

## システム情報工学研究科修士論文概要

年 度	平成 26 年度	学位名	修士( 工学 )
専 攻	知能機能システム	専攻	著者氏名 守谷 一朗
指導教員氏名 宇津呂 武仁			
論文題目 質問回答サイトおよびウェブからのノウハウ知識の相補的収集			
論文概要 <p>インターネット上には様々な情報があり,多くのユーザはウェブページから日常の行動に役立つ知識を得ている.知識を得るための代表的なウェブサイトとして,Wikipediaをはじめとする百科事典サイトやYahoo!知恵袋をはじめとする質問回答サイトが挙げられる.特に,質問回答サイトでは,「花粉症の対策方法」や「結婚式でのスピーチの仕方」といったユーザの日常の行動に役立つノウハウ知識が多く掲載されている.一方で,質問回答サイトやウェブ上に含まれる情報は膨大であり,ユーザにとって役立つノウハウ知識を集約して提示することが求められる.そこで,本研究では,ある検索対象についてのノウハウ知識を網羅的に収集し,集約・俯瞰する手法を確立する.ここで,質問回答サイトには多くのノウハウ知識が含まれているが,質問回答サイトだけでは十分でないことが想定される.例えば「結婚式の電報の文例」を考えると,質問回答サイトから得られる限定的な文例情報だけでなく,電報の文例を専門的に扱っているサイトが紹介している網羅的な文例情報を合わせて参照した方がより有益である.そこで,本論文では,質問回答サイトおよび一般のウェブページという二種類の情報源を併用することにより,ノウハウ知識を相補的に収集し,集約・俯瞰する手法を提案する.</p> <p>本研究では,まず,質問回答サイトから収集した質問回答事例,および,検索エンジン・サジェストを索引として収集されたウェブページの混合文書集合に対してトピックモデルを適用することにより,話題のまとまりを生成する.次に,話題のまとまりごとに二種類の情報源の記事を確率の高い順に10件ずつ目視にて分析し,3件以上同一とされる内容の話題を情報源ごとに抽出する.次に,抽出した話題を,「ノウハウ知識」,「ノウハウ以外の知識」,「意見」,「その他」の4つに分類することで,ノウハウ知識を手手で選定する.最後に,得られたノウハウ知識を内容ごとに人手で大分類にまとめる.一例として,「花粉症」に関するノウハウ知識を収集した結果においては,合計55個の話題が収集された.収集された話題の中には,「花粉症の温熱治療のための吸入器」のように,ウェブページのみから得られるノウハウ知識が合計で19個あり,全話題の約35%となった.このように,本研究において,質問回答サイトのノウハウ知識を集約し,さらに,質問回答サイトには含まれないノウハウ知識をウェブページから補えることを示した.</p>			
審査日	平成 27 年 1 月 29 日		
審査員	(大学名 職名)	(学位)	(氏名)
主査	筑波大学 教授	博士(工学)	宇津呂 武仁
副査	筑波大学 教授	博士(工学)	中内 靖
副査	筑波大学 准教授	博士(工学)	矢野 博明