



HOKKAIDO UNIVERSITY

Title	セマンティック・ウェブ上のリソースの探索的検索・統合・可視化分析フレームワーク [論文内容及び審査の要旨]
Author(s)	朴, 斌
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第11750号
Issue Date	2015-03-25
Doc URL	https://hdl.handle.net/2115/58785
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Bin_Piao_abstract.pdf, 論文内容の要旨



学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称 博士（情報科学） 氏名 朴 斌

学 位 論 文 題 名

セマンティック・ウェブ上のリソースの探索的検索・統合・可視化分析フレームワーク

(Interactive Framework for Exploratory Search, Integration, and Visual Analysis of Semantic Web Resources)

近年, 各分野におけるセマンティック・ウェブ技術の普及が進むことによって, RDF(Resource Description Framework)に基づいて記述されたデータやサービスの量が急激に増加することになり, 今では, DBpedia をはじめ様々な分野の約 1000 を超える RDF データセットがウェブ上に公開されるようになった. これらの RDF データセットは RDF クエリ言語である SPARQL(SPARQL Protocol and RDF Query Language)を用いて問い合わせることができる. これに併い, 意味構造を持つこれらの大量データやサービスに対し, ただアクセスしてブラウジングするだけでなく, 様々な種類のデータやサービスを統合利用して興味深いデータ集合やそれらの間の関係を検索・探索・分析しながら, 新たな知見を発見したいという要求が高まってきている.

セマンティック・ウェブ上に集積された膨大なデータ集合や多種のサービスは, 多様なデータを組み合わせることで, データの創発的な価値を生み出す可能性を高める一方, データ集合やサービス自体の理解を困難にしており, その多様な活用を難しくしている. セマンティック・ウェブを有効に扱うためには, セマンティック・ウェブ上のデータやサービスを自在に検索, 統合, 分析できることが重要である. しかし, それを可能にするには以下の 4 つの問題点を解決しなければならない.

(1) 検索要求定義や分析を行うためのシナリオの想定・作成が困難

セマンティック・ウェブ上の情報を利用して実世界の課題を解決する際, 例えば「転勤先での 3 世帯同居可能な新しい家を探したい」のように明示されている条件だけでは課題の解決に至らず, 通勤時間や, 家賃, または家周辺にある教育施設, 医療施設等のような様々な潜在的条件をも一緒に考慮し, 分析する必要があるような曖昧な課題は, それを解決するのに必要とするあらゆる条件を予め想定した上で, 適切な検索要求の定義や分析を行うためのシナリオを作成するのは困難である.

(2) 複雑な SPARQL 検索要求文のインタラクティブな定義が困難

適切な SPARQL 検索要求文を作成するには, 検索対象 RDF データのスキーマ, つまり, どのようなクラスが存在し, それらはどのような意味的關係によって関連づけられているのかという構造を事前に把握し, それに基づいて検索要求文を作成する必要があるが, 多くの RDF データセットは固定したスキーマを持たず, 巨大化しつつある各分野の RDF データに対して, そのスキーマの全体像を把握することは困難になりつつある. また, スキーマを知っていたとしても, SPARQL に関する深い知識を持たない一般ユーザには, 論理演算を含む複雑な SPARQL 検索要求文を正確に素早く作成することは困難である.

(3) 多様なデータやサービスの組み合わせ統合が困難

セマンティック・ウェブにおける多様なデータやサービスの組み合わせ統合には, 以下の 4 つの場合がある.

- a) 同一 RDF データセット中での多様なデータの組み合わせ統合
- b) 異なる RDF データセット間の多様なデータの組み合わせ統合
- c) 異なるサービス間の多様なデータの組み合わせ統合
- d) RDF データセットとサービス間の多様なデータの組み合わせ統合

ここで, a) と b) の場合は統合しようとするデータ集合のスキーマに関する知識が必要となり, 特に b) の場合, 使用されている語彙も異なることが多いことから, 異なるデータ集合の組み合わせ統合は困難である. また, c) と d) の場合は, サービスに関する知識や高いプログラミング能力が必要となる.

