

# 農作業基本オントロジーに基づく米の生産費統計調査の自動化

## Automatic Statistical Processing for Production Cost of Rice Based on Agriculture Activity Ontology

朱成敏<sup>1\*</sup> 武田英明<sup>1</sup> 法隆大輔<sup>2</sup> 竹崎あかね<sup>2</sup> 吉田智一<sup>2</sup>  
 Sungmin JOO<sup>1</sup> Hideaki TAKEDA<sup>1</sup> Daisuke HORYU<sup>2</sup>  
 Akane TAKEZAKI<sup>2</sup> Tomokazu YOSHIDA<sup>2</sup>

<sup>1</sup> 国立情報学研究所

<sup>1</sup> National Institute of Informatics

<sup>2</sup> 農業・食品産業技術総合研究機構

<sup>2</sup> National Agriculture and Food Research Organization

**Abstract:** This paper proposes automatic statistical processing for census of agriculture using agriculture activity ontology. The spread of agricultural IT systems has been accompanied by generating data such as farming information or agricultural activities. These data is expected to be used for census of agriculture by statistical processing. However, data in agricultural IT systems is nonetheless not easy to integrate since the languages to describe data are not unified. In this paper, we propose automatic statistical processing system based on agricultural activity ontology as core vocabulary for data integration between census for agriculture and agricultural IT systems.

## 1 はじめに

農作物生産費統計は農作物の一定単位を生産するために消費される経済費用に関する調査であり、農作物の生産計画はもちろん国の食糧や経済政策の策定、予測においても重要な指標である。その調査方法は対象農家に作業日誌や現金出納帳などの関連資料を配布し、自計調査による記入を基本としている。集計は職員による地域毎のデータ統合を行う。しかし、農家には農業ITシステムが普及されており、そのデータから直接集計することが可能であればデータの記入や集計の時間が短縮され、調査による負担が軽減される。しかし、データを統合するためには農家が運用している農業ITシステムのデータ項目が統計調査の項目に対応していなければならない。

そこで、本研究は農作業基本オントロジーを用いて農作業名称の基準を提示し、統計調査の自動化について考察する。また、農家の実データを用いて米の生産費統計調査における集計システムを実装し、自動化への可能性を検討する。

## 2 農産物生産費統計調査の現状

### 2.1 概要

農産物生産費統計調査は農林水産省によって毎年行なわれている調査であり、米や小麦など12種の重要農産物の生産費の実態を明らかにし、生産や経営に関する対策の策定における資料を整備することを目的としている。作物の生育期間によって開始時期は異なるが1年間単位で調査を行う。調査項目としては作物の生産活動に投入された費目別の費用、労働時間、肥料などの資材の費用の生産活動における項目と就業者数、耕地面積などの経営に関する項目がある。調査は現金出納帳、作業日誌、経営台帳を調査対象の農家に配布し、記入してもらおう方法で行われ、集計対象ごとに加重平均法により算出する。調査結果は農林水産省のホームページにて公表される<sup>1</sup>。

### 2.2 問題点と農業ITシステムとの連携

農産物生産費統計調査は国の農業政策において貴重な調査であるが、データ集計において農家の協力が必要となり、また、負担となる可能性もある。現在、農産物生産費統計調査ではデータを収集するために調査対象

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

<sup>1</sup> [http://www.maff.go.jp/j/tokei/kouhyou/noukei/seisanhi\\_nousan/](http://www.maff.go.jp/j/tokei/kouhyou/noukei/seisanhi_nousan/)

の農家に現金出納帳、作業日誌、経営台帳を配布している。例えば、作業日誌は作業区分、作用社名、1日当たり標準労働時間、作業日、参考の項目がある。調査対象の農家はこれらの項目を記入し、提出をする。また、場合によっては職員による面接調査が併用されることもある。こういった場合、農家は統計調査のために作業日誌を記入する必要があり、農業 IT システムを用いて営農管理を行っている場合、データ入力への負担が二重となる。こういった場合、農家が営農管理のために運用している農業 IT システムのデータと集計システムを連携することで迅速で且つ正確な集計が可能となる。

