

オントロジの共通理解はどこまで進むのか？

セマンティックWebコンファレンス2007
パネルディスカッション「オントロジの使い方、作り方」資料
2007.1.25

日本電信電話株式会社
NTT情報流通プラットフォーム研究所
佐藤 宏之

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

セマンティックWebの魅力

- 異なる環境でのデータの再利用
 - 意味を持たせたデータの流通が可能※
- 異なるデータの統合
 - URIとオントロジを活用したデータ間のマッピングが容易

※XMLもタグで意味を持たせたデータである。しかし、これは特定のシステムでの処理を前提としている。これに対してセマンティックWebの基盤となるRDFは、任意のシステムでの処理を可能とするため、三つ組(トリプル)に代表される知識表現のデータモデルをシステム間の共通理解の範囲に置こうとしている。

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

オントロジとセマンティックWeb

- オントロジ
 - セマンティックWebだけのものではない
 - 解きたい問題や状況にあわせてさまざまな種類を使い分けて良い
 - light weight, heavy weight, domain ontology, upper ontology...
- セマンティック“Web”
 - The Semantic Web is a **web of data**.
 - さまざまなユーザ・組織の参加によってこれを実現するには、「使える」オントロジの共通理解 (common understanding) が求められるのでは？
 - 共通理解は、システム間、システムとユーザ間、そしてユーザ間でも！

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

共通理解のためのポイント(1) 誰とどのレベルまで共通理解を図るか？

- データ構造とタグによる意味の表現
 - XML
- 少し高度な知識表現
 - RDF、RDF Schema (語彙記述言語)
- より高度な知識表現
 - OWL (オントロジ記述言語)
- さらにより高度な知識表現
 - OWL+SWRL (ルール記述言語)

- もっとさらにより高度な知識表現
 - 結局、自然言語？+常識...

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

共通理解のためのポイント(2) 語彙やオントロジの共有

- 既存の語彙やオントロジ定義はたくさんある
 - Swoogle、SchemaWebなどに
 - しかし、RSS、Dublin Core、FOAFのように安心して使える語彙やオントロジは少ない
- Web利用ユーザの「自然な要求に応えた」評判の良い語彙やオントロジが欲しい
 - 程度表現オントロジ
 - コンテンツの利用パターンに基づき、コミュニティの知恵を語彙の選定に利用しているmicroformatsは興味深い
 - レビューの共有、イベントスケジュールの公開、ロケーション情報...
- メンテナンスされる共通の拠り所となる概念体系が欲しい
 - CycやWordNetに相当するクラス階層をFolksonomy的に作りあげるオントロジに期待

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

共通理解のためのポイント(3) オントロジ適用によるメリットの理解

- 「現実的な場面で」どのように役に立つ(楽になる)のか
 - owl:TransitiveProperty
 - 例えば、情報流通プラットフォーム研究所はNTTに「所属」し、プロパティ「所属」が推移的であるというオントロジ記述がある
 - 私は情報流通プラットフォーム研究所に「所属」と書くだけで、システムは私がNTTに所属していることを理解してくれる
 - owl:InverseFunctionalProperty
 - データの名寄せに使えると便利
- 効果の予測が容易か、ユーザエクスペリエンスを向上できるかが問題

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

メリットとコストのトレードオフ

- オントロジを使う
 - 計算量やメモリ使用量が増大
 - 思いもよらない推論結果、矛盾
- オントロジを作る
 - 知識を表現することで精一杯
 - OWLの文法に従って記述しただけでは本人にしか理解できないものになる恐れ

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.

おわりに

- 何でも自由に記述できる高度なオントロジ「仕様 (specifications)」と「共通理解 (common understanding)」の間にギャップがある
- さまざまなユーザ・組織を巻き込んだ“Web of Data”の実現に向けて少しだけオントロジの共通理解を底上げしたい
- オントロジを用いた推論により「新しい関係を発見」しなければいけないということはない
 - 知識表現のグラフパターンマッチングにおいて少し形状が異なっても柔軟に対処できるとうれしい cf. NTTの研究
 - Web of Data、まずはマッピングしてつなげること

But can I see my photos in a calendar to see what I was doing when I took them? Can I see bank statement lines in a calendar?

Why not? Because we don't have a web of data. Because data is controlled by applications, and each application keeps it to itself.

W3C Semantic Web Activity <http://www.w3.org/2001/sw/> より ※下線は著者

(c)2007 NTT Information Sharing Platform Laboratories. All rights reserved.