

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. 研究の背景

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。標準化して、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]。よって、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には、目的と実現する考え方のアーキテクチャーが異なり、マスターは構造も用語も異なるため現実感が無いため、共通的な方法論ができないかを考える。

作物名称は、研究、行政、流通、農家により変わる。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、流通は流通に都合の良いネーミングのため、農家は営農の都合に良い分類のための用途により視点が異なる。よって、用語も異なる。

これを目的に応じた用語の変換（翻訳）を実現する共通な仕組みができないかに着目する。用語の変換をするために用語をタグで結び、関連用語を抽出することを考える。これと同じような事例を探し、参考として研究する。そして、2008 年の Wikipedia の記事構造からの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、オープンソース (GPL) で配布されているため、MediaWiki を用いれば、自分専用の Wikipedia を設置・運用することができる [4]。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

農業情報の用途により視点、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. 研究方法

研究方法は、MediaWiki のサーバーを立ち上げる、MediaWiki に知識を登録をする、登録した情報から、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。

以上の 4 つの方法で研究する。

4. 成果物のイメージ

類似名称のある作物や意味が同じ作物固有の名称がある作物など目的に応じて最適な語彙の翻訳を可能とすることで、用語と世界をつなぐことができる。

5. 進捗状況

現在の進捗状況は、検索システムのプロトタイプの開発のための準備をしている。

6. 今後の計画

今後は、MediaWiki を立ち上げ、データを入力してデータを抽出できるかを実験する。

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・下位関係抽出の論文, 2008. https://www.jstage.jst.go.jp/article/jnlp/16/3/16_3_3_3/_pdf.
- [4] ニコニコ大百科. MediaWiki とは. <http://dic.nicovideo.jp/a/mediawiki>.