

Web API ドキュメントからの情報抽出による  
プログラムライブラリ作成支援  
Information Extraction from Web-API Documents  
for Supporting Generation of Web-Programming Libraries

学籍番号：201121722

氏名：高井 正成

Masanari TAKAI

WebアプリケーションがWeb APIという形式で他のアプリケーションに対してその機能を提供する事例が増加している。Web APIリポジトリであるProgrammable Webには現在8000件以上のWeb APIが登録されており、Web APIの登録数は年々増加している。

Web APIは各々に互換性がなく、Webアプリケーション開発者はWeb APIを利用する場合、それぞれの仕様に沿ったプログラムライブラリを作成する必要がある。また、新たなWeb APIが公開された場合複数の開発者がそれを利用しようとして、複数の同様の機能を持ったライブラリが作成されてしまう。Web APIの仕様が変更された場合は、ライブラリを修正しなければならないが、修正漏れがあったライブラリは一部もしくは全てが利用不可能になってしまう可能性がある。以上のような理由からWeb APIのライブラリを開発者が作成するのは非効率であるといえる。そこで本研究では開発者がライブラリを自分で作成する手間を省くため、Web APIのオンラインドキュメントからライブラリの作成に必要な情報を機械的に抽出しライブラリを生成する手法を提案する。抽出にはWeb APIのオンラインドキュメントに見られる規則性を利用する。これまで開発者が自分で作成していたライブラリを機械的に生成することにより、開発者のライブラリ作成にかかる手間を省く。

本研究では実際に提案手法を実装したプロトタイプシステムを作成し評価実験を行った。Programmable Webに登録されているオンラインドキュメント30件を対象とし、そのオンラインドキュメントをドキュメントの記述方法の違いからタイプ1と2に分類した。タイプ1はWeb APIを利用する際重要な要素が提示される順番が決まっており、表や箇条書きによりある程度情報が整形して記述されているもので、タイプ2はWeb APIの1つの機能に関する情報が複数のWebページに記述され、各情報は表や箇条書きの形式は用いられず、整形されていない文章内で説明されるものである。評価実験の結果として、タイプ1のオンラインドキュメントに対して十分な抽出精度が得られることを示した。またプロトタイプシステムを用いないライブラリの開発よりも用いた開発の方が、タイプ1において作業量が少なくなり、以上で述べたライブラリ作成時の諸問題を解消するのに有効であることを示した。

研究指導教員：阪口 哲男

副研究指導教員：森嶋 厚行