

NTT持株会社ニュースリリース

(ニュースリリース)

平成23年2月21日

日本電信電話株式会社
NTTレゾナント株式会社

文書に3階層のタグを自動付与する「階層的オートタギング技術」の実証実験を開始 ～ボタンのオン/オフ操作で興味のあるQ&A文書に辿り着く「gooラボ QA.ON/OFF」～

日本電信電話株式会社(東京都千代田区、代表取締役社長:三浦 惺、以下「NTT」)と、NTTレゾナント株式会社(東京都港区、代表取締役社長:中嶋 孝夫、以下「NTTレゾナント」)は、ポータルサイト「goo」で提供しているQ&Aサービス「教えて!goo」の質問回答データを分析し、その内容から文書の内容を端的に表す「カテゴリタグ」「主題タグ」「キーワードタグ」という3種類のタグ(内容を端的に表す短い文字列)を自動的に付与することで、目的の文書の絞り込みを容易にする「階層的オートタギング技術」の実証実験「QA.ON/OFF(キューイー オンオフ)」を、実証実験サイト「gooラボ」にて、本日より開始いたします。

1. 「階層的オートタギング技術」開発の背景

多数の文書を整理する際には、文書をカテゴリ毎に整理するという手法が一般的でした。しかし、カテゴリによる整理では、整理する側にとっても複数の話題を扱う文書をどのカテゴリに整理するか判断が難しく、検索する側も意図するカテゴリに目的とする文書が含まれておらず、発見しにくいといった課題がありました。

一方で、ブックマークサービスなどのソーシャルメディアにおいては、ユーザが文書に対して自由にタグを付与して整理する方法が発展してきました。多数のユーザによって総称的なものから具体的なものまで多岐に渡るタグが付与されると、多様な観点で目的とする文書を見つけ出せるようになります。しかし、これまでに蓄積されてきたタグが付与されていない大量の文書に、人手で多様な観点をタグを付与することは現実的ではありません。

本技術は、文章の内容を端的に表す3種類の意味レベルのタグを自動的に付与することにより、カテゴリ分けされた文書に対して多様な観点をタグによる文書整理を容易に実現するものです。

2. 本実証実験の概要

QA.ON/OFF(キューイー オンオフ)」は、ユーザが「教えて!goo」の質問・回答(Q&A)文書を検索する際、ナビゲーションボタンとして、本技術により予め各Q&A文書に付与された3種類のタグ(別紙1☞参照)を活用し、それぞれのボタンをオン/オフして検索内容を絞り込んでいくことで、目的とする文書に辿りつくための手助けをするサービスの実証実験です。(別紙2☞参照)

本実証実験では、階層的オートタギング技術により付与されたタグが、ユーザの検索行動において、目的の文章を検索するための有効な手段になりうるかについて、評価することを目的としています。

3. 「階層的オートタギング技術」の特徴

従来技術は、タグを付与するための学習データとして、ユーザが各文書に付与したタグのデータを必要としましたが、本技術では3つの新要素によりカテゴリ分けされた文書集合のみを学習データとして、高いタギング精度を実現することができます。

(1)各カテゴリで主に扱われる話題を自動発見する「主題語発見機能」(別紙3☞参照)

(2)カテゴリと(1)で発見した主題の集合をタグの候補として、文書中には含まれていない文字列でも、文書内容から統計的

に推定してカテゴリタグ・主題タグの組を付与できる「階層分類タギング機能」(別紙4☐ 上図参照)

(3)入力文書の構造を考慮して、文書中の具体的な文字列をキーワードタグとする「抽出タギング機能」(別紙4☐ 下図参照)

カテゴリタグをもとにしたカテゴリボタンは検索の視点の切り替えに役立ち、主題タグとキーワードタグをもとにしたメインピックボタンとサブピックボタンのオン/オフは検索対象の絞り込み/拡大に役立ちます。

4. 利用方法

「goo」トップ(<http://www.goo.ne.jp/>)から「gooラボ」をクリックし、「QA.ON/OFF」のデモサイトをクリック。トップページに検索キーワードを入力、もしくは注目ワードをクリック。

以下のURLから直接ご利用いただくことも可能です。

URL : <http://qaonoff.labs.goo.ne.jp/>

5. 各社の役割

(1) NTT

本実験に必要な、階層的オートタギング技術を提供します。本実験により、商用環境における同機能の技術的検証を行うとともに、精度向上のためのデータを得ることにより、今後の開発に反映します。

(2) NTTレゾナント

本実験に必要な、Q&Aデータを「教えて！goo」より提供します。本実験により、利用動向やユーザのご意見を基に、ビジネス性の評価・検討を行います。

6. 今後の予定

NTTおよびNTTレゾナントは、本実験により得られたデータを活用し、本技術の実用化を目指すとともに、様々な検索技術の開発と高度化に取り組むことで、お客様にとって、より使いやすいポータルサイトの提供に取り組んでいきます。

別紙・参考資料

- ▶ [別紙1](#)☐
- ▶ [別紙2](#)☐
- ▶ [別紙3](#)☐
- ▶ [別紙4](#)☐

本件に関するお問い合わせ先

■ NTT サイバーコミュニケーション総合研究所 (NTT サイバーソリューション研究所)

企画部広報担当

Tel: 046-859-2032

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

■ NTTレゾナント

広報担当

Tel: 03-6703-6250

E-mail: pr@nttr.co.jp

ニュースリリースに記載している情報は、発表日時点のものです。現時点では、発表日時点での情報と異なる場合がありますので、あらかじめご了承くださいとともに、ご注意をお願いいたします。

NTT持株会社 ニュースリリース

▶ [最新ニュースリリース](#)

▶ [バックナンバー](#)

▶ [English is Here](#)

NTT持株会社 ニュースリリース内検索

1997 ▼ 年 04 ▼

月 ~

2021 ▼ 年 11 ▼ 月

NTTグループの情報は
こちらからもご覧いただけます。



▲ [このページの先頭へ](#)

▶ [更新履歴](#) ▶ [サイトマップ](#) ▶ [お問い合わせ](#) ▶ [著作権](#) ▶ [プライバシーポリシー](#) ▶ [情報セキュリティポリシー](#) ▶ [ウェブアクセシビリティポリシー](#) ▶ [個人情報保護について](#)