

メタデータスキーマ作成のための
メタデータ語彙探索支援システムの構築
A Retrieval Support System for Metadata Terms
to Assist the Metadata Schema Design Process

学籍番号：201321643

氏名：西出 頼継

Yoritsugu NISHIDE

Linked Open Data(LOD)が普及したことにより、分野を越えたメタデータの利用が頻繁に行われるようになった。LOD として公開されるメタデータの利用を支援するためには、メタデータの構造や制約といったメタデータスキーマを作成してメタデータと併せて公開することが求められる。メタデータスキーマを作成する際は、既存のメタデータ語彙で定義されたクラスやプロパティなどのタームを再利用することが望ましい。しかし、メタデータ語彙の数は非常に多く、各タームの定義を調べ適切なタームを選択する作業は大きな負担となる。Linked Open Vocabularies(LOV)等のメタデータ語彙探索支援システムは存在するものの、探索の条件として扱える属性は少なく、キーワード以外に、例えば記述対象に関する情報を用いた探索などは行えない。また、既存手法で扱える探索条件だけでは、探索結果に含まれる多くのタームに対して定義や用途を調べていく優先度を、探索要求を考慮して提示することが難しく探索に時間が掛かる。そこで本研究では、探索条件に一致するタームの発見とその選択支援を目的とした、メタデータ語彙探索支援手法を提案する。

本手法では、既存のシステムで実現されているキーワード探索に加え、記述対象に関する情報と、共起するプロパティなどのタームを条件とする探索を実現する。また、各探索条件を考慮したタームのランキングを実現する。本手法の実現のために、記述対象のクラス 455 件の分類と、既存メタデータインスタンス 12,941,323 件の分析、ランキングのためのスコア算出式の定義を行った。さらに、本手法を実装したメタデータ語彙探索支援システムを構築した。本手法と既存手法とで行った比較評価実験では、本手法の方が探索条件に適したタームを推薦できていることが分かった。また、この実験から入力するキーワード選定の難しさが問題になっていることが分かり、キーワード検索に対して事前に候補を分類しておくか、類似キーワードを推定しておくことで入力を支援できると考察した。

研究指導教員：杉本 重雄

副研究指導教員：永森 光晴