

# LODを使った 社内情報と外部情報の連携

- 条 照宣(株式会社富士通研究所)  
高梨 益樹(富士通株式会社)

## 中/長期の目標(2010年の報告)

テーマ:

セマンティックWeb技術を適用することで、  
業務システムの再定義を行う。

セマンティックWeb技術のビジネス化。

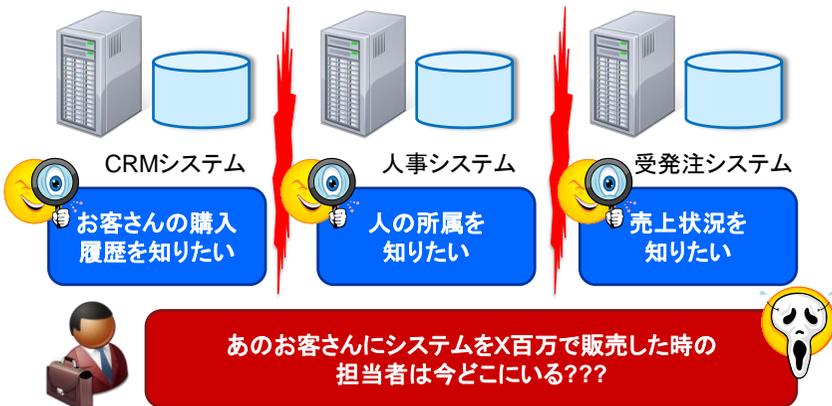
# インターネットツール

目的に応じて、必要なサービスを組み合わせることで、目的の情報を簡単に入手することができる



# 社内ビジネスツール

社内ツールは業務を遂行するために作られたシステムであり、目的以外のことに応用できない



## 予定(今年の報告)



- ▶ **メタ情報の充実**  
フォークソミ的な知識編集を普及させる方法。  
データ洗い替え(組織変更、研究体系変更)の自動化。
- ▶ **効果測定方法の確立**  
どのような指標を“見える化”するのか
- ▶ **適用部門の拡大**  
研究部門→コールセンターなどデータと業務がパターン化  
可能な業種/業務から
- ▶ **リポジトリ**  
MySQLが最適解か？ RDFリポジトリの理想型はSPARQL？
- ▶ **外部データとの連携**  
企業ユーザーに必要な情報(ロコミなど)がLOD化されるか。  
自社の製品情報をLODとして公開できるか。

## システム設計方針

- **システムの目標**
  - 「ある技術について詳しい研究員に連絡を取る」を簡単に調べられる
- **メタ情報の充実**
  - 富士通研究所内の情報をLinked Data化し保持する
    - ・ 人物情報(氏名、所属、内線番号、メールアドレスなど)
    - ・ 技術情報(人手の活用)
    - ・ 文書情報(タイトル、著者、概要など)
- **リポジトリ**
  - MySQLを使いシステムを構築してみる
- **外部データ連携**
  - DBPediaをターゲットになるべく外部データへのリンクを含める

## システムの特徴

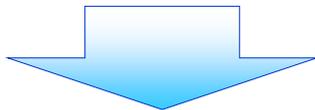
- URI検索
  - キーワードによる補完
- ネットワーク図を用いたLinked Dataの可視化



- Linked Data追加・修正
  - 検索結果が無い場合には、その場で追加できる
  - 新たなリンクを利用者が追加できる

## 実証実験

- 過去の問合せ事例を使い検索速度の評価を行う
    - 数名の被験者に対して、3種類の過去事例を与え、
      - ・ Linked Dataを使ったシステム
      - ・ 従来システム
- 2つのシステムで検索を実行してもらい、その検索時間を計測した。



**従来のシステムを利用するよりも  
40%～60%の時間を短縮することができる**

## 問題点

### ■ 外部公開データとのリンク

- 単語レベルでは公開されている一般的なものとマッチしない
- 日本語の公開データが少ないので英語化する必要がある



**日本語のデータハブが必要**

### ■ 統制語彙

- データをリンクさせるために辞書的な語彙が必要になる

### ■ 初期データ

- 運用を開始するために、ある程度の基本データセットが必要になる



**適用先を広げて確認が必要**

## 適用先の拡大

### ■ 社内複数部署に対して適用開始

### ■ 見えてきたこと

- 思っていた程、語彙辞書は必要ないことがわかってきた
- 外部データも公開されているLinked Dataだけでなく、Feedのようなフロー情報も利用できる
- 意味によるリンクよりも事実によるリンクが重要になる
- 定形業務よりも不定形業務の方がマッチする



**パターン化できない業務で、オントロジではなく  
ライブな事実に基づいたデータ連携**

# 今後の展開

部分に力点をシフトしていきたいと考えている

▶メタ情報の充実

フォークソミ的な知識編集を普及させる方法。  
データ洗い替え(組織変更、研究体系変更)の自動化。

▶効果測定方法の確立

どのような指標を“見える化”するのか

▶適用部門の拡大

研究部門→コールセンターなどデータと業務がパターン化  
可能な業種/業務から

不定形業務へ拡大

▶リポジトリ

MySQLが最適解か？ RDFリポジトリの理想型はSPARQL？

MySQLの限界

▶外部データとの連携

企業ユーザーに必要な情報(ロコミなど)がLOD化されるか。  
自社の製品情報をLODとして公開できるか

Feed情報をターゲット

shaping tomorrow with you