学位論文題名

A Unified Framework For Organizing, Accessing, and Federating Web Resources

(ウェブリソースの組織化, アクセス, 連携統合のための
統一的フレームワークに関する研究)

学位論文内容の要旨

近年のコンピュータ技術とネットワーク技術の発展により, ワールド・ワイド・ウェブ上のリソースは種類も量もますます急速に増加し続けている。これらのウェブリソースには, 文書やデータだけでなく, アプリケーション・ツールやサービスなどが埋め込まれた複合文書などが含まれている。本論文第1章に述べているように, このような状況下において, 目的の情報に効率よくアクセスするためには, 用途に合わせて情報を組織化し, 組み合わせる新たな関係を定義し, それらの関係をウェブ上に保存するための枠組みを提供する必要がある。しかしながら, 表形式, 階層構造, 索引付けといった従来の情報組織化手法では, ウェブリソースのように多種多様なリソースの管理やアクセスには不十分である。

また, ウェブリソースの高度な再利用のためには, 適切なウェブリソースの発見, それらのリソースの相互運用と協調連携を支援する必要がある。セマンティック・ウェブ技術やウェブサービス技術は, あらかじめ決められた範囲におけるウェブリソース連携統合を支援するが, エンドユーザがアドホックにサービスを発見し, それらの間に連携を定義することは支援しない。

本論文では, エンドユーザによるアドホックなウェブ上の情報リソースの組織化, アクセス, 統合のための統一的な環境に関する基盤技術を確認した。具体的には, マルチメディア・複合文書を, 様々な情報リソースを保存し管理するための「地」として用い, その地文書において, 直接操作によって複数のリソースを関連付け, 再構成することを可能とした。

はじめに, 第2章では本研究の基盤となる既知の要素技術について解説し, 第3章では, 「トピカモデル」と呼ばれる, 様々なコンピュータ内のリソースを表現するための新たなデータ・モデルを提案している。トピカモデルは単純で厳密な構造を要求しないハイパー・テキストモデルに, 厳密な関係構造を要求する関係データ・モデルを統合した中間的な情報管理と情報アクセスのモデルである。トピカモデルは, 多種多様な情報リソース間に定義された関係に基づくアクセス手法を提供する。第4章では, 既存のHTMLによって記述されたウェブ文書を, トピカモデルを用いて表現することにより, URLだけでは指定できないウェブアプリケーションの構成要素や, ナビゲーションによってたどり着くことができるリソースを識別することが可能になる。

次に第5章では, ウェブ文書の一部を再構成し, 目的の情報へのアクセスを自動化するため, 「C3W(Clipping, Connecting, and Cloning on the Web)」と呼ばれるフレームワークを提案する。本枠組みでは, ウェブ文書の一部をセルとして切り抜き, 新たな複合文書を作成し, それらの切り抜きの間にスプレッドシートの式を入力することによって新たなデータ連携を定義する仕組みを確認した。切り抜かれたセルは, 切り抜き元のウェブ文書における入出力関係を保ったまま切り抜かれるため, 出力用のセルには, 入力用セルの値に応じた結果が表示される。複数の異なる探索結果を比較するために, 切り抜かれたセルを複製するための枠組みも提供している。本枠組みを用いて, ワープロ文書やスプレッドシートといったオフィス文書を, これらの切り抜きの埋め込み先として用いる新

技術についても述べ、それらの文書と切り抜かれたセルとの機能連携について述べる。この新技術により、構築されたカスタム複合文書の更なる再利用と流通が可能となる。

第6章では、ウェブリソースの組織化、アクセス、統合の統一的枠組みとして「関係アノテーション・フレームワーク」を提案している。この枠組みでは、すでにウェブ上に存在する様々なウェブ文書をウェブリソースの組織化、アクセス、統合のための「地」として働く空間として用いる。ユーザは、地空間として用いられる任意のウェブ文書において、この文書中の一部分を任意に複数個選び、ウェブリソースを格納するための「場」として働くように定義することができる。これらの場の間には関係が定義され、その関係を用いて様々なウェブリソースを関連付けて保存することができる。リテラル値や URI、さらにウェブ文書の構成要素がこれらの場に格納可能である。格納された関係とリソースは、その「地」文書の注釈(アノテーション)として保存される。これにより、元の「地」文書を変更することなく、様々なウェブリソースを関連付けることが可能となる。ユーザはこれらの関係を利用して、情報探索を行うことが可能となる。

本論文で提供される枠組みを用いて、ユーザは細分化された様々なウェブリソース間に新たな関係を定義し、その関係をウェブ文書上に保存することが可能となる。これにより、エンドユーザによる直接操作を用いた、アドホックなウェブリソースの組織化、アクセス、連携統合が可能となる。

学位論文審査の要旨

主 査 教 授 田 中 讓
副 査 教 授 原 口 誠
副 査 教 授 栗 原 正 仁

学 位 論 文 題 名

A Unified Framework For Organizing, Accessing, and Federating Web Resources

(ウェブリソースの組織化, アクセス, 連携統合のための
統一的フレームワークに関する研究)

