



## 2008 年度下期未踏 IT 人材発掘・育成事業 採択案件評価書

### 1. 担当PM

勝屋 久 PM (Venture BEAT Project 主宰)

### 2. 採択者氏名

チーフクリエイター: 水野 淳太(奈良先端科学技術大学院大学  
情報科学研究科 自然言語処理学講座 博士前期課程 2 年)  
コクリエイター: 村田 祐一(テルモ株式会社 開発技術部設備技術課 社員)

### 3. プロジェクト管理組織

日本エンジェルズ・インベストメント株式会社

### 4. 委託金支払額

5,357,000 円

### 5. テーマ名

情報検索のための学習成長型対話システムの開発

### 6. 関連Webサイト

なし

### 7. テーマ概要

情報が氾濫しているといわれる現在、情報を検索する手段はいくつかあるが、主要なのはウェブ検索である。しかしウェブ検索では情報収集に対してユーザ側の努力を要求する。つまり、良い情報を得るためには良い検索を行う必要があり、それは検索

クエリの設定という経験に基づいた能力を要求する。

さらに、得られた情報が最良のものであるかは判断できない。もっと良い検索が行えるときに、過剰な検索クエリによって得られるはずの情報が失われていることがある。

それに対し、ここで提案する自動情報収集システムは、ユーザの要求する情報についてユーザの想定以上に網羅的に収集し、ユーザの必要な情報だけを選別して提示するシステムである。さらにユーザ特化を適用することで、ユーザがどのような情報を日常的に要求しているかを判定し、必要な情報のみを積極的に絞り込む。これらの機能を対話型インターフェースで提供することで、計算機に不慣れなユーザでも容易に情報を入手することを実現する。

また、対話型インターフェースは対話履歴を得ることができ、その履歴を活用することで、よりユーザの要求に適合した検索が行えると考えている。

対話によって高精度な情報検索が行えるとなれば、ユーザは積極的に対話を行うようになるだろう。そしてその対話は情報検索をさらに高精度化する。情報検索と対話という二つの要素を組み合わせることで、優れた情報検索が行える対話システムを構築する。

## 8. 採択理由

本提案プロジェクトである情報検索のための学習成長型対話システムの開発のベースとなる技術は自然言語処理である。一般的に自然言語処理を対話システムに適応するのは難しいとされているが、未踏のテーマとしては技術面そしてビジネスモデルを構築する面において挑戦的であるが試みる価値はあると考えられる。未だ競合しそうな対象はなく、未踏のテーマとしても良いと考えられる。

## 9. 開発目標

### 1) ユーザ発話からの検索クエリの自動構築

一般的なウェブ検索では、ユーザの入力する検索クエリは単語の集合である。近年のウェブ検索では文での入力が可能になりつつあるが、文中の単語のうち、どれをクエリにすべきであるかは自明ではない。最も簡単な方法は入力された文に対して形態素解析を行い、特定の品詞の単語のみクエリとして利用するというものであるが、本システムではさらに以下の工夫を施した。

- ・新語や未知語による解析誤りを防ぐ
- ・品詞別の選択
- ・係り受け解析の利用

## 2) 検索クエリ拡張

1) 節で構築した検索クエリはユーザの入力した文から抽出しているため、ユーザにとって未知であるが有用なクエリを利用することはできない。そこで、2 つの方針で構築したクエリに関連する単語を自動的に追加できるようにした。

- ・ユーザ非依存のクエリ拡張
- ・ユーザの趣向に合わせたクエリ拡張

## 3) 複数の検索結果からのシステム発話の生成

1) 節、2) 節で、ユーザの入力した発話文に対して、ユーザの趣向を反映した拡張クエリを得ることが出来る。最終的に得られるクエリの一部を以下に示す。

## 4) ウェブブラウザ上で利用するためのインタフェースの構築

前述までの対話システムはコンソール上で利用可能であるが、ユーザビリティ向上のためにウェブブラウザ上で利用できるよう、インタフェースを構築する。本システム専用のインタフェースとして Ruby on Rails で構築した(図 10)。専用のインタフェースでは、検索エンジンの指定、クエリ拡張の有無、およびユーザの趣向を利用するかどうかなどを選択することができる。また、動画検索や商品検索では、動画および商品の画像も表示される。このインタフェースはウェブサーバ (apache2) を介して外部から利用することが可能である。もう一つのインタフェースとして twitter API を利用して twitter 上で実現した(図 11)。twitter にて本システムのアカウントを作成し、ユーザは自分のアカウントから本システムのアカウントを follow する。そして、本システムに話しかける。つまり“@システムアカウント名ユーザ発話”という形式でつぶやくと、システムは¥@ユーザアカウント名システム発話”として出力する。twitter の場合、検索エンジンを指定するといった設定を送ることが難しいため、すべての検索エンジンを利用し、クエリ拡張はユーザの“つぶやき”の履歴を2) 節で述べた方法で行う。

