

説明に用いられる視点に着目した知識メタデータの構築*

堀田 久貴 (学籍番号 200721557)

研究指導教員：中山 伸一

1. はじめに

知識の流通量の増加は、目的の知識がどこにも記述されていないという問題が回避されるというメリットが得られる一方、本当に欲しい情報がノイズに埋もれてしまうというデメリットも発生させてしまう。そのため、より効率的に知識を得るための方法が求められる。

本研究では、情報の探索におけるメタデータや文献探索における構造化抄録を参考に、用語の検索結果にどのような知識が含まれているかを明確にし、内容レベルでの絞込みを行うための知識メタデータの作成を行う。また、知識メタデータに基づいた包括的な記述のためのテンプレートを作成し、このテンプレートに基づいた電子百科事典の作成について検討する。

なお、本研究においては、用語を概念として、概念について記述される個々の説明文を知識、それらの集合を情報として扱う。

2. 構築方法

知識メタデータは、まず、国語辞典の説明文から説明の視点の抽出を行い、その独立性を検討することにより原版をまとめる。

次に、電子百科事典にその知識メタデータを付与する作業を行い、問題点の発見を行う。

最後に、それを反映した最終版を作成する。

なお、原版の作成には新潮現代国語辞典第二版を用いた。知識メタデータ（原版）の付与はエンカルタキッズ百科の1470語を対象にして、スーパーニッポニカ、ポプラディア、マイペディアを用いて行った。

3. 結果

3.1 知識メタデータの構築

国語辞典から23の説明の視点を抽出し、「モノ」「コト」「動的表現」「静的表現」の指標を用いて独立性の検討を行い、19の項目からなる知識メタデータ（原版）を作成した（表2参照）。

電子百科事典の説明文にこの知識メタデータの付与を行った結果、「モノ」「コト」の扱いが実際の説明文にそぐわないこと、及び直感的な運用が難しいという問題点が明らかになった。

「モノ」「コト」の視点はずし、場所のような一部の説明の視点が他の説明の視点に従属的な項目を除いて、表1に示す14の項目からなる知識メタデータ（最終版）を作成した。

表1 知識メタデータ（最終版）の項目

項目区分	項目
本質的特長 性質 発生変化 利用 性質	外見
	構成・成分
	全体
発生変化	出来方
	変わり方
利用	機能
	用途・用法
歴史	変遷
概念的関係	上位概念
	下位概念
	対比
用語的關係	等価の用語
	反対語

* “Construction of knowledge meta data based on categorization of explanation” by Hisaki HOTTA

また、運用を直感的に行うためのツールとし

て、エンカルタキッズ百科事典の用語の区分を基に知識メタデータと用語区分の対応表を作成した。

3.2 電子百科事典の説明文の傾向

電子百科事典マイペディアの説明文に付与した知識メタデータの状況を分析した結果、項目毎の使用回数は表 2、使用項目数毎の用語頻度は図 1 のように得られた。

表 2 電子百科事典における付与された知識メタデータ(原版)の項目の傾向

項目	用いられた回数	用いられた割合
出来方	55	4.2%
変わり方	54	4.1%
外見	147	11.1%
構成・成分	792	59.9%
全体	26	2.0%
機能	102	7.7%
変遷	817	61.8%
用途	151	11.4%
用法	88	6.7%
性質	421	31.8%
上位概念	1089	82.4%
場所	413	31.2%
値・量	507	38.4%
等価の用語	596	45.1%
反対語	34	2.6%
対比	109	8.2%
現象	74	5.6%
状態	50	3.8%
下位概念	52	3.9%

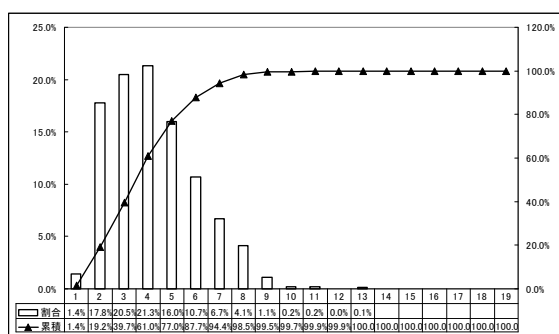


図 1 電子百科事典の一つの用語に対して付与された知識メタデータの項目数

4. 考察

作成した知識メタデータ(最終版)は、フレームや意味ネットワークなどの知識表現法による一般的な知識の表現の形式を全て包含しており、変遷等の時系列を扱うような知識表現法で

は扱いにくいタイプの知識の記述も扱えていることが確認できた。また、特定の概念について 14 全ての項目についての説明文を作成することができた。以上のことから知識メタデータの妥当性を示した。

また、知識メタデータ(原版)の項目毎の付与状態の分析の結果、4 つが特に多く使われている傾向と、電子百科事典に項目のあった 1322 語では、1 語当たりで用いられている説明の視点が 7 個以下の用語が 94.4% に及ぶ傾向が明らかになった。このことから、電子百科事典の説明文は包括的とは言えないことが分かった。この結果を受け、複数の用語について包括的な記述の為のテンプレートによる説明文の作成を行い、実際に説明文が作れることの確認と記述のためのテンプレートの有効性を確認した。

5. おわりに

知識メタデータを作成し、その有効性の確認をした。また、運用のための対応表を作成した。

さらに、知識メタデータをテンプレートとして説明文を作成することによる包括的な百科事典の作成について提案した。

今後の展望としては、限定的で専門性の高い知識へ知識メタデータが対応できるかの検討が考えられる。

文献

- [1] 堀田久貴, 他. 知識表現としての説明の仕方の分類:国語辞典と生物教科書からの検討. 情報知識学会誌. 2008, vol. 18, no. 2, p. 83-86.
- [2] 斉藤孝. 「記録・情報・知識」の世界. 東京, 中央大学出版会, 2004. 322p.
- [3] 村上陽一郎, 他, 大澤光編著. 感性工学と情報社会:「感性工学は情報社会の課題にどう取り組もうとしているのか」. 東京, 森北出版. 2000. 219p.