

平成19年度科学研究費補助金実績報告書（研究実績報告書）

1. 機関番号 1 4 6 0 3 2. 研究機関名 奈良先端科学技術大学院大学
3. 研究種目名 特定領域研究 4. 研究期間 平成19年度～平成20年度
5. 課題番号 1 9 0 2 4 0 5 7
6. 研究課題名 経験マイニング：Web文書からの個人の経験の抽出と分類

7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
6 0 2 7 2 6 8 9	フリガナ イヌイ, ケンタロウ 乾, 健太郎	情報科学研究科	准教授

8. 研究分担者(所属研究機関名については、研究代表者の所属研究機関と異なる場合のみ記入すること。)

研究者番号	研究分担者名	所属研究機関名・部局名	職名
	フリガナ		
	フリガナ		
	フリガナ		
	フリガナ		
	フリガナ		

9. 研究実績の概要(国立情報学研究所でデータベース化するため、600字～800字で記入。図、グラフ等は記載しないこと。)

下欄には、当該年度に実施した研究の成果について、その具体的内容、意義、重要性等を、交付申請書に記載した「研究の目的」、「研究実施計画」に照らし、600字～800字で、できるだけ分かりやすく記述すること。また、国立情報学研究所でデータベース化するため、図、グラフ等は記載しないこと。

商品やサービスなど、様々な事物（トピック）に関する個人の経験を広くWeb文書集合から抽出し、データベース化する技術を開発する。具体的には、指定されたトピックに関連する個人の経験の記述をWeb文書集合から収集し、述語項構造に基づく表現形式に構造化するとともに、事態タイプ（ポジティブ/ネガティブな出来事・状態、入手・利用等の行為など）や事実性情報（当該事態の時間情報とそれに対する話者態度）といった意味情報を解析する。これにより、Web上に散在する膨大な数の個人の経験が意味的な索引とともにデータベース化され、〈トラブル〉や〈要望〉といった意味による柔軟な検索が可能になる。19年度は課題全体の設計を行うとともに、次の3点で成果を得た。

①事実性解析：まず複数のトピックに関するブログ記事を人手で分析することによって、タスクの要件を洗い出し、テンス・アスペクトの一般化である時間情報とモダリティの一般化である態度情報の解析の組み合わせからなる課題を設計した。つぎに、約4千件の訓練事例を作成し、factorial CRFにより一定の精度で解ける見込みを得た。

②事態タイプ分類：事態タイプの中核となる評価極性、とくに名詞の評価極性をコーパスから自動獲得する実験を行い、既存の方法を大幅に改善できることを示した。

③応用システム：以上の成果を利用し、経験に基づくブロガー検索システムを開発した。このシステムは、特定の商品に関する人々の経験を〈興味あり〉〈買った〉〈満足〉などの経験クラスで検索し、結果をブロガー単位で表示する。例えば「iPod touchに興味を持っている人」、「実際に買った人」、「満足している人」を集めることができ、マーケティング等の情報源として有効活用できると考えている。

※ 成果の公表を見合わせる必要がある場合は、その理由及び差し控え期間等を記入した調書(A4判縦長横書1枚)を添付すること。

10. キーワード

- | | | |
|---------------|----------|----------|
| (1) テキストマイニング | (2) ブログ | (3) 経験 |
| (4) 自然言語処理 | (5) 意味解析 | (6) 情報抽出 |
| (7) 事実性解析 | (8) 評価極性 | |

(裏面に続く)

11.研究発表（平成19年度の研究成果）

〔雑誌論文〕 計（ 2 ）件

著者名	論文標題			
Nozomi Kobayashi, Kentaro Inui, and Yuji Matsumoto	Extracting Aspect-Evaluation and Aspect-Of Relations in Opinion Mining			
雑誌名	査読の有無	巻	発行年	最初と最後の頁
Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)	有		2007	1065-1074

著者名	論文標題			
乾健太郎, 鳥澤健太郎	WWWからの知識獲得			
雑誌名	査読の有無	巻	発行年	最初と最後の頁
日本語学	無	27 (2)	2008	48-61

〔学会発表〕 計（ 4 ）件

発表者名	発表標題		
乾健太郎, 原一夫	経験マイニング：Webテキストからの個人の経験の抽出と分類		
学会等名	発表年月日	発表場所	
言語処理学会第14回年次大会論文集	2008.3.20	東京	

発表者名	発表標題		
東山昌彦, 乾健太郎, 松本裕治	述語の選択選好性に着目した名詞評価極性の獲得		
学会等名	発表年月日	発表場所	
言語処理学会第14回年次大会論文集	2008.3.20	東京	

発表者名	発表標題		
原一夫, 東山昌彦, 乾健太郎, 松本裕治	経験マイニングのための事実性解析		
学会等名	発表年月日	発表場所	
情報処理学会第70回全国大会	2008.3.14	筑波	

発表者名	発表標題		
原一夫, 乾健太郎	事態抽出のための事実性解析		
学会等名	発表年月日	発表場所	
情報処理学会研究報告自然言語処理研究会	2008.1.22	東京	

〔図書〕 計（ 0 ）件

著者名	出版社		
書名	発行年	総ページ数	

12. 研究成果による産業財産権の出願・取得状況

〔出願〕 計（ 0 ）件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	出願年月日	国内・外国の別

〔取得〕 計（ 0 ）件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	取得年月日	国内・外国の別

--	--	--	--	--	--

13. 備考

※ 研究者又は所属研究機関が作成した研究内容又は研究成果に関するwebページがある場合は、URLを記載すること。

<http://cl.naist.jp/~inui/research/EM/>