近年、ウェブの発達により、膨大な数の多様な情報やサービスが知的リソースとしてウェブ上に出版され蓄積されているが、これらの情報やサービスを、各人が種々の観点のもとに、相互に関連するもの同士をまとめて新たに組織化し、他者が組織化したものも含めて自在にアクセスできるようにウェブ上で再構成するための技術は未発達である。サービスを含むリソースの相互関連付けは、単に関係を定義するにとどまらず、各サービスが提供する機能を連携統合することができなくてはならない。オープンな環境で多様なリソースが公開され、その高度な再利用が望まれている現在、このようなウェブ上のリソースの組織化、アクセス、連携統合の技術の研究開発と確立が望まれている。

本論文は、このような現状にあるウェブ上のオープンリソースの世界において、本学で開発されてきた知識メディア技術を基盤に、これらのリソースの自在な組織化、アクセス、連携統合のための統合技術を新たに提案し研究開発したものである。

まず、2章において、ウェブリソースとそれらの間の意味関係を扱うための基盤技術のサーベイを行い、3章と4章において、トピカと名づけた組織化とアクセスのためのアーキテクチャを提案している。トピカは、相互に関連するリソースの同種の n 項組が多数存在するとき、これを関係表として既存のウェブ文書に埋め込むことによって、関係表に基づくリソース間の連想検索を支援すると共に、この関係表自体の検索探索を、それが埋め込まれているウェブ文書に対するナビゲーションやサーチエンジンによるキーワード検索を用いて支援しようとする試みである。関係表の管理検索は従来はデータベースによって行われてきたが、ここで想定しているような環境では、リソース間の関係表は多数の利用者によって自在に動的に定義され、その種類は多様で、数も限定することができない。このような場合には、データベースのスキーマを予め定義することができず、多様な多数の関係表の中から目的の関係表を検索探索すること自体が大きな問題となる。このような場合には現在のデータベース技術は無力である。本論文では、このような関係表を構成する各属性を、既存の特定のウェブ文書中の、分節化されているマルチメディアコンテンツやテキスト項目に対応付け、これらの分節化項目をトポスと名づけ、トポス間の関係と、属性間の関係を対応付けることにより、関係表をこのウェブ文書に埋め込むことを提案している。利用者は、トポスを介して属性地にアクセスでき、特定のトポスに対して制約をかけることによって関係表に制約をかけ、他のトポスを介してアクセスされる属性値を制限することができる。この関係表のアクセスは、この関係表の文脈となるウェブページをサーチエンジンを用いて検索したり、ページ間のナビゲ-

ションを用いて探索することにより、支援することができる。適切な既存のウェブページを用いるなり、新たにウェブページを作成することにより、異なる利用者間で文脈を共有することもある程度可能である。3章では、トピカモデルを提案し、4章でトピカモデルをウェブ技術を用いて実現するための方法論を述べている。

5章では、ウェブリソースの一部を相互の関連性を保ちながら、オブジェクトとして切り出す技術として、C3Wと名づけられた知識メディア技術の新しいアプリケーションフレームワークを提案している。C3Wでは、ウェブ上の情報ならびにサービスを区別することなく、ウェブ文書中の任意のマルチメディアコンテンツやテキスト項目、ウェブアプリケーションの入力項目や出力の一部を、各々ウェブクリップと呼ばれるオブジェクトとして、もとのウェブページやウェブアプリケーションにおける相互の入出力関係を保持しながら複数個切り出すことができる。利用者は、特殊なブラウザ上でマウス操作のみを用いて、任意の分節化されているオブジェクトを切り出すことができる。一連のナビゲーションの中で切り出されたウェブクリップは、相互の入出力の依存関係を保持するので、切り出された後も、入力項目を更新すると、これに依存するウェブクリップが自動的に更新される。さらに、2つの独立なナビゲーションから切り出されたウェブクリップの n 項組の間に、一方の入力項目ウェブクリップの値が、他方の出力項目ウェブクリップの値に等しくなるよう制約を定義することにより、これらの2組の n 項組を連携統合することができる。

6章では、先行する章において議論されたアーキテクチャと技術を基盤にし、これらを統合して、C3Wを用いて切り出されたウェブクリップ群をリソースとして、任意のウェブページ上にトポス群を定義し、そこに関係表としてこれらのリソース群を埋め込むための関係アノテーションアーキテクチャを統合アーキテクチャとして提案している。このアーキテクチャは、ウェブに対するアノテーションシステムを関係アノテーションに拡張しているとともに、アノテーションの内容として、ウェブ上の任意のリソースを相互の関連性、依存性を保ちながら埋め込むことができる点に新規性があり、トピカモデルをウェブ上で実現することにより、ウェブ上のオープンリソースの組織化、アクセス、連携統合に対する統一的支援フレームワークを実現するものである。

これを要するに、著者は、ウェブリソースの組織化、アクセス、連携統合に関する新しい統一的フレームワークに関して新知見を得たものであり、ウェブ工学、知識メディア工学に対して貢献するところ大なるものがある。よって著者は、北海道大学博士(工学)の学位を授与される資格あるものと認める。