現在の集計方法のもう一つの負担はデータの統合である。調査対象の農家がそれぞれ記入した資料からデータを統合する際にデータ項目となる語彙が統一されていないことから、人による判断が必要となる。現在は担当職員の判断によって語彙を整理して行っている。これは農家の慣習や地域による用語の揺らぎが生じるからである。例えば、表 1 は研究目的での利用許可を得た農家法人の実データから抽出した稲作に関わる作業名称と農林水産省の「米及び小麦の生産費」の作業分類一覧表 [1] から提示された作業名称との対応表である。このように農家が作業名称として使っている用語と標準語彙の差があり、農業 IT システムからデータを収集する際に作業日誌の作業名称を農作業の標準語彙へ変換する必要がある。

表 1: 稲作の作業一覧.

作業日誌の作業名	統計調査の作業分類
1 番耕起	耕起
2 番耕起	
3 番耕起	
肥料散布	施肥
追肥	追肥
畦畔除草	畦畔除草
稲刈り	稲刈り
整地	整地
畔塗・修繕	畔塗修繕
堆肥散布	堆肥散布
除草剤散布	除草剤散布
スタブル耕	スタブル耕
プラウ耕	プラウ耕
田植え	田植え
その他播種	は種
代かき	代かき

## 2.3 農産物生産費統計調査の自動化

前説で述べたように農産物生産費統計調査を農業 IT システムと連携し、効率良く行うためには以下の 2 つの条件が必要とされる。

- 農業 IT システムからのデータ集計
- 作業名を標準語彙へ変換

そこで、本稿では稲作を対象に農業 IT システムから発生したデータを農産物生産費統計調査の調査項目に対応させ、集計を行う自動化について考察する。農業 IT システムと統計調査の連携が実現されると、農家にとっては統計調査における手間が省け、また集計機関にとっては集計にかかる時間が短縮されるメリットが予想される。また、事例の収集により生育期間中に生産費用や生産量の予測も可能となり、環境要因による変化に対応する政策の修正も可能となる。農家においてはリアルタイムで営農状況を把握することができ、農家の生産性増大にも良い影響を与えられられる。

## 3 統計調査のための語彙の処理

本章では農業 IT システムと農産物生産費統計調査の連携による自動化に必要なとされる実データの作業名称と統計調査の統計項目間の対応について考察する。

### 3.1 作業データの分析

農産物生産費統計調査を自動化するためにまず、農業 IT システムを運用している農家法人の実データを入手した。導入されている農業 IT システムは農研機構による作業計画・管理支援システム (PMS, 以下 PMS)<sup>2</sup> である。PMS はデータを対外的に交換するために CSV やエクセルのようなスプレッドシート用形式形式から EU による農業データの標準規格である agroXML<sup>3</sup> 基盤の XML 形式まで用いることが可能である [3]。また、agroXML は作物の生産、土地の利用、経営管理など、営農における詳細な情報を記入することができるため、PMS から抽出したデータは作業の内容や日付、労働時間、関連する資機材の詳細など作業に関わる様々な情報が記録されている。

本研究で対象とした作業データは、研究目的での利用許可を得た農家法人の実データであり、2015 年 10 月 25 日から 2016 年 10 月 11 日まで 1 年間の 2,247 件の稲作に関わる作業の内容が記録されている。作業別の記録件数を表 2 に表す。

<sup>2</sup><http://www.aginfo.jp/PMS/>

<sup>3</sup><http://www.aginfo.jp/PMS/FarmXML/agroxml/agroxml.html>

表 2: 作業別のデータ件数.

