

データ連携のためのオントロジーを用いた農業分野の用語標準化

Standardization of Agricultural Vocabulary for Data Federation based on Ontology

朱 成敏^{*1} 武田 英明^{*1} 法隆 大輔^{*2} 竹崎 あかね^{*2} 吉田 智一^{*2}
Sungmin JOO Hideaki TAKEDA Daisuke HORYU Akane TAKEZAKI Tomokazu YOSHIDA

^{*1}国立情報学研究所
The Japanese Society for Artificial Intelligence

^{*2}国立研究開発法人農研機構中央農業総合研究センター
National Agriculture and Food Research Organization

In recent years, the spread of agricultural IT systems has been accompanied by generating various data. These agricultural data is expected to optimization of process and prediction of yield in agricultural field. To analyze these data, data sharing and integration among agricultural IT systems is essential. However, since there is a lot of multi-meaning vocabulary and no guideline for vocabulary in agriculture field, each of the agricultural IT systems has been operated based on its own vocabulary. In this paper, we propose using ontology as core vocabulary for agriculture that becomes the guidelines for data federation between agricultural IT systems.

1. はじめに

近年、日本の農業分野における IT 技術は作物の栽培から営農管理、流通に渡り幅広く応用されている。また、様々な IT システムが農業現場に普及され、今後も高齢化による農家の離農や作物生産の効率性のために IT システムの導入とその利活用が増えていくと予想される [内閣府 15a]。また農業 IT システムからは様々なデータが発生されている。これらのデータを蓄積し、解析することで農作業の適切化、収穫量の予測が可能となり、農家のコストを抑える共に、農作物の収量と品質の向上も期待される。

しかし、データを利活用するためには IT システム間のデータの連携が必要となるが、データ連携において重要な基準となる用語情報が標準化されていないため、IT ベンダーはそれぞれ独自の定義を行い、農業 IT システムに導入している。こういった背景から内閣府 IT 総合戦略本部は農業分野における情報創成・流通の促進のために農業情報の相互運用性等の確保に資する標準化や情報の取扱いに関する政府横断的な戦略を策定した [内閣府 15b]。

そこで、本稿ではオントロジーに基いた農業分野の用語標準化を提案する。まず、対象を農作業とし、要求される言語的特徴を考察する。それに基づいて構築したオントロジーを用いて語彙リストを作成し、用語標準化の手段としてのオントロジーについて考察する。

2. 背景

本章では農業分野における IT 技術について概観し、用語標準化の必要性と問題点について述べる。

2.1 農業分野における IT システム

現在、農業分野において注目される IT 技術は GPS、センサーネットワーク、ロボット、POS など多くの分野にわたる。また、作物の生産、収穫など営農管理の記録する管理システムも幅広く普及されている。農業分野に IT 技術を導入することで生産の効率化、生産性と品質の向上を確保することができ、農業全体の価値の上昇も期待される。そして、これらの IT システムからは様々なデータが発生されており、データを解析す

ることによって農家は作業の最適化や収穫量の予測、営農計画など様々な応用が可能となる。また、データを統合・連携し、分析することによって地域、または国の農業全体における状況を把握することができ、気象条件など環境要因における対策や生産政策などを決定することが容易となる。

日本では農業就業人口の減少によって 1 農家当りの農地面積が増加している。また、農業従事者の高齢化が進めており、農家の生産所得も低下している^{*1}。こういった背景から生産の効率化のために様々な農業 IT システムの開発と普及が行われている。代表的な農業 IT システムでは富士通の Akisai^{*2}、日立ソリューションズの GeoMation Farm^{*3} などがおり、現在営農管理に係る IT システムだけで 38 システムが導入されている。また、今後も様々な農業分野において IT システムの開発と普及が予想される。

2.2 データ連携における問題点

様々な農業 IT システムから発生したデータを統合・連携するためには基準となるデータ項目が必要である。基準となるデータ項目を中心にデータを収集し繋げることでデータの統合と連携が可能となる。しかし、それぞれの農業 IT システムはベンダーが独自で定義したデータ項目に基づいて開発されたため異なるシステム間のデータ統合が困難である。こういった背景から平成 27 年に内閣府 IT 総合戦略本部は「農業情報の標準化に関する個別ガイドライン」として「農業 IT システムで用いる農作業の名称に関する個別ガイドライン (試行版)(案)」[内閣府 15c] と「農業 IT システムで用いる環境情報のデータ項目に関する個別ガイドライン (試行版)(案)」[内閣府 15d] を提示した。

