

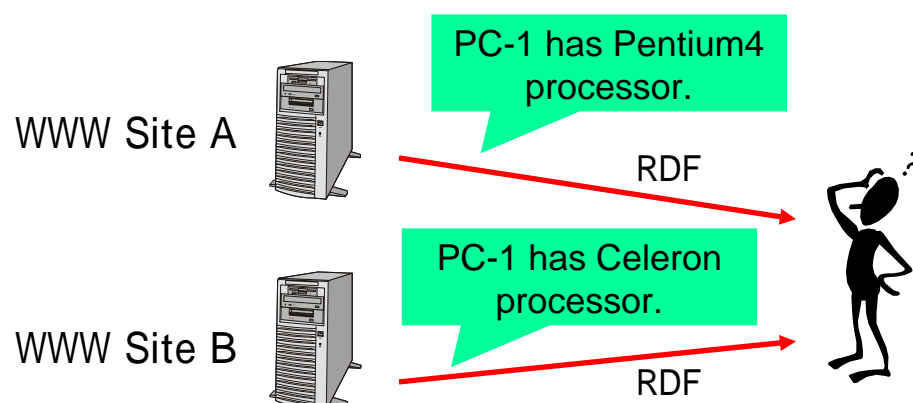
RDF共有ブックマークを使用した RDF情報の信頼性表現モデルと その応用システム

日本電気株式会社
白石 展久



RDF記述の信頼性情報の必要性

- RDF情報は、誰でも作成・発信が可能
- RDFで書かれた情報には、信頼性が高いものとそうでないものがある
- 複数のRDF情報源が、互いに矛盾した情報を発信した場合、それらのRDF情報を使用するシステムは混乱をきたす



信頼性の低いRDF情報の混在による被害

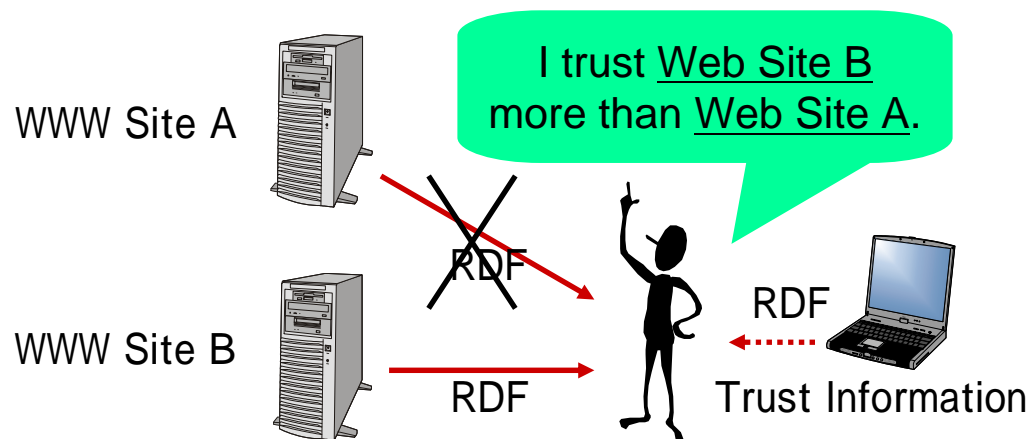


- 既存のWebベースのシステムは、何らかの信頼性情報を暗黙に仮定
 - 「連携データの範囲があらかじめ規定・制限されている」、「人間が目視によって確認する」 etc.
 - セマンティックWebベースのオープンなシステムにおいては、これらの信頼性情報を明示的に規定し、保証する必要がある
- オープンなRDF情報を自動的に連携するシステムでは、信頼性の低い情報の混入による被害は、より深刻
 - 連携対象となる情報の範囲が限定されない
 - 信頼性の低い情報を、人間が目視でチェックする機会がない

RDFによるRDF信頼性情報の表現



- 当該RDF情報が格納されていたURI(情報源URI)によって、信頼性(信頼値)を決定
- 当該RDF情報の「オリジナル性」を表すパラメータではない
 - 「当該RDF情報がどのURIによって『支持されているか』」を表す