作業名	記録件数
1 番耕起	13
2 番耕起	166
3 番耕起	121
肥料散布	254
追肥	320
畦畔除草	515
稲刈り	307
整地	48
畔塗・修繕	233
堆肥散布	90
除草剤散布	32
スタブル耕	95
プラウ耕	8
田植え	3
代かき	34
その他播種	8
(計)	2,247

記録されている作業名の中では農作業の標準語彙と一致する作業名もあるが、以下の作業は対応する語彙を検討する必要がある。

- 1 番耕起, 2 番耕起, 3 番耕起, その他播種
- 肥料散布, 畔塗・修繕

まず、肥料散布, 畔塗・修繕については同義語, または表記の問題である。1 番耕起, 2 番耕起, 3 番耕起は耕起の下位概念に思われるが、同じ地番に対して3つの耕起が連続で行われていないため数字の意味が不明瞭である。また、作業日の順序が番号順と異なる点や使用された機械の機体番号が一致している点など、これらの3つの作業が目的の異なる別作業であるとは言い難い。従って1 番耕起, 2 番耕起, 3 番耕起は「耕起」に対応させることが妥当だと考えられる。その他播種が8件全部同じ内容で使用された機械は田植え機, 資材としては鉄コーティングをした種子であり, 鉄コーティング湛水直播の栽培方法の直播作業だと思われる。従ってその他播種は「は種」に対応させることが適切である。

このように作業日誌に作業名の揺らぎが生じる原因は2つに大別される。まず、一つ目の原因は別名, 同義語などの表記の問題である。これは作業の目的や手段など, 作業の定義は変わっていないため用語を自動的に置換することで処理が可能となる。2つ目の原因は農家が独自の作業名を用いるからである。こういった場合は作業の詳細から判断し, 標準語彙となる作業名に対応させる必要がある。

### 3.2 農作業基本オントロジー

農作業基本オントロジー (AAO, Agriculture Activity Ontology) は農業 IT システム間データの連携・統合のために開発された農作業名称のた語彙体系である [4]。農作業基本オントロジーでは農作業を「目的」, 「行為」, 「対象」, 「場所」, 「手段」, 「道具」, 「時期」, 「作物」の属性の値を用いて定義し, 概念間の包含関係に基づく階層構造として構築された。また, 農作業名称の読み, 同義語, 英名も収録されている。185 語が収録されていた公開版 (2015 年 5 月公開) から 5 回の更新を通じて収録語彙数は 410 語となった (表 3)。また, 123 語の同義語が収録されており, 同義語を標準語彙へ変換する API も提供している<sup>4</sup>。

表 3: 農作業基本オントロジーの更新履歴.

バージョン	公開日	語彙数
0.94	2015-05-12	185
1.00	2015-11-02	301
1.10	2016-02-12	330
1.31	2016-04-22	367
1.33	2016-09-23	374
1.41	2017-01-05	410

農作業基本オントロジーは農林水産省による「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)[2]」と農産物生産費統計調査の調査項目に対応しており, 農業 IT システムの作業名を農作業基本オントロジーの作業名に対応させ, 統計調査項目の用語に変換することでデータの集計が可能となる。そのため, 本研究ではまず, 農家の作業データから作業名を抽出し, 農作業基本オントロジーの語彙に変換を行う。その後, 変換された作業名毎に労働時間の集計を行い, 農作物生産費統計調査における米の作付規模別生産費用の調査項目である作業別労働時間として算出する (図 1)。

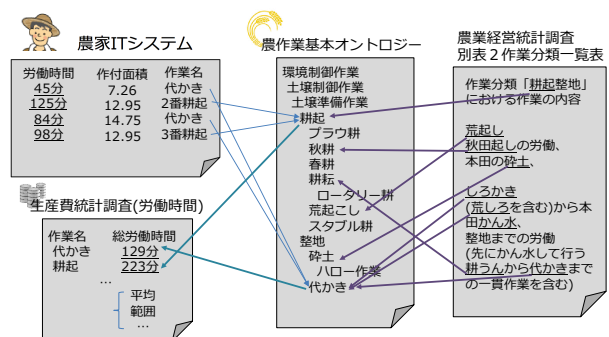


図 1: 農作業基本オントロジーに基づく集計.