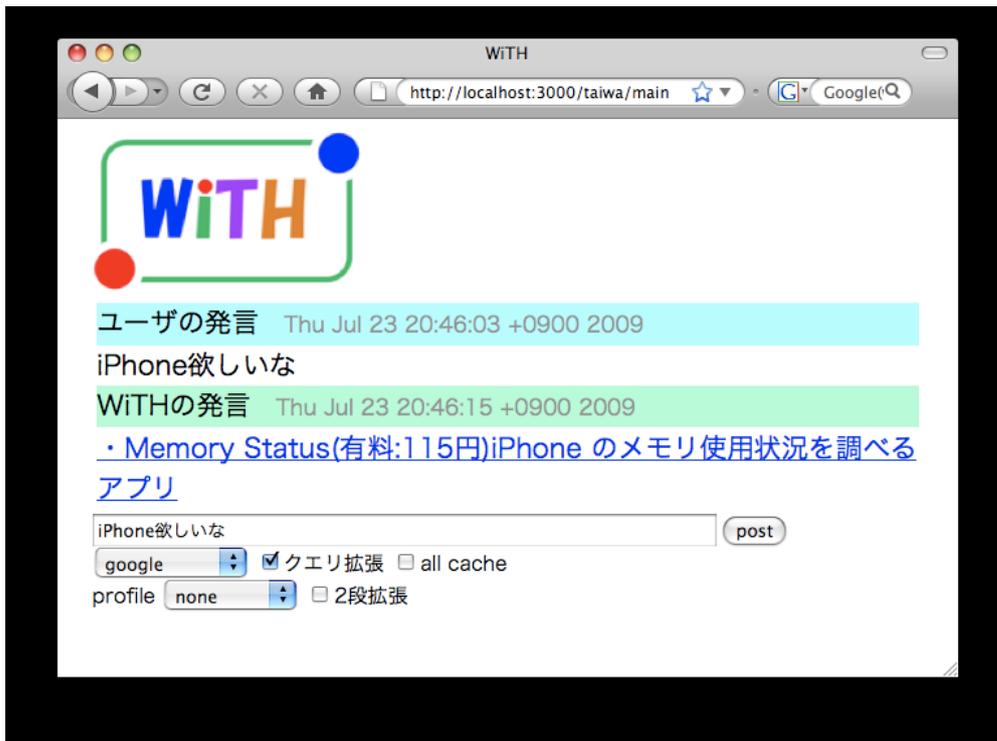


図 10 Ruby on Rails による専用インタフェース



図 11 twitter 上での本システム

## 10. 進捗概要

技術的にも、実用的にもハードルの高いプロジェクトで開発当初から、どのように目に見える形でかつ有効に具現化をするかで試行錯誤をしてきたが、合宿・中間報告会を行い、その具体的な方向性が都度みえてきた。また、開発が進むにつれ、当プロジェクトに携わるチーフクリエイター水野淳太氏とコクリエイターの村田祐一氏は役割分担・意思疎通に課題が生じたが、合宿をきっかけにプロジェクトを根本から見直し、目指す方向・優先順位・お互いを尊重した円滑なコミュニケーションの意識・役割分担の明確化を自立的に見直し、開発を手がけて、実証すべくテスト・評価も期間中に完了をした。開発期間中のプログラム開発に対する情熱も感じられた。

## 11. 成果

当該プロジェクトの成果物において、幾つかのケースで実験して、下記の概略成果がみられた。そして当該開発ソフトウェアの特徴でもある。

### 1) 検索エンジンとの比較

開発成果の特徴を、検索エンジンにウェブ検索と商品検索のどちらを利用するか大きく分け、検索エンジンとの違いを明らかにする。

#### 【ウェブ検索の場合】

本システムの発話は従来のウェブ検索と比較すると、本システムはユーザの関心を引きやすい結果を提示できると結論づけられた。(ウェブ検索と同様のことが動画検索にも言える。)

#### 【商品検索の場合】

ある有名商品検索システムと比較した場合にユーザの趣向をとらえることで、普通に検索しているだけでは到底見つけられないような商品を検索することが出来る点がみられた、本システムの大きな特徴である。

### 2) 対話システムとしての特徴

対話システムとしての特徴に、ユーザのクライアント側が速度の遅いネットワークであっても大量のウェブ検索結果を解析した結果を高速に入手できる点があげられる。twitter の場合はもちろんだが、Ruby on Rails によるインタフェースをウェブ上で利用する場合でも、ユーザのクライアントと対話システムを実行するサーバとの間で通信される内容は、ユーザ発話と一部のパラメータ情報(利用する検索エンジンの種類など)、システム発話のみである。ウェブ検索を含めた一連の処理はすべてサーバ側で実行されるため、PHS や 3G のように高速ではないネットワーク

を利用して多数のウェブページを検索できない場合でも、クライアント側にはほとんど負荷を与えずシステム発話を得ることができる。

### 3) twitter での可能性

本システムは twitter 上で利用できるため、一つの特徴として、「どんな話題に対しても気の利いた返事を返してくれる」ということがあげられる。ユーザは必ずつぶやき返してくれる follower として、本システムを楽しむことが出来る。また、つぶやきの履歴からユーザの趣向を推定することが出来るので、趣向の似た別のユーザを紹介することも実現可能である。(システムの構成上、可能であるが、実現はできていない。)これは、従来の対話ボットでは実現できていない機能である。

## 12. プロジェクト評価

本プロジェクトは、有効な情報検索の行えるパーソナルな対話システムの開発をめざし挑戦的であり、面白いテーマだった。実証テストでもある程度の評価がでて、難しいテーマで完了させたことは評価できる。実用の観点では幾つかの課題は残るが、この技術をベースに今後も継続的に質を高め、研究だけではとどまらず、一般公開から事業化へつなげてほしい。

また、人材育成的な観点においては本プロジェクトの経験をとおして、自己表現力、コミュニケーション能力、市場・ユーザー志向の考え方の方法、ビジネスマインド、創造力、アントレプレナーシップを向上できたと考えられる。

## 13. 今後の課題

今後社会・市場に広める上での考慮点は下記となる。

- 1) 当該ソフトウェアの性能向上
    - ・クエリ拡張手法の性能向上
    - ・システム発話選択の性能向上
  - 2) ユーザビリティの完成度
  - 3) 事業化推進のためのプロモーション・プロジェクト体制
- 今後も必要な時点で個別にアドバイスを行ってゆきたい。