマット仕様の公開によって

- 誰でもがWebコンテンツのレーティングが可能
 - 検閲ではなくコンテンツプロバイダー、利用者自身が様々なコンテンツの評価が可能
- 評価データが交換可能
 - フィルタリングプログラムとデータを分離し利用者の選択を広げる
 - データを作成、配布する事業者
- Web空間に多くのレーティングデータが存在
 - ソシアルインフォメーションフィルタリングの実現

W3C／慶應での今後の活動

- ワークショップの開催
- 標準化のための活動の開始を推進
- 慶應大学として
 - ブロッキング以外の道を同時に探る
 - 優良サイトを優先的に検索と同値
 - 精度の高いソーシャルフィルタリングアルゴリズムの開発
 - フィルタリングソフトウェアのサンプル実装
 - ISWC 2004において発表

まとめ

- セマンティックWebの技術を用いてソーシャルフィルタリング技術を開発
 - 応用: 出会い系サイトのブロック
- W3Cの標準への向けての活動
- 連絡先
 - 吉田: erika@sfc.keio.ac.jp
 - 北川: kaz@sfc.keio.ac.jp