温度や降水量など国際基準が明確な環境情報に比べて農作業名称の場合は作業の目的や対象となる作物、作物の成長段階などによって複合的な意味を持つ作業が多く存在するため語彙体系として明確な整理が困難な部分がある。内閣府 IT 総合戦略本部から提示された農作業名称のガイドラインの場合も農作業名称における語彙の多様性を考慮して対応をしている。まず、農作業を大・中・小項目に分類し、小項目は利用者が適宜拡張することができるようになっている。また、すべての中項目に

連絡先: 朱成敏, 国立情報学研究所, 〒101-8430 東京都千代田区一ツ橋 2-1-2, joo@nii.ac.jp

*1 <http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001129518>

*2 <http://jp.fujitsu.com/solutions/cloud/agri/>

*3 http://www.hitachi-solutions.co.jp/geomation_farm/

表 1: ガイドラインの例：「4. 基肥施肥」と「6. 追肥施肥」

大項目	中項目	小項目
4 基肥施肥	1 運搬	
	2 配合	
	3 施肥	
	4 養液栽培における定植時に行う養液作りや施肥	
	5 その他	緑肥すき込み, 秋落ちを防ぐための客土, 施肥溝掘り等
6 追肥施肥	1 運搬	
	2 配合	
	3 施肥	穂肥散布等
	4 養液栽培における定植後の養液の濃度管理や補充など	
	5 その他	施肥溝掘り等

「その他」の項目があり、柔軟に対応することが可能である。しかし、農作業名称のガイドラインが持つ柔軟性はデータの統合・連携をするための基準として適切ではない部分がある。表 1 は大項目である「4. 基肥施肥」と「6. 追肥施肥」の下位項目の例である。2つの大項目の下位項目である中項目に同じ名称の作業を多く存在しており、大項目を参照しない場合、作業を特定することが困難である。このような曖昧な構造はデータ統合・連携のための基準として適切ではない部分があり、また今後新しい農作業名称を追加する際にも効率的な対応が困難である。

2.3 用語標準化のためのオントロジー

前章で述べたように農業分野における用語標準化としての語彙体系は以下の2つの条件が必要とされる。

- 語彙の多様性への対応
- 言語体系としての明確な構造

そこで、本稿では農業分野の用語標準化のためにオントロジーを用いた用語標準化を提案する。オントロジーは概念間の関係をより詳細にし、定義した語彙の間を相互に結びつけることができる。また、概念と語彙を分離することで多義的な意味を持つ概念を記述することも可能である。そして、オントロジーは矛盾を見つける発展性の機能を提供しており、農業分野の語彙を持つ多様性に対応する明確な語彙体系の提供が可能だと考えられる。

3. 農作業名称の標準化のためのオントロジー

我々は農作業における語彙標準化のために農作業を網羅した「農作業基本オントロジー」を構築した。本章ではその設計と構築について述べる。

3.1 農作業オントロジーの設計

農作業オントロジーを構築するためにまず農作業名称の収集を行った。収集対象は内閣府 IT 総合戦略本部が平成 27 年

度に提示した「農業 IT システムで用いる農作業の名称に関する個別ガイドライン(試行版)(案)」に収録された農作業を中心に、農業現場での営農データと、農林水産省の統計調査[農林水産省 12]を加え、農作業名称を網羅した。そして、農作業の構造化、多義的概念、別名の表記を農作業オントロジーに要求される特徴とし、「農作業基本オントロジー」の設計を行った。そして、オントロジーの構築は記述論理に基づいて行い、より明確に構造化した。本節ではその設計について詳しく述べる。

3.1.1 農作業の構造化

まず、農作業の概念を構造化するために農作業を目的に注目し、作物生産を目的とする「作物生産作業」と営農管理を目的とする「営農管理作業」の2つを上位概念とした。実際に圃場で行われる農作業は作物生産作業に含まれ、また目的によって「作物生育作業」、「環境制御作業」、「収穫後作業」、「作物生産支援作業」の4つの下位概念を持つ。このように上位階層では目的の属性が持つ値を基準として分類を行った。

