

図書館情報専門職養成機関に関する資料目録の Linked Open Data 化

林 由莉亜

筑波大学情報学群知識情報・図書館学類は、図書館司書や図書館に関する専門的な知識を要する専門職を養成する機関(以下、図書館情報専門職養成機関)のひとつである。筑波大学図書館情報メディア系の教員らが、「21世紀図書館情報専門職養成研究基盤アーカイブ」の構築に着手し、その過程で知識情報・図書館学類の前身校の資料の目録を作成した。本研究はこの資料目録を対象とする。

この資料目録を、機械が読解・処理可能なデータである Linked Open Data (以下、LOD)化させる試行が過去に行われ、LOD 化の要件の一部を満たした資料目録が作成された。しかし、この段階では多くの課題が残されていた。これを今後の図書館情報専門職養成機関に関する研究を行う上で、例えば外部の研究者が図書館情報専門職養成機関の歴史について調査する際など、有効なデータとして活用が可能な LOD にすることが本研究の目的である。

資料目録を LOD 化し活用可能にするために、残されている課題や、リンクの起点となり得る有効な情報の構造を分析した。分析の結果、LOD における情報の単位であるトリプルの、その属性を記述する語(以下、述語)に、一般的に利用されている語彙とは別に独自に定義した語彙のみを使用している点、リンクの起点になると考え難い情報が、キーワードとして記述されている点が挙げられた。

独自に定義した語彙のみを使用している点について、独自語彙の内、意味が同一であり置き換え可能な述語は一般的な述語に置き換えた。意味が同一と考えられる述語は他のデータセットにも用いられている一般的な述語を利用することで、トリプルの内容が人にとって分かりやすくなり、また、複数の LOD データの集合(以下、データセット)から同時に同一の意味を持つトリプルを検索することが出来るようになる。同一の意味の述語が見つけれられないものや、一般的な語彙では表現出来ない意味の述語は、独自に定義した語彙のままとした。LOD 化の試行を行う前の資料目録には記述のない、リンクの起点になると考え難い情報が記述されている点に関しては、それらの削除を行った。

以上の修正を行った 57,692 件のトリプルからなる資料目録が、LOD として十分に活用可能かどうかを評価するため、DBpedia を含め 4 件のデータセットを用い、リソースや文字列を用いてリンクすることが可能かを評価した。結果、人名、組織名、施設名といった固有名詞を 152 件確認した中で、「つくば市立図書館」のように完全一致した文字列や、「国文学研究資料館資料」中の「国文学研究資料館」のように部分一致した文字列を属性に持つリソースから、他のデータセットの該当固有名詞の項目へ 123 件リンクを行うことが出来た。以上のように、今後の図書館情報専門職養成機関に関する研究を行う上で、有効なデータとして活用が可能な資料目録の LOD にすることができた。

(指導教員 阪口哲男)