

## 情報知識学会2007年度第2回研究会報告

日時：1月13日（土）14時半～17時  
会場：科学技術振興機構西日本支所  
論題：シソーラスのオントロジ化のための用語関係の拡張について（発表当日、発表内容によりふさわしい、この論題に修正されました）  
発表者：韓相吉氏（韓国・大林大学）  
共催：日本図書館研究会情報組織化研究グループ（2008年から整理技術研究グループは名称変更しました）  
出席：14名  
内容：

発表者らのグループではこれまで、各種の情報検索システムで用いられるシソーラス（件名標目表を含む）を開発してきた。また、それらの経験をもとに約58万語からなる大規模シソーラスも開発している。今回は、用語関係に関わる独自の取り組みに焦点をあてた発表であった。

### 1. 今日の情報環境とシソーラス

- 今日の情報環境におけるシソーラスには、情報検索システムで索引用語を統制する道具という役割だけでなく、全文データベースにおける自動索引環境への対応、オントロジ環境への対応、用語データベースや知識ベースとしての役割、なども求められる。それらを果たすためには、用語関係を拡張し、新しい多様な用語関係を構築する必要がある。
- 拡大する役割への対応を考えると、構築及び活用の限界も明らかになってきた。利用方法に対する無知、初期構築後の更新体制の貧弱さ、典拠の概念の欠如、汎用シソーラスの国家的な開発計画の不在、などである。

### 2. シソーラスの用語関係とその問題点

- 国際標準である ISO2788 では、同義関係（UF, USE）、階層関係（BT, NT）、関連関係

（RT）が定義されている。このうち階層関係は属関係・事例関係・全体部分関係に分けることもできるとされている。しかしこれだけでは用語関係が不足しているといわざるをえない。例えば階層関係、関連関係はノードラベルやファセットによる概念拡張が必要（ISOにもノードラベルの考え方は示されているが十分な規定がない）であるし、同義関係にも様々な種類を表現する拡張が必要である。

- また、スコープノートや分類記号の付与なども何らかの関係性を表すものである。このような付加情報を「付加関係」と名付けているが、これもまた拡張の必要がある。

### 3. 同義関係の拡張

- 同義関係について、用語の変化に伴う前後関係、北朝鮮（北朝鮮）語と韓国語の関係、一般語と学術用語との関係、にそれぞれ記号を割り当てている。また、外国語にはそれぞれ言語コードを付与する。
- さらに同義語の追加情報を格納するエリアを設け、俗語・敬語・卑称語・幼児語・女性語・男性語など、ディスクリプタに対する各同義語の性格を、明確に定義できれば記述している。

### 4. 階層関係の拡張

- 階層関係の設定には通常、複数の区分原理が用いられるので、そのままの状態では階層構造を用いた一括検索を行うとノイズが発生することがある。ROOTシソーラス等ではノードラベルを用いて整理を行っているが、ファセット（共通の属性を持った概念範疇）の導入が必要である。我々のシソーラスでは「範疇関係」「意味役関係」「属性関係」の3種のファセットを導入している。
- 「範疇関係ファセット」は階層の種類を明示するもので、各関係を全体部分関係・事例関係・一般化関係に分類する。このうち「一般化関係」（属関係にあたる）はさらに、「意味的一般化」（「気象災害」－「冷

害)のような場合)と「造語的一般化」(「信号」－「デジタル信号」のように複合語となる場合)に分類する。

- 「意味関係ファセット」は複合語(すなわち造語的一般化)の場合に、修飾部(「デジタル信号」なら「デジタル」)の果たす役割を示すものである。「デジタル信号」の場合は「方式」役割であり、その他に「根源」「対象」「道具」「時間」「目標」「場所」「行為者」「受患者」がある。
- 「属性関係ファセット」も複合語に関わるものである。こちらは修飾部に当たる単位概念の属性を示すものであり、「分野・理論・方法」「感覚・感情」「物質・材料」「状態・性質」「内容」「行為」など約20種がある。

## 5. 関連関係の拡張

- 関連関係を、「反意」「同義」「上位」「下位」「一般」と区分する。このうち同義・上位・下位は、何らかの理由で同義関係あるいは階層関係と設定されなかった、類似の関係である。
- さらに「一般」関係は、「構成要素」「機能」「道具」「附属」「手段」「用途」「類型」「処理」に種類分けされる。例えば、「ネットワーク」に対して「プロトコル」は「道具」関係、「電子商取引」に対して「XML文書」は「手段」関係である。
- また、人名・地名・団体名・作品名・商品名などの固有名詞に対して、「属性関係」を付与する。例えば、人名であれば異名、略歴、生没年、著作、受賞等のデータを属性要素として付与している。

## 6. 付加関係の拡張

- 各用語の意味属性を「概念ファセット」として付与する。これは4.で述べた「属性関係ファセット」と同一のものをを用いるが、属性関係ファセットが複合語の修飾部を対象とするのに対して、概念ファセットは対象用語自らの属性に対して付与される。階

層関係では上位語の概念ファセットは下位語に継承される。

- また、各用語の種類を表す「用語範疇」を付与している。「主題語」「人名」「地名」「統一書名」などであるが、例えば主題語なら「法律名」「動物名」「記念物名」などさらに細かくしている部分がある。
- 分類記号の付与を積極的に行っている。経済的である、検証された分類体系に基づくものである、他の分類体系との互換性もはかれる、等の利点が考えられるためである。

## 7. おわりに

- シソーラスの構築作業は労働集約的で、終わりが無い。
- シソーラスはオントロジにすることで進化する。用語関係の拡張はオントロジ化への一つの道程である。
- 用語と意味関係の情報はオントロジへの変換が可能である。また大規模な用語集合から分類記号を利用して分野別及び用途別オントロジを構成していくことも可能と思われる。
- 今後、用語数の画期的な拡大、多様で標準化された概念関係名の設定、実在的な知識の表現に欠かせない属性関係の把握と定義、名詞だけでなく形容詞及び形容動詞のためのシソーラス構築、などが課題である。