そして、下位階層では農作業をより具体的に定義するために行為、対象、場所、手段、時期を属性とし、その値によって分類を行った。また、特定作物に限定される作業は作物の属性を用いて定義を行った。属性の値は上位概念から下位概念に継承される。例えば、「脱穀」は目的の値を上位概念である「出荷規格調整作業」から継承し、「出荷規格調整」を目的とし、「取り離す」を行為とし、「穀粒」を対象とする作業として定義される。

このように収集された農作業名称を属性の値を用いて分類を行い、明確に構造化した。図 1 は農作業基本オントロジーの構造を表す。

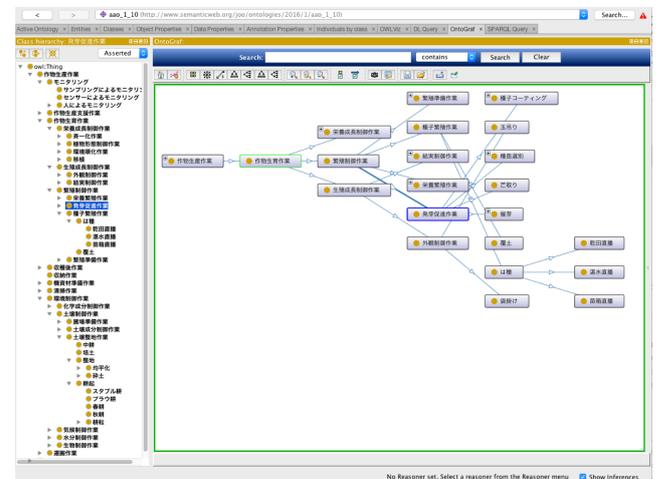


図 1: 農作業基本オントロジー ver.1.10

3.1.2 多義的概念

農作業基本オントロジーは農作業を定義するために属性の値を用いて分類を行う。作業を行う根本的な理由を優先して属性である目的による細分化を行う。しかし、農作業の中では複数の目的によって行われる作業が一部存在する。例えば、生育期間中に畑を耕起する「中耕」は土壌整地のためにも行われるが、雑草を抑制するためにも行われる。また、土の表面に有機物を敷く「マルチング」は温度の保持や調整を目的として行われるが、雑草を抑制する目的もある。このように複数の目的で行われる農作業のために複数の上位概念を持つ多義的な概念を

考慮して構築を行った。

3.1.3 別名の表記

農作業の名称の中では同じ作業でも様々な別名が存在する場合があります。例えば、種を播く作業である「播種」は「種まき」とも呼ばれる。苗を外気に晒して強くさせる「ハードニング」は「硬化」とも呼ばれる。また、作業を記入する場合、「播種」と「は種」、「剪定」と「せん定」のように常用漢字ではない漢字をひらがなで表記する場合もある。このように複数の表記を持つ農作業のために農作業基本オントロジーでは概念と表記を分離した。

3.1.4 記述論理による定義

農作業基本オントロジーは記述論理に基いて構築を行った。例えば、枝を切ることで形を整えたり、養分を効率良く成長させて成長を促進する作業である「せん定」は以下のように表現される。

代かき ≡ 農作業
目的 (植物形態制御 ⊂ 生殖成長)
行為: 切る
対象: 枝 (1)

また、せん定は別名の表記として剪定が存在しており、次のよう表現される。

せん定 ≡ 剪定 (2)

このように農作業基本オントロジーは記述論理に基づいて構築されており、OWL[W3C 04] への変換と RDF スキーマによる語彙定義を用いて定義することが可能である。

3.2 農作業基本オントロジー ver1.10

農作業基本オントロジーは補完と修正を行い、現在 ver1.10 が最新版である。最初公開版である ver0.94^{*4} は目的と手段、作物を属性として農作業名を定義した [竹崎 15]。ver1.00^{*5} では目的、行為、対象、場所、手段、時期、作物に属性を増やすことでより明確な分類が可能となった。また記述論理に基いて定義をし、概念の矛盾や同一性を明確にした [朱 15]。最新版である ver1.10^{*6} では農林水産省による統計調査を参考し、作業名の追加・補完によってより実用性を高めた。

