

2005年8月4日

(報道発表資料)

日本電信電話株式会社
NTTレゾナント株式会社

「gooラボ」上で新たな情報収集機能の共同実験を開始

～最新のニュースやブログ記事の中から求めるジャンルの情報を的確に収集する「関心事（かんしんじ）アンテナ」～

日本電信電話株式会社（以下NTT、本社：東京都千代田区、代表取締役社長：和田紀夫）と、ポータルサイト「goo」を提供するNTTレゾナント株式会社（以下NTTレゾナント、本社：東京都千代田区、代表取締役社長：資宗克行）は、最新のニュース記事やブログ記事の中から関心事の情報をもれなく提示するサービス「関心事アンテナ」の商用環境での検証・評価のための実験を、ポータルサイト「goo」上に開設した実験サイト「gooラボ」（<http://labs.goo.ne.jp/>）上で開始します。

今回、NTTサイバーソリューション研究所（以下NTT研究所）が世界で初めて開発した「日本語概念フィルタリング技術（concept based web search）」は、インターネット上で言葉を検索する時に未知語（固有名詞等の検索データベースに登録されていない新しい言葉）を検出すると、データベース上の類似する言葉や前後の文章から自動的にその意味内容を認識し、検索データベースに追加することにより、インターネット上でユーザが必要とする情報を漏れなく検索することが可能となる技術です。

本実験は、この技術を活用し、新しい情報収集サービスの実現を目指す取り組みです。幅広く情報を収集するための「アンテナ」としてキーワードとなる単語や文章を登録します。あらかじめ設定した20以上の情報ジャンルに関する「アンテナ」と、ユーザ自身が個別に登録可能な「オリジナルアンテナ」により、情報を検索するための適切なキーワードが分からない場合でも、最新のニュース記事やブログ記事の中から、ユーザの求める情報収集が可能となります。

なお、実験期間は、本日から2005年11月7日までを予定しています。

1. 背景と目的

既存のウェブ検索サービスでは、ユーザの検索をサポートするさまざまな取り組みを行っているものの、入力語を含むページを検索結果として提示することから、ユーザが適切なキーワードを思いつかない場合は求める情報を入手できないという課題がありました。

このような課題に対して、NTT研究所が今回開発した日本語概念フィル

タリング技術は、ユーザが入力したキーワードが含まれるかどうかではなく、キーワードや文章が指し示す概念を判別し、ユーザが求める情報をもれなく探すことを可能としました。これにより、ユーザは自分の求める情報に関するキーワードや文章を思いつくままに登録することで、的確な情報をもれなく探し出すことができます。

今回提供する「関心事アンテナ」は、この日本語概念フィルタリング技術を最新のニュース記事やブログ記事から探し出すアンテナとして応用したサービスです。例えば、「芸能人の結婚」のような、必ずしも人名などキーワードが分からない場合でも、そのジャンル内の情報をもれなく探し出して提示してくれる、これまでにないサービスです。

両社は、共同実験を通じ、ウェブ検索サービスを補完する世界的にも先進のサービスとして「関心事アンテナ」の技術面・運用面での実用性検証を行い、併せてこの分野におけるNTTおよび「goo」のプレゼンス向上を図ります。

2. 「関心事アンテナ」の概要

(1) 関心の高いジャンルのアンテナを用意

ニュースやブログの中で特に関心が高い20以上のアンテナをあらかじめ用意しており、これを利用して、「gooニュース」および「gooブログ」掲載の記事の中から各アンテナの指し示すジャンル別に過去1ヶ月分の記事を読覧可能です。記事は、ニュース記事、ブログ記事別に、アンテナが判断する適合度順に並べられます。また、アンテナに関係の深い記事を書いているブロガーを「おすすめブロガー」として表示します。

(2) 個別の関心事は「オリジナルアンテナ」で情報収集

予め用意されているアンテナに加え、関心事をアンテナとして個別に登録することもできます。登録する内容は、関心事についてのキーワードに加え、指し示すジャンルについての日本語の文章でも登録できます。また、「このアンテナで調べる」ボタンをクリックすることによって、登録する前にアンテナの集める記事を確認できます。

(3) 関心度合いによりアンテナを微調整

「オリジナルアンテナ」に複数のキーワードを登録する場合、「詳細設定」ボタンを押して表示されるスライダーにより、キーワード毎のアンテナ感度の強弱を個別に調整できます。また、キーワード毎に用意されたチェックボックスを利用して、そのキーワードを必ず含むように指定することもできます。

(4) 類似記事を探し出すことも可能

収集したニュース記事やブログ記事からピックアップした任意の記事について、内容が類似する記事をポップアップ表示で即座に確認すること

ができます。

3. 技術のポイント

ユーザが入力したキーワードと意味内容が類似した文書を網羅的に情報収集する概念検索手法を取り入れたシステムは、これまでも提案されています。しかし、これまでの概念検索では、単語間の意味内容の類似関係を求めるための辞書が必要であるため、インターネット上の情報のように、辞書に登録されていない未知語が次々と書き込まれるような情報源の検索では、未知語が検索精度を低下させることから適用が難しい状況でした。

今回開発した日本語概念フィルタリング技術は、情報収集対象となる文書中から辞書に登録されていない未知語が検出されると、その文書に含まれる既に辞書に登録されている単語間の意味内容の類似関係から、検出された未知語に関する意味内容の類似関係を自動的に求めて辞書に追加できます。これにより、これまで概念検索が苦手としていたインターネットの検索サービスにも概念検索手法を応用できる可能性が格段に高まりました。また、単語間の意味内容の類似関係を精度よく求める独自手法も合わせて開発し、従来のキーワード検索に比べ、検索結果の網羅性を向上させることが可能になり、適当なキーワードが思い浮かばないようなあいまいな検索の場合でも、もれなく情報を探し出すことができるようになりました。

4. 各社の役割

(1) NTT

本実験サービスに必要なNTT開発の技術を提供します。これにより、商用環境における同機能の有用性および有効性を検証します。

(2) NTTレゾナント

実験終了後の商用サービス導入に向け、利用動向やユーザのご意見を基にビジネス性を評価します。

5. 今後の予定

NTTは、実験の技術的な評価とフィードバックを行うことで、いつでも必要な情報に様々な形で出会うことができるための「サイバー・ナビゲーション技術」の研究開発促進を図ります。また、NTTレゾナントでは、本実験で得られたデータをもとに、情報収集対象となるサービスの拡大や「goo」でのサービス化も視野に入れビジネス性の検討を行う予定です。

・別紙 [「関心事アンテナ」概要](#)

本件に関する問合せ先

NTTサイバーコミュニケーション総合研究所

企画部広報担当 河野・山下

Tel : 046-859-2032

E-mail : ckoho@lab.ntt.co.jp

NTTレゾナント

ポータル事業本部 田畑・栗山

Tel : 03-5224-5500

E-mail : pr@nttr.co.jp

NTT ニュースリリース 