(4) 複雑な構造をもつ膨大な RDF データに対する分析が困難

膨大なデータの分析において、多様な視点から探索的に可視化分析を行う連携多重ビュー (Coordinated Multiple View) に関する研究が多くなされてきた。しかし、スキーマの全体像を事前に予想できない RDF データに対しては、適切な連携多重ビューを構築し、分析を行うことは困難である。

本論文では、上記の 4 つの問題点を解決するためにセマンティック・ウェブ上のリソースを探索的に検索、統合、可視化分析できる ESIASW(Interactive Framework for Exploratory Search, Integration, and Visual Analysis of Semantic Web Resources) フレームワークを提案した。ESIASW フレームワークは、機能を持つ視覚的部品に対する直接操作を用いて、セマンティック・ウェブ上のリソースに対する検索、統合、分析を探索的に繰り返すことを可能にし、膨大なデータ集合やサービス自体に対する理解を深めていきながら、課題を解決するためのシナリオそのものを発見し、新たな知見発見に至る全過程を統括的に支援する枠組みである。

セマンティック・ウェブ上のリソースの検索において、ESIASW フレームワークでは、視覚的部品に対する直接操作によって、多視点から探索的に RDF データのスキーマを抽出・展開・可視化することを可能にしている。これにより、ユーザは検索要求定義に必要なスキーマを、関心のある 1 つまたは複数のクラスをアクセス始点として、それと何らかの意味的関係を持つクラスを探索的に逐次抽出することによって、必要に応じて部分的かつ探索的に展開して利用することができる。展開・可視化したスキーマを直接操作することによって、複数の異なる視点からインタラクティブに検索条件を同時に指定したり、変更したりすることを可能にする。また、視覚的部品に対する直接操作によるクラス間の論理演算を可能にすることで、必要なスキーマをユーザ自身が新しく定義し、論理演算を含む複雑な検索要求を定義することを可能にしている。

セマンティック・ウェブ上のリソースの統合において、ESIASW フレームワークでは、同一作業空間内で、異なる複数の RDF データのスキーマを展開・可視化したり、必要なセマンティック・ウェブ・サービスを呼び出したりすることを可能にしている。ユーザは、展開・可視化された RDF データスキーマを意味的関連や論理的関連に基づいて 1 つのスキーマに統合して利用したり、異なるセマンティック・ウェブ・サービスを意味的関連に基づいて直接結合したり、RDF データスキーマとセマンティック・ウェブ・サービスを意味的関連に基づいて直接結合することで、多様なデータやサービスを自在に組み合わせ統合することができる。

セマンティック・ウェブ上のリソースの分析において、ESIASW フレームワークでは、RDF データのスキーマを探索的に抽出・展開しながら、それに基づいて連携多重ビューを動的に構築することを可能にする。ユーザは構築された連携多重ビューの上で複数のデータ集合をインタラクティブに選択することができる。ESIASW フレームワークでは、連携多重ビュー上の各選択操作に独自の色を割り振り、その結果を Brushing&Linking 技術と色の組み合わせ技術を用いて表すことによって、異なるデータ集合を相互比較したり、それらの間の相関関係を探し出したりすることを可能にする。

まず、最初の章で、セマンティック・ウェブ上のリソースの検索、統合、分析に関する背景や問題点を述べた後、本論文の貢献について述べている。

第 2 章では、セマンティック・ウェブとそれを支える基盤技術について述べた後、インタラクティブな可視化環境を構築するのに用いる知識メディアアーキテクチャとそれに基づいて開発されたシステム Webble World について述べている。

第 3 章では、RDF データモデルとその拡張について述べた後、提案する ESIASW フレームワークの概要と ESIASW フレームワークが提供する各視覚的部品について述べている。

第 4 章では、ESIASW フレームワークを用いたセマンティック・ウェブ上のリソースの探索的検索について述べている。

第 5 章では、ESIASW フレームワークを用いたセマンティック・ウェブ上のリソースの統合について述べている。

第 6 章では、ESIASW フレームワークを用いたセマンティック・ウェブ上のリソースの探索的可視化分析について述べている。

第 7 章では、関連研究とその問題点を明らかにし、提案する ESIASW フレームワークのユーザビリティに関して既存の関連研究と比較しながら評価を行っている。