RDFの信頼性情報は普遍ではない

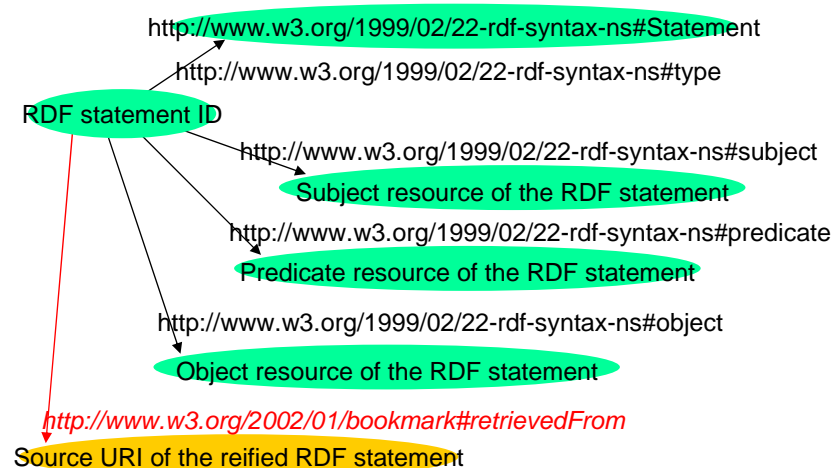


- RDF情報の信頼性はそれを使用するユーザによって異なる
 - ユーザがその情報が真であると「信じるかどうか」
 - ユーザがその情報を重要であると「信じるかどうか」
 - RDF情報の重要性・・・GoogleのPageRankと似た概念
- RDF情報の信頼性をRDFブックマークで表現する
 - 情報源URIをRDFブックマークによって重み付けする
 - RDFの信頼性情報自体もRDFで表現することにより、通常のRDF記述とRDF信頼性情報とを容易に連携させることが可能。(セマンティックWebのメリット)

RDFブックマークによるRDF信頼性表現モデル: bookmark:retrievedFromプロパティ



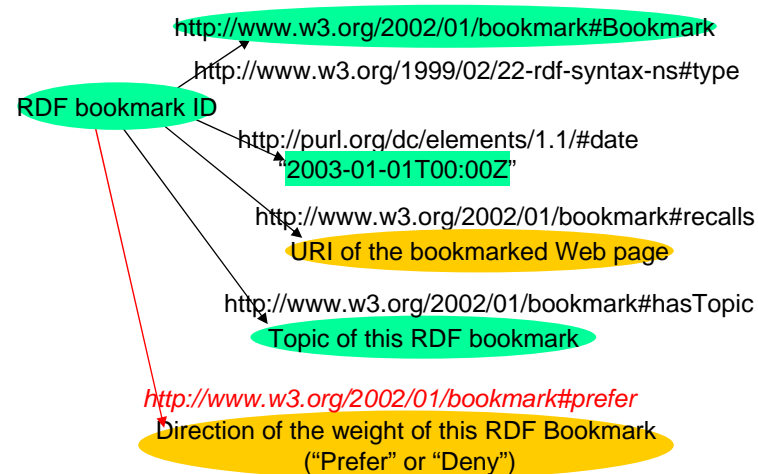
- bookmark:retrievedFrom属性は、「RDF情報の情報源URI」を表す
- RDF情報は、Reification(具体化)によって複数のRDFリソースによって表現される。本プロパティには、Reificationによって具体化された当該RDF情報の情報源のURIが格納される。



RDFブックマークによるRDF信頼性表現モデル: bookmark:preferプロパティ



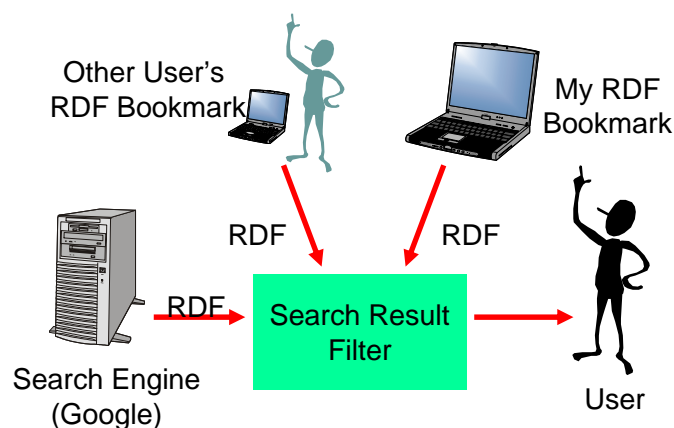
- bookmark:preferプロパティは、「ブックマークの重み付け方向」を表す
 - prefer: bookmark:recallsに格納されたURIに「賛同」(プラスの重み付け)
 - deny: bookmark:recallsに格納されたURIに「非賛同」(マイナスの重み付け)
- 本プロパティは、「prefer」か「deny」かの2値
 - ユーザの主観や、その時々での判断の揺らぎを排除