<sup>4</sup><http://cavoc.org/aao.php>

## 4 統計調査の自動化

本研究で提案する生産費統計調査の自動化のために統計システムを実装した。本システムは農業 IT システムの作業データを指定されたデータ項目に合わせて格納し、格納されたデータから農作業基本オントロジーの作業名称とデータの作業名のマッピングを行う。そして、マッピングされた作業名毎に労働時間の集計を行い、その結果を表示する。

### 4.1 データの入力

データの入りは CSV 形式で行われる。農作業基本オントロジーは農作業を目的、行為、対象、場所、手段、道具、時期、作物の 8 つの属性を持つ値から定義している。これらの属性と agroXML の共通項目である機資材を含む道具、作物、場所、作物、時期 (日付) を作業名とともに指定し、作業データをシステムに格納する (図 3)。

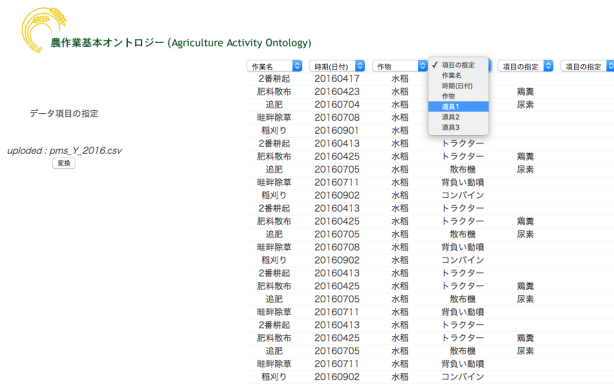


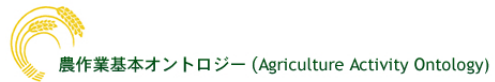
図 2: 作業データの項目指定。

### 4.2 作業名のマッピング

本システムは 2 つの方法で作業名のマッピングを行う。まず、農作業基本オントロジーが持つ同義語を用いて作業名を整理する。これは共通農業語彙<sup>5</sup>が提供している農作業標準語彙への変換 API を用いる。これにより今回の実験対象となる作業データ数の 2,247 件のうち 1,939 件が農作業基本オントロジーの作業名に対応付けられた (図 4)。

マッピングに失敗した残り 308 件のデータの作業名は 1 番耕起, 2 番耕起, 3 番耕起, その他播種であり, 2.3 で述べたように農家法人が独自で定義した作業である。こういった独自に定義された作業名はデータ項目の特徴を考慮し手動でマッピングを行う (図 4)。システム

<sup>5</sup>http://cavoc.org/



作業名(作業者数)	総労働時間	作業者平均	1日平均
スタプル耕(3名)	103時間 53分	34時間 37分	01時間 26分
ブラウ耕(1名)	12時間 30分	12時間 30分	04時間 10分
代かき(1名)	44時間 59分	44時間 59分	08時間 59分
堆肥散布(2名)	119時間 42分	59時間 51分	03時間 31分
整地(3名)	116時間 09分	38時間 43分	01時間 36分
施肥(1名)	101時間 47分	101時間 47分	05時間 59分
田植え(1名)	06時間 59分	06時間 59分	06時間 59分
畦畔修繕(1名)	05時間 17分	05時間 17分	00時間 13分
畦畔除草(5名)	165時間 16分	33時間 03分	01時間 08分
稲刈り(3名)	276時間 29分	92時間 09分	02時間 25分
追肥(3名)	147時間 00分	49時間 00分	02時間 27分
除草剤散布(2名)	12時間 57分	06時間 28分	02時間 09分
未指定作業 308 件	294時間 56分		

ダウンロード [CSV]

図 3: 同義語 API によるマッピング結果。

はデータから道具、作物、場所、作物、時期 (日付) を抽出し、マッピングを行う判断資料として提示する。管理者は提示された項目を選択し、脳作業基本オントロジーの農作業名と指定することによってマッピングを行う。



図 4: 属性による作業名のマッピング。

### 4.3 データの集計

農作業基本オントロジーの作業名とマッピングされた作業データから労働時間と労働者の情報を集計し、作業別の平均労働時間を表示する。また、その結果を CSV 形式で出力し、ダウンロードすることも可能である (表 5)。



