

階層的自動タグ付けによるエキスパート検索エンジンの研究・開発

○大平 倫宏^{*1)}、富山 真一^{*1)}

1. 目的・背景

現在、企業内や研究グループ内等の問題解決においては、特定の専門知識を活用することが重要になっている。このためには、そのグループ内で、適切な知識を持った専門家を見つける必要がある。近年では、専門家を検索するためのエキスパート検索システムが考案されているが、現状で提案されているシステムでは、検索語が類義語を持つ場合や、一人の従業員が複数の専門分野等を持つ場合において、正しく機能しない場合が存在する。本研究では、類義語に対応した階層的タグ付けを利用することで、検索語が類義語を持つ場合でも専門家をリストアップでき、かつ一人の専門家が複数の専門分野を持つ場合でも、階層的タグと検索語の関連度からふさわしい専門家をランキング可能なエキスパート検索の方法の研究・開発を行った。

2. 研究内容

(1) 実験方法

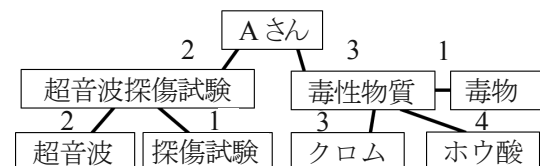
従来、よく用いられる方法として、図1のように専門家に対するタグを列挙し、そのタグとの関連性を元に検索を行う方法がある。それに対して、本研究では、社内文章やウェブ上の情報を基に、自動的に図2のような階層的タグ付けを行う方法を開発した。類義語は横方向に、カテゴリ分類は縦方向にタグをリンクしている。検索時には、検索語と人物名のタグ間の距離の総和を利用して、ランキングを行う。

(2) 結果及び考察

図1に示す従来型の方法と、図2に示す提案法の比較を行った。比較方法としては、あらかじめ検索語とそれに対する各専門家の関連性を定めておき、ランキングの正当性を示す指標である NDCG@5 を用いて評価した。NDCGの値が大きいほど正当性が高い。表1から、提案法が従来法に比べ、優れていることが分かる。特に、従来法では図1のような例で、検索語として「超音波」と「害虫駆除」を選んだ場合に、Aさんを高くランク付けすることが多い。それに対して提案法では、「超音波探傷試験」に関連する「超音波」と「毒性駆除」に関連する「害虫駆除」を区別して捉えるため、Aさんを不当に高くランク付けすることはない。タグを階層化することにより、一人の専門家が複数の専門分野を持つ場合でも、正確に評価が可能となる。

専門家名	Aさん
タグ	超音波、探傷試験、 毒性物質、害虫駆除

図1. 従来の専門家へのタグ付け方法



(数字はタグ間の距離を表す)

図2. 階層的なタグ付け方法

表1. NDCG法による比較

	従来法 (図1)	提案法 (図2)
NDCG@5	0.612	0.814

3. 今後の展開

新たな分野での製品開発を行う企業においては、その分野の専門家とコンタクトを取ることが重要である。また、成長企業では、人数が増加するにつれて、誰がどの分野の専門家かを把握することは困難となる。このような開発型中小企業に向けて、本研究成果の活用を働きかけていきたい。

*1)情報技術グループ