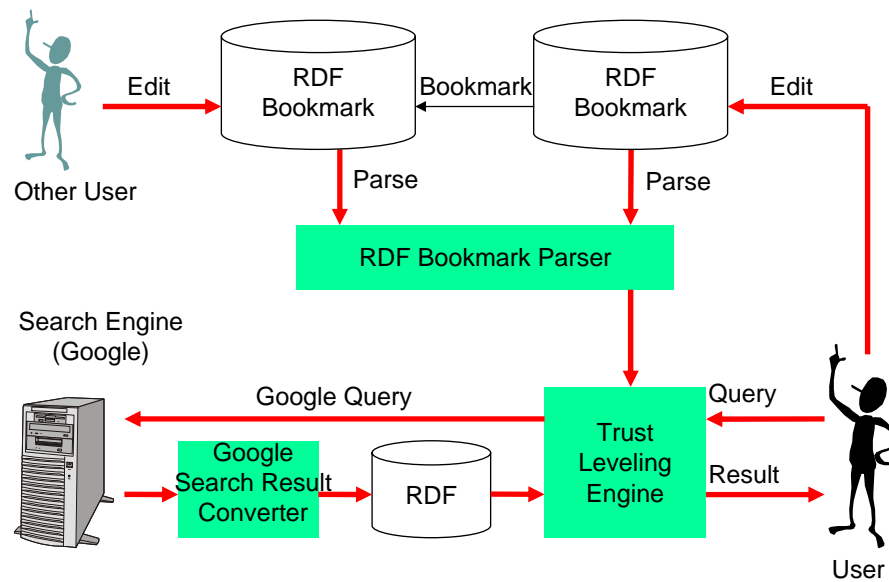
プロトタイプ: RDF共有ブックマークと 検索エンジンの自動連携



- 検索エンジンの検索結果には、ノイズが含まれる
 - ユーザは、複数の検索結果の中から目的のURIを抽出する作業が必要
- RDF共有ブックマーク情報によって、検索結果をフィルタリングし、より洗練された検索結果を提供
 - RDFで提供された検索結果情報に対して、RDFで提供されるブックマークより導かれる信頼性情報を自動的に付与し、ユーザに提示する



本プロトタイプ機能ブロック概要



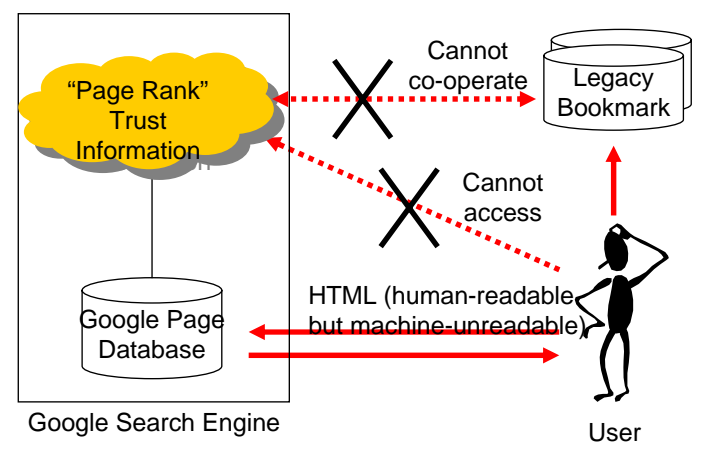
本プロトタイプでの実装機能

- Google検索結果のRDFコンバータ
 - 信頼情報付与エンジンでの処理を可能にするため、Googleから戻される検索結果を解析し、RDFに変換する。
- RDFブックマークパーサ
 - ユーザのRDFブックマークおよびそれよりブックマークされた他のユーザのRDFブックマークを解析する。
- 信頼レベル付与エンジン
 - Google検索結果RDFとRDFブックマーク解析結果を基に、信頼レベル付けされたGoogle検索結果をユーザに出力する。



