

農業における Wiki を活用する知の構造化

プロジェクトマネジメントコース・ソフトウェア開発管理グループ 矢吹研究室 1242034 氏名 小池 克人

1. 序論

平成 27 年 3 月 10 日に発表された農業情報の標準化に関する個別ガイドラインでは、農業の情報の相互運用性を確保するインターオペラビティーとポータビリティ確保の標準化が必要とされている [1]。

しかし、異なる企業や団体の意思を一つにまとめる作業には困難が伴う。標準化したデータをやり取りするプロトコルやデータ形式は各企業の利害がぶつかるため、困難である [2]。故に複数のシステムの間でマスターを統一しようとする共通語彙は、目的と実現する考え方のアーキテクチャーが異なる。マスターは構造も用語も異なるため標準化することができない。今回は共通的な方法論に統一することができないかを考える。

作物名称は、研究や行政などにより変わる。研究は研究目的に応じた分類のため、行政は行政の目的に応じた分類のため、などの用途により視点と用語が異なる。

今回の研究では、異なる作物名称を目的に応じた用語に変換（翻訳）するための、共通する仕組みがないかということに着目した。用語の変換をするために用語を上位下位で結び、関連用語を抽出することを考える。関連用語の抽出をしている事例を探し、参考にして研究する。そして、自然言語処理に Wikipedia 上位下位関係抽出の論文を参考にして、Mediawiki を利用する [3]。MediaWiki は、Wikipedia に使用されているフリーソフトで自分の Wiki を自分のサーバーに設置することができる。そのため MediaWiki を利用し、目的に応じて最適な語彙の翻訳を可能とする翻訳システムの開発を研究する。

2. 目的

本研究では、農業情報の用途により用語が異なる語彙を目的に応じた最適な語彙への翻訳ができるような仕組みを作ることが目的である。

3. 手法

本研究では、MediaWiki を用いて知識を登録する。登録した情報から、上位下位関係抽出ツールを用いて単語間の関連情報を抽出をする。抽出した情報を使って、用語の翻訳を試みる。

4. 結果

Mediawiki に様々な書き方をした解析することで成功するパターンを発見することに成功した。カテゴリに出荷統計作物名などを記載して、記事の見出しのレベルを上げて書くことで目標とする解析結果を表示できた。

5. 考察

成功パターンを発見できたため、結果に記載したように Wiki を編集すれば、Wiki を編集したことがない人でも容易に編集することができ、農業情報の標準化できると考えられる。

6. 結論

本研究では、上位下位関係抽出ツールを用いて、関連情報を抽出することに成功した。どのように運用していくかが今後の課題である。

参考文献

- [1] 内閣官房情報通信技術 (IT) 総合戦略室. 農業情報の標準化に関する個別ガイドライン等について. https://www.kantei.go.jp/jp/singi/it2/senmon_bunka/nougyou_guideline/siryou1.pdf.
- [2] 井原敏宏. 第 2 段階に入る農業 IT、複数企業・組織が連携して標準化や協業を推進, 2015. <http://itpro.nikkeibp.co.jp/atcl/column/14/346926/032900208/?ST=system>.
- [3] 隅田飛鳥, 吉永直樹, 鳥澤健太郎. 自然言語処理による wikipedia からの上位・下位関係抽出. 自然言語処理, Vol. 16, No. 3, pp. 3–24, 2009.