農作業はそれぞれの URI を持ち、その名前空間では農作業が持つ属性の値や上位・下位概念、別名の表記が表示される。図 2 は「代かき」の URI^{*7} の例である。そして、属性の値を用いて自動生成した語彙リストも公開している。語彙リストの詳細は次章で詳しく述べる。

4. 語彙リストの構築

本章では最新版である農作業基本オントロジー ver1.10 を用いて作成した語彙リストについて述べる。

4.1 語彙リストの構成

農作業基本オントロジーは各々の農作業を目的、行為、対象、場所、手段、時期、作物の 7 つの属性の値を用いて定義した。脳作業オントロジーの概念である農作業の名称では 2 つ



図 2: 「代かき」の URI

の種類が存在する「田植え」や「乾田直播」のように農業現場で実際行われる作業の名称と農作業を分類するために上位概念として存在する抽象的な名称がある。本稿では前者を用語、後者を概念と呼ぶ。例えば、用語として定義された「畦畔除草」と「圃場周り除草」の上位概念である「物理的除草抑制作業」は物理的手段を用いた雑草抑制作業の総称する概念である。そして、概念と用語はその属性を上位概念から継承し、定義に用いられた属性の値も包含関係を持つ。

我々は現在農作業基本オントロジー ver1.10 から自動生成された概念リスト、用語リスト、属性リストをオンライン上で公開している^{*8}。図 3 はその語彙リストの例である。上位概念と下位概念、属性へのリンクを辿ることで概念と用語の意味が理解出来るように実装された。

4.2 語彙リストの管理

農作業の名称は栽培方法の改良や環境の変化によって新しい用語が追加される場合も考えられる。こういった新規用語の追加や既存用語の修正などの管理は有識者の見解や現場の意見を参考して農作業基本オントロジーに反映している。また、農作業基本オントロジーは記述論理に基づいて構築されたため OWL[W3C 04] による記述も可能である。そして、OWL は標準のデータ形式として用いられているため機械による推論処理が容易となる。推論処理を行うことで新規用語が追加される際に適格な分類と矛盾を見つけることが可能となる。先行研究 [朱 16] では OWL 推論システムである SWCLOS[Koide 11] を用いて新規用語の上位概念として最も特異な概念を推論することができ、その有用性を確認した。

5. 今後の課題

今後の課題としてはまず、収録された農作業の定義に対する検証が挙げられる。農作業の名称は農業現場によってその意味が異なる場合があり、有識者と共に概念の定義と属性の値について考察が必要である。現在、その検証作業を一部進めており、今後合意を通じて順次更新していく予定である。

*4 2015 年 5 月 12 日公開, 185 語収録

*5 2015 年 11 月 2 日公開, 301 語収録

*6 2015 年 2 月 12 日公開, 330 語収録

*7 <http://cavoc.org/aao/ns/1/代かき>

*8 <http://www.cavoc.org>

a) 概念リスト	
雑草抑制作業 (ざっそうよくせいさぎょう) 英名: activity for weed control 雑草制御を目的とする生物制御作業 上位語, 概念: 生物制御作業 下位語, 概念: 物理的雑草抑制作業 化学的雑草抑制作業 生物的雑草抑制作業 中耕 マルチング 同意語: 除草	階層 4/7
資材消毒 (しざいしょうどく) 資材消毒を目的とし, 資材を対象とする消毒する機資材準備作業 上位語, 概念: 機資材準備作業 下位語, 概念: 鉢消毒	階層 4/7
施設管理作業 (しせつかんりさぎょう) 英名: maintenance activity of facilities 施設管理を目的とする施設機材管理作業 上位語, 概念: 施設機材管理作業 下位語, 概念: ハウス組み立て ハウス修繕 果樹棚補修 誘蛾灯管理 トンネル組み立て	階層 4/7
b) 用語リスト	
種肥散布 (ほこえさんぷ) 追肥作業の一つで 穀物において 生育期間中に窒素分を用いて穂を対象として行う作業 上位語, 概念: 追肥	階層 7/7
補植 (ほしょく) 英名: activity for supplementary planting 移植作業の一つで 補植を目的として行う作業 上位語, 概念: 移植 下位語, 概念: 浮き苗直し	階層 5/7
圃場周り除草 (ほじょうまわりじょそう) 雑草制御を目的として圃場周りで除草する作業 上位語, 概念: 物理的雑草抑制作業	階層 6/7
c) 属性リスト	
畦畔除草剤: 化学的手段の一種, または一部 種わら: 有機物の一種, または一部 空気組成: ガス環境の一種, または一部 空気組成制御: 空気組成の一種, または一部 窒素分: 肥料の一種, または一部	