従来の検索エンジン

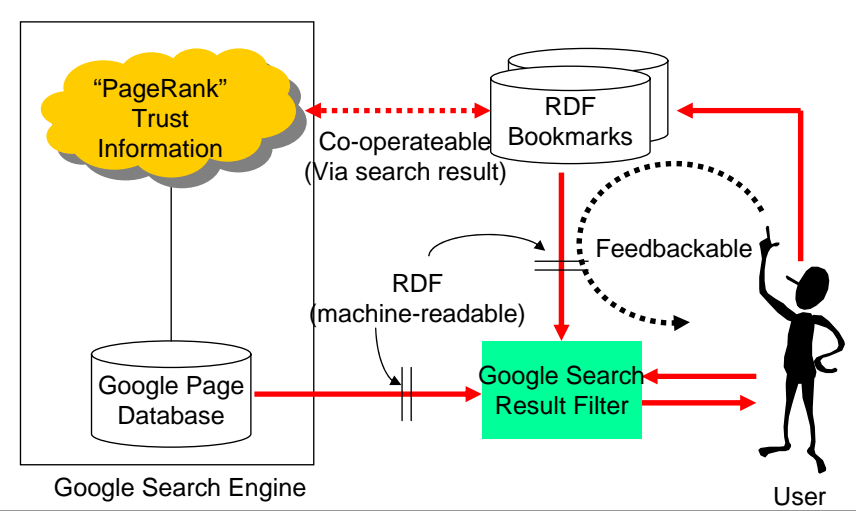
- 検索結果に対するフィードバックは不可能
 - 検索エンジン内の信頼性情報 (PageRank) に、ユーザはアクセスできない
- 検索エンジン内の信頼性情報と、ユーザの信頼性情報とが別々に管理されている
 - これらの信頼性情報は互いに連携することができない
 - ユーザはこれらの情報を、手動で別々に管理



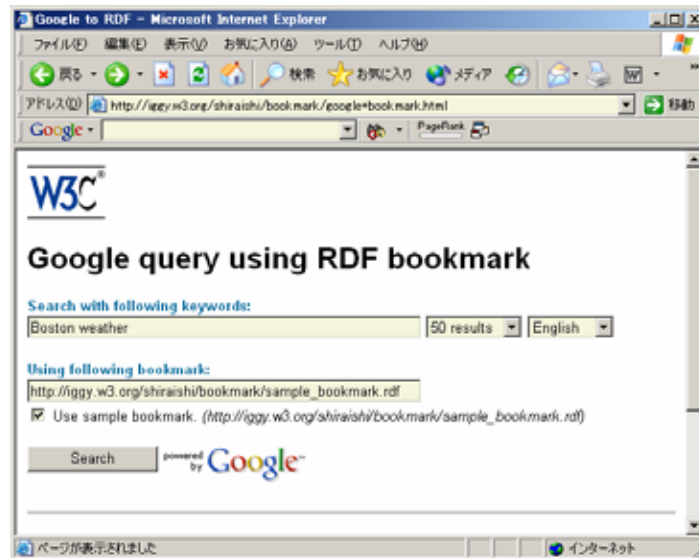
RDF共有ブックマークによる検索エンジンフィルタ



- ブックマーク内容を編集することにより、検索結果に対してフィードバックが可能
- 他のユーザによる信頼性情報 (ブックマーク) を再利用可能
- 検索結果を通じて、検索エンジン内の信頼性情報 (PageRank) と、ユーザによる信頼性情報 (ブックマーク) との連携が可能



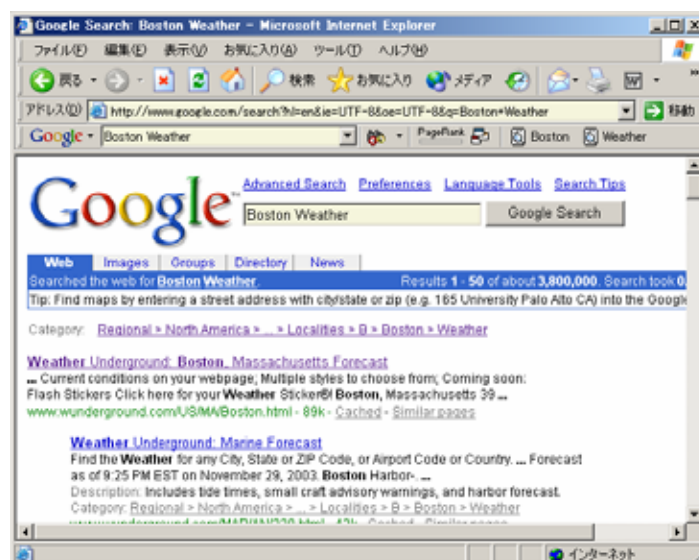
本プロトタイプのデモ (時間があれば)



