

オントロジーを利用した異種情報間の相互利用に関する研究

小野 雅史, 長井 正彦, 織田 竜也, 高野 誠二, 吉田 英嗣, 柴崎 亮介, 小口 高
 東京大学 空間情報科学研究センター

連絡先: <maono@csis.u-tokyo.ac.jp> Web: <http://shiba.iis.u-tokyo.ac.jp/~integration/>

(1) 動機: 従来, 主に紙面上に存在していた様々な一次情報(文字/記号/計測量など)が, 人手あるいはセンサなどの工学デバイスを介してデジタル化され, Web の普及とも相まって, それらの多くが Web 空間へと(デジタル化された一次情報として)投射されている。この流れに沿って, Web 空間に存在するデジタル一次情報の総量/種類は現在も増え続けている。

そうしたデジタル一次情報の一部として, 地球環境情報が存在する。ここで, 一口に地球観測情報といっても多種多様であり, 書式/性質/表現/内容などの様々な相違がある。そのため, 各固有の地球観測情報をそれぞれ相互に組み合わせることは, 例えばデータ形式自体の制約条件などが強いがため, 容易ではない。しかしながら, このように多種多様な地球環境情報を統融合し, 科学的/社会的に有用な二次情報(気候変動の影響評価など)に変換して社会に還元することができれば, 地球温暖化/大規模自然災害のような地球環境問題に向けての対応の促進につながる。

そこで, 地球環境情報のように, 有用ではあるが多種多様であるがゆえに必ずしも潜在可能性を引き出しきれない情報に対して, 相互利用できる手法を開発し社会の利益に結びつけること, これが我々の基本命題となっている。

(2) アプローチ: 一口に情報といっても, 様々な相違点がある。例えば, 辞書と分類体系とでは情報の管理構造が異なっており, 学术论文とニュースとでは情報内容の性質が異なっている, あるいは英語とタイ語とでは言語の表現体系が異なっている。

そこで, より具体的なアプローチとして次のように考えている。リソースとして, 公共機関が作成した土壌分類体系農業分野における現地調査などのデータ, タイ語のニュースサイトなどを想

定する。それらリソースをもとに, 独自に構築した専門用語辞書, 逆引き辞書などを使った情報アクセスのためのインターフェイスを開発する。そうして, ユーザが知識情報を発見するためのアプリケーション事例を提案する。

(3) 意義: 相互利用のための方法論として, 我々はオントロジーに注目している。オントロジーの定義に対しては, 古くはギリシャ哲学における Aristotle から最近では AI における Gruber や Guarino らの議論に至るまで, 実に様々な解釈が存在するが, 要点としては, ある特定の定義域における概念集と語彙集を論理により明示的に記述し体系化すること, と整理できる。そこで, 我々はこのオントロジーの方法論を用い, 複数の知識情報体系の孤島をまず作成し, しかるのち孤島の間には橋を架けるための手法を研究している。このアプローチは, 近年タコツボ化されたと言われた知識情報源に横串をさすことで, 横断的な利用可能性を切り開くという意味において, 有意義である。

(4) 特徴:

- ・ 多種多様な資源データ: 現地調査/文献情報/ニュース記事/英語やタイ語の Web サイトなど, さまざまな情報源を扱う。
- ・ 情報の発見支援: 人名/地名などの固有名や, 専門用語をブックとして, 専門用語辞書を使ったインターフェイスを通して発見するシステムを開発。
- ・ 新たな知識への導引: 既知の知識(一次情報)の比較/参照を通して, 人の発見あるいは方法論の適用により, 応用的な二次情報への変換を支援する。

(5) その他: 本研究は, 科学技術振興調整費「地球観測データ統合/情報融合基盤技術の開発」(代表: 柴崎亮介)の支援を受けている。



図 1: オントロジーによる相互利用性・情報サービス機能の開発