図 3: 語彙リストの例

今回構築した農作業基本オントロジーは農作業全般における作業を収録した。今後作物に特化した語彙の整理のために作物に特化したオントロジーを農作業基本オントロジーから拡張・分離が出来るように作物の属性を補完していきたい。

本稿で提案した農作業基本オントロジーは用語の標準化を目的として構築された。今後は日付や作業者情報など農作そのものに対して記述をする要素を追加し、農業 IT システムに導入する予定である。

6. おわりに

本稿では農業分野におけるデータ連携のための用語標準化について考察し、解決方法としてオントロジーを用いる手法を提案した。また、用語標準化のために農作業のオントロジーを設計・構築を行った。農業分野の用語標準化に農業オントロジーを用いることで概念の明確な定義や農業分野が持つ用語の多様性に対応することが可能となった。そして、明確な構造は新規用語の追加において適切な分類を行うことができ、効率的な更新も期待される。こういった成果から農業分野の用語標準の条件として提示した「語彙の多様性への対応」と「言語体系としての明確な構造」に対応できたと考えられる。その成果として本稿で提案した「農作業基本オントロジー」は平成 28 年度に内閣府 IT 総合戦略本部が提示した「農業 IT システムで用いる農作業の名称に関する個別ガイドライン (本格運用版) (案)」*9 に反映されたことが挙げられる。今後、農作物、農薬など農業分野における用語標準化にオントロジーを積極的に取り組

み、構築する予定である。

謝辞

本研究は、総合科学技術・イノベーション会議の SIP (戦略的イノベーション創造プログラム) 「次世代農林水産業創造技術」(管理法人: 農研機構生物系特定産業技術研究支援センター) の支援を受けて行った。

参考文献

- [内閣府 15a] 内閣府, 農業情報創成・流通促進戦略に係る標準化ロードマップ(案), <http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/nougyou/dai10/siryoul-2.pdf> 2016 年 3 月 9 日参照
- [内閣府 15b] 内閣府, 農業情報創成・流通促進戦略, <http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/senryakuzenbun_140603.pdf> 2016 年 3 月 9 日参照
- [内閣府 15c] 内閣府, 農業 IT システムで用いる農作業の名称に関する個別ガイドライン (試行版) (案), <http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/nougyou/dai10/siryoul-3.pdf> 2016 年 3 月 9 日参照
- [内閣府 15d] 内閣府, 農業 IT システムで用いる環境情報のデータ項目に関する個別ガイドライン (試行版) (案), <http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/nougyou/dai10/siryoul-4.pdf> 2016 年 3 月 9 日参照
- [農林水産省 12] 農業経営統計調査/農産物生産費/確報/平成 24 年産米及び麦類の生産費/年次/2012 年/「利用者のために」, 農林水産省, 政府統計の総合窓口, <<http://www.e-stat.go.jp/SG1/estat/List.do?lid=0>>
- [竹崎 15] 竹崎あかね, 法隆大輔, 武田英明, 朱成敏, 吉田智一, 情報共有に資する稲作業語彙体系の試作, 農業情報学会 2015 年度年次大会, 個-14, 2015.
- [朱 15] 朱成敏, 武田英明, 法隆大輔, 竹崎あかね, 吉田智一, 農業 IT システム間データ連携のための農作業基本オントロジーの構築, 電子情報通信学会知的環境とセンサネットワーク研究会, 75, 2015.
- [朱 16] 朱成敏, 小出誠二, 武田英明, 法隆大輔, 竹崎あかね, 吉田智一: 記述論理に基づく農作業オントロジーの設計と応用, 第 38 回人工知能学会セマンティックウェブとオントロジー研究会 (SIGSWO), 06, 2016
- [Koide 11] Seiji Koide, Theory and Implementation of Object Oriented Semantic Web Language, NII 論文 ID:500000547373, 2011
- [W3C 04] OWL Web Ontology Language Guide, W3C Recommendation 10 February 2004, <https://www.w3.org/TR/owl-guide/>

*9 平成 28 年 3 月 31 日公開