

# メタデータを用いたパソコン検索システムの試作

セマンティックWeb コンファレンスデモ展示

2003年11月17日(月)

ITエージェント株式会社

Copyright © 2003 by IT Agent Corp. Japan

## なぜパソコン検索システムなのか

★パソコンのスペックには似たような語句が並ぶため、ほしい検索結果が埋もれてしまう

例



1. DVD-RもDVD-RWもDVD+RWも検索されてしまう
2. 「倍速」という語句が強調され、4倍速以外も検索されてしまう
3. ドライブメーカー、雑誌の紹介記事、用語解説、動作推奨環境のページが多く検索される
4. メーカーの表記は「パーソナルコンピュータ」なのでこのキーワード設定では検索されない

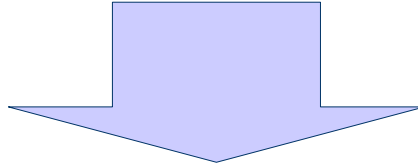
Copyright © 2003 by IT Agent Corp. Japan

※ **単純なパターンマッチングではない**

意味、係り受けの判断、解釈も必要になる  
パソコンの構成要素が予め定義されている必要がある  
字面で似て非なるものの排除、フィルタリングが必要になる  
同義語、類義語、略称を考慮した検索が必要になる

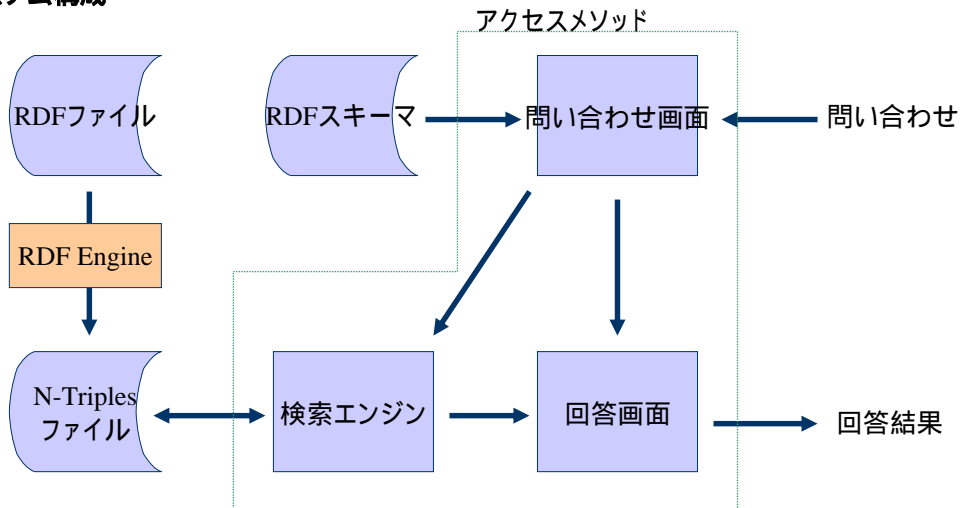
※ **ユーザの要求意図を汲み取る**

目的の情報だけが検索されるマイニング的要素が必要になる



**Semantic Web技術の応用**

システム構成

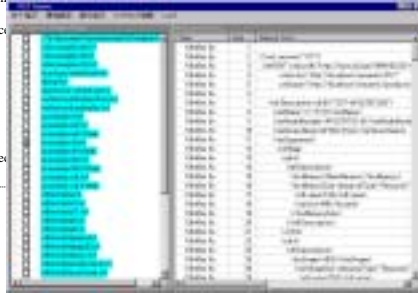


## データの流れ

```
<rdf:Description rdf:ID="ZHIHF55F40WJU">
  <ita:Maker>日立製作所</ita:Maker>
  <ita:ModelNumber>HF-W55F-40WJ-U</ita:ModelNumber>
  <ita:SeriesName></ita:SeriesName>
  <ita:Equipment rdf:parseType="Collection">
    <rdf:Description>
      <ita:Memory rdf:parseType="Collection">
        <rdf:Description>
          <ita:Main rdf:parseType="Resource">
            <ita:value></ita:value>
            <ita:Unit></ita:Unit>
          </ita:Main>
        </rdf:Description>
      </ita:Memory>
    </rdf:Description>
  </ita:Equipment>
  <ita:HomePage rdf:resource="http://www2.ec.idaten.net/jp/shosai.asp?HINCODE=ZHI-HF55F40WJU">
</rdf:Description>
```

RDFファイル

RDF Engine



N-Triplesファイル

```
<#ZHIHF55F40WJU> <ita:Equipment> <_:RDFEngine37> .
<_:RDFEngine37> <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> <_:RDFEngine38> .
<_:RDFEngine37> <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> <http://www.w3.org/1999/02/22-rdf-syntax-ns#nil> .
<_:RDFEngine38> <ita:Memory> <_:RDFEngine39> .
<_:RDFEngine39> <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> <_:RDFEngine40> .
<_:RDFEngine39> <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> <http://www.w3.org/1999/02/22-rdf-syntax-ns#nil> .
<_:RDFEngine40> <ita:Main> <_:RDFEngine41> .
<#ZHIHF55F40WJU> <ita:HomePage> <http://www2.ec.idaten.net/jp/shosai.asp?HINCODE=ZHI-HF55F40WJU> .
```

## 問い合わせ応答

```
<rdf:Property rdf:ID="OpticalDrive">
  <rdf:label xml:lang="en-US">Optical Drive</rdf:label>
  <rdf:comment>光学ドライブ</rdf:comment>
  <rdf:domain rdf:resource="#Equipment">
  <rdf:range rdf:resource="#OpticalDrive">
</rdf:Property>

<rdf:Property rdf:ID="CDROM">
  <rdf:label xml:lang="en-US">CD-ROM</rdf:label>
  <rdf:comment>CD-ROM</rdf:comment>
  <rdf:domain rdf:resource="#Equipment">
  <rdf:range rdf:resource="#CDROM">
</rdf:Property>
```

RDFスキーマ

問い合わせ画面

検索結果回答画面

詳細表示画面

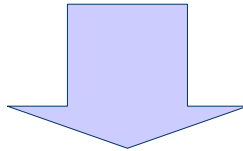


## Semantic Web技術を応用する利点を確認されたのか？

現時点ではシステムに対する要求を満たすレベルではない

しかし

- ・スキーマを利用すると単なるキーワード検索よりも目的の情報に到達しやすい
- ・目的に合致しない情報は検索結果として返ってこない



全文検索型サーチエンジンを補完するものとして有効なのではないか

### 🌟 RDF記述の自動化

現在RDF記述されていないページへのRDF記述付与  
ページ作成時のRDF記述自動付与  
スキーマの自動生成

### 🌟 システムの統合化

統合化された状態でないと試作の域を出ない

### 🌟 検索のインテリジェント化

検索キーワードの文脈自由自然言語対応  
TPO、ユーザプロフィールに応じた検索対応  
実世界情報処理