

## 研究成果最適展開支援事業 (A-STEP) FS ステージ (シーズ顕在化) 事後評価報告書

プロジェクトリーダー (企業責任者) : (株) ウィリルモバイル

研究責任者 : 京都大学 黒橋 禎夫

研究開発課題名 : 言語解析技術のインターネット広告の基盤技術としての活用可能性の検証

### 1. 研究開発の目的

本研究では、日本語解析に使われる言語解析技術が、インターネット広告市場において活用される web サイトと広告をマッチングさせる技術の要素技術となるかどうかについて検証を行う。

現在、インターネット広告市場では、人 (web サイト) に対して広告や宣伝が展開されているが、web サイトの内容と広告のマッチング技術が欠けているため、広告の訴求力が低く、web サイト訪問者にとって有益な情報が提供できていない点が問題視されている。そこで、本事業では京都大学で研究されている言語解析技術を活用する事で、web サイトのページ内容と広告の内容を解析し、その解析結果から最適なマッチングを図ることを目的に研究開発を行う。

### 2. 研究開発の概要

#### ①成果

本研究ではマッチングエンジンの基礎技術となる、広告カテゴリの特徴語の自動獲得技術、口語調テキストの頑健解析と新語・専門語の自動獲得技術の研究と開発を行った。プロジェクトの目標として、マッチングエンジンのβ版システム開発を目標とした。

期間中に、広告カテゴリの特徴語の自動獲得技術については、282 のカテゴリ選定を行い、それぞれに対応した特徴語を抽出し、100%の達成となった。

口語調テキストの頑健解析と新語・専門語の自動獲得技術の研究と開発については、文章の構成から品詞を判断し、未知語の検出を行う技術の確立と、Wikipedia の情報から新語や固有名詞、人名を抽出する技術が確立できたため 90%の達成となった。

マッチングシステムβ版については開発が完了し、市場でのテスト運用過程まで達成しており、実用化に繋がる成果が得られた。

#### ②今後の展開

今後は広告配信システムの改良とサービスのプロモーションを中心に進めていく予定。

市場でのテストの結果、本研究とは別の開発事項も多く挙がっているため、実用化に向けた開発も追加し、進めていきたい。本研究に関する研究開発では、口語対応、新語・未知語への対応を引き続き大学側と連携して進めていきたい。

### 3. 総合所見

モデルが提案され、必要なデータが収集されるという点で、当初研究計画は遂行された。また、これに基づきブログや携帯サイトにマッチした広告を自動的に選んで配信するための技術開発で、マッチング率を上げるために、大規模な広告コーパスの収集と関連語の大規模な抽出を行い、広告カテゴリの抽出を行った。自然言語処理の適用によりマッチング度が向上されたとの試行結果が得られているが、今後、本手法の検証方法の確立と計数的評価が必要と思われる。