

# 構造/テキスト Web データを対象としたハイブリッド問合せ言語の設計

## Design of a Hybrid Query Language for Structured and Text Web Data

学籍番号: 201021747

氏名: 袖山 広輝

Hiroki SODEYAMA

近年、テキストとそれに並存する構造データの 2 つのデータの組で構成された Web データが広く普及しつつある。これまで、これらのデータに問合せを行うためには、構造データを対象とした SPARQL 等の言語に従った構造化問合せ (structured query) を用いるか、テキストページを対象としたブーリアン検索等で利用されるキーワード問合せ (keyword query) を用いるかのいずれかを選択する必要があった。構造化問合せは複雑な条件を指定した高度な問合せが記述可能であるが、相応のスキルを要求するため利用可能なユーザが限定されてしまう。一方、キーワード問合せは単純であるためライトユーザに広く普及しているが、複雑な条件を記述できない。このため、ライトユーザに普及しているが表現力の低いキーワード問合せを利用するか、高度な問合せを行うために一般には難しい構造化問合せ言語を学び利用するか、というオール・オア・ナッシングの選択をせざるをえなかった。この二者択一という状況を緩和できれば、高度な問合せをライトユーザの身近なものにできる可能性がある。

本稿では、キーワード問合せと構造化問合せをシームレスに融合した問合せ言語である Gradation 問合せ言語 (Gradation Query Language. 以下, Gradation とする) を提案する。キーワード問合せのみを利用するようなライトユーザであっても、キーワードに加えて簡単な検索オプションをしばしば利用することに着目し、Gradation を設計した。Gradation では、キーワード問合せをベースとして、簡単な追加記述によってテキストに並存する構造化データに対する問合せ条件を表現することで、キーワード問合せと構造化問合せとのシームレスな融合を実現する。Gradation の特徴は、単にどちらも表現できるというだけでなく、双方の条件指定を混在させられる点にある。これにより、要求やスキルに応じて、単純なキーワード問合せから高度な構造化問合せまでを広くカバーするとともに、“pay-as-you-go” スタイルの問合せを実現する。すなわち、問合せ記述にかかるコストや問合せ条件の精確さを、ユーザが自身の要求やスキルに応じて選択し、利用することを可能にする。

本稿では、Gradation の提案に加え、Gradation による問合せの処理システムの実装について説明する。また、Gradation がリレーショナル完備であることの証明を行う。さらに、pay-as-you-go スタイルの問合せの Gradation による実現可能性を評価するために行った実験の結果について説明する。

研究指導教員: 森嶋 厚行

副研究指導教員: 永森 光晴