農作業基本オントロジー (Agriculture Activity Ontology)

作業名(作業者数)	総労働時間	作業者平均	1日平均
スタブル耕(3名)	103時間 53分	34時間 37分	01時間 26分
ブラウ耕(1名)	12時間 30分	12時間 30分	04時間 10分
代かき(1名)	44時間 59分	44時間 59分	08時間 59分
堆肥散布(2名)	119時間 42分	59時間 51分	03時間 31分
播種(1名)	06時間 57分	06時間 57分	06時間 57分
整地(3名)	116時間 09分	38時間 43分	01時間 36分
施肥(1名)	101時間 47分	101時間 47分	05時間 59分
田植え(1名)	06時間 59分	06時間 59分	06時間 59分
畔塗修繕(1名)	05時間 17分	05時間 17分	00時間 13分
畦畔除草(5名)	165時間 16分	33時間 03分	01時間 08分
稲刈り(3名)	276時間 29分	92時間 09分	02時間 25分
耕起(3名)	287時間 59分	95時間 59分	02時間 31分
追肥(3名)	147時間 00分	49時間 00分	02時間 27分
除草剤散布(2名)	12時間 57分	06時間 28分	02時間 09分
未指定作業 0 件	00時間 00分		

ダウンロード [CSV]

図 5: 作業別労働時間の最終集計.

## 5 考察

本研究では農研機構が開発した作業計画・管理支援システム (PMS) のデータを用いた。PMS は agroXML 基盤のデータ形式で交換用データを生成することが可能であり、データの再利用に適している。一方、ベンダーによってはデータをエクスポートする機能が実装されていない場合も確認され、本研究の提案手法の適用が困難な場合も予想される。

統計調査を自動化するために農作業基本オントロジーで農作業を定義する属性の値を参照し、作業データのデータ項目との連携を試みたが、作業データに記録されていた項目が作業を区別する判断資料としては適切ではなかった。例えば、耕起の下位概念であるブラウ耕とスタブル耕はそれぞれブラウとスタブルという機会を用いて行う耕起である。しかし、作業データでは両方トラクターと記録しており、作業名以外では区分が不可能な場合もある。

今後の課題としては異なる農業 IT システムのデータを用いてデータを統合し、集計を行うことが挙げられる。また、本稿では農家が独自で定義した農作業はデータを提示し、人の判断によりマッピングを行ったが、今後は事例を収集し、適用条件を学習することによって集計を自動化させたい。

## 6 むすび

本稿では農業 IT システムのデータを対象に米の生産費統計調査の自動化を行う方法について考察し、集計システムの実装を行った。本研究の成果によって農業 IT システムの営農管理データから作業名を農作業の標準語彙と対応させ、データを集計することができた。従来の集計方法より効率よく、かつより詳細なデータを集計することが可能となり、今後の農業分野に対する統計調査の集計方法として、その可能性を確認した。

## 謝辞

本研究(の一部)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

- [1] 農林水産省, 農業経営統計調査報告 米及び小麦の生産費,  
< [http://www.maff.go.jp/j/tokei/kouhyou/noukei/seisanhi\\_nousan/data/pdf/seisanhi\\_komemugi\\_riyo\\_u.08.pdf](http://www.maff.go.jp/j/tokei/kouhyou/noukei/seisanhi_nousan/data/pdf/seisanhi_komemugi_riyo_u.08.pdf) > 2017年2月14日参照
- [2] 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 (本格運用版),  
< [http://www.kantei.go.jp/jp/singi/it2/senmon\\_bunka/shiryo/siryoyou04.pdf](http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/shiryo/siryoyou04.pdf) > 2017年2月14日参照
- [3] 吉田 智一, 木浦 卓治, 南石 晃明: 農業生産工程管理データ表現形式 FIX-pms の開発, 農業情報研究, Vol. 22, No. 2 p. 103-116, 2013.
- [4] 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 回人工知能学会セミナー ウェブとオントロジー研究会 (SIGSWO), 06, 2016.