

Ontologyを用いた非文字資料の検索

木下研究室

羽生 敏英 (201002748)

1 まえがき

非文字資料とは、文字媒体として記録することなく受け継がれてきた感覚や経験といった文化に関する民族学的資料を指す。本稿では「只見町インターネット・エコミュージアム」(以下、只見町 IE)に記載された民具の一部を対象とし、民具データベース(以下、民具 DB)を構築していく。

非文字資料は、同名の民具であっても採取地域等によって差異が出てくる。同様に異なる名前の民具であっても実際には同じものを指している場合がある。よって非文字資料を取り扱う際には、単純に民具名だけでなく登録番号によって区別する必要がある。そうした個別の民具に注目することで、個々の民具の持つ差異を検索キーに対して、より正確に検索結果へ反映することができる。そのためにメタデータによって紐付された意味検索が求められる。

本稿では、Ontology 構築支援ツールの1つである Protege 4.3 を用いて民具 DB を構築する。図1で示すように入力された検索キーに対して、民具 DB による Ontology の他に、例えば Wikipedia Ontology のような既存の Ontology を組み合わせることで「只見町 IE」の検索性の向上を考える。

2 提案手法

民具 DB では、まず大分類、中分類、小分類の3段階の階層構造を基本構成としている。これまでの研究では、民具クラスの1つ下のクラスとして分類番号が位置していたため、それがどの民具を指しているのか視覚的に分かりづらく、また新しく民具情報を追加する際に無駄な部分が多くあった。

今回の民具 DB では、図2のように「ネコミノ」という民具が大分類に生活用具、中分類に衣類、小分類に防寒具に分類され、「ネコ」という民具の別名であることを示している。

また各民具に対しては、使用目的、採取地域、材料についての情報をネットワーク構造的に関連付けている。図3では登録番号「A3903」という民具は「ネコ」という名称の防寒具で、布とワラを材料とし、叶津という地域で山樵や施肥といった作業に使用されていたことがわかる。これらの情報はメタデータによって色分けされ、各民具の持つ関係性を視覚的に示している。

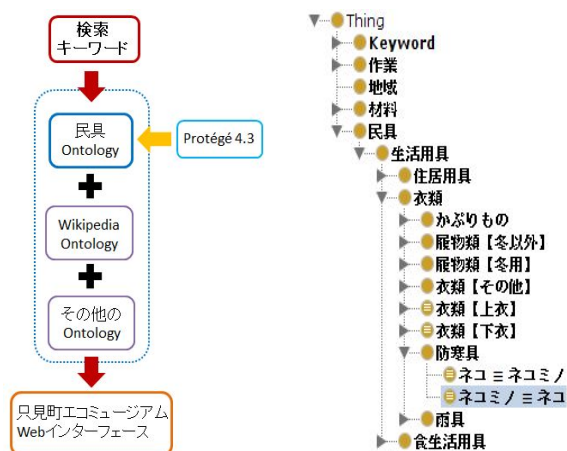


図 1: 只見町 IE の改善案 図 2: 民具 DB の構造

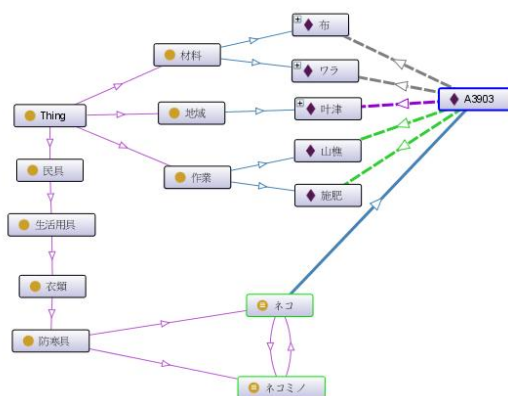


図 3: 民具の関係

3 結果

各民具に対して、個別に概念を設定することによって検索キーの増加を行うことができた。例えば、「只見町 IE」では、検索キーに「服」あるいは「仕事着」と入力しても該当する検索結果を返すことはなかったが、民具 DB では、検索キー「服」に対し衣類【上衣】、衣類【下衣】に属するものを、検索キー「仕事着」に対しては使用目的に「仕事着」を含むものをそれぞれ検索結果として返すことができた。

民具 DB では階層構造とネットワーク構造を組み合わせ、民具資料の体系化を目指した。そして今回、民具 DB の雛形となるものを構築することができた。しかし問題点として情報追加の煩雑さが残されている。今後の課題として、異なる Ontology との併合による語彙情報の増加や推論による民具情報の自動的な追加が挙げられる。