以下のURIにて公開中。

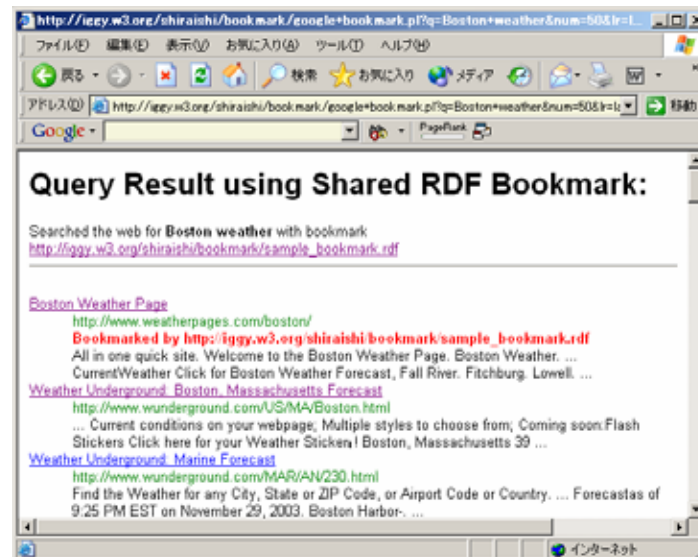
<http://iggy.w3.org/shiraishi/bookmark/google+bookmark.html>

デモスクリーンショット1: Googleでの検索結果



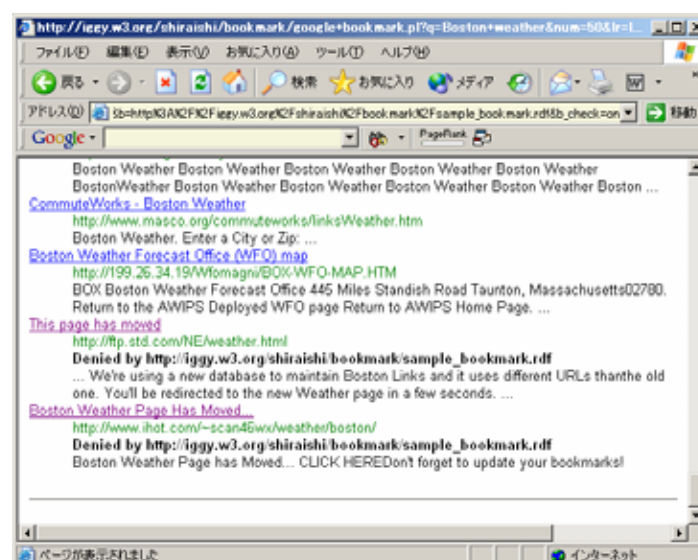
“Boston”と“Weather”をキーワードとした
Googleでの検索結果

デモスクリーンショット2: RDF共有ブックマークとの連携 ~ preferプロパティ



Preferプロパティによる検索結果の重み付け

デモスクリーンショット3: RDF共有ブックマークとの連携 ~ denyプロパティ



Denyプロパティによる検索結果の重み付け



今後の研究課題

- 本システムの信頼性情報付与メカニズムを、検索エンジンの検索結果以外の、一般のRDF記述に対して適用する。
 - ex. Web文書(勧告・技術文書等)の翻訳の分散RDFデータベースの構築とその信頼性情報付与システム
- 本プロトタイプの更なる機能拡張
 - ブックマーク内のhasTopicプロパティを信頼値の計算に考慮する



更に詳しくは、、、

- “The RDF Trust Model using RDF Bookmark and its Application”
Nobuhisa Shiraishi, NEC Corporation,
WWW Conference 2004: WF3 Contents Labeling:
Technical and Socio-Cultural Challenge and Solutions,
May 18 2004, New York City, New York, USA
 - ◆ 原文:
<http://web.sfc.keio.ac.jp/~kaz/www2004/papers/ns/www2004.xml>
 - ◆ 邦訳:
<http://web.sfc.keio.ac.jp/~kaz/www2004/papers/ns/www2004.ja.xml>