

Title	マークシート処理システムの拡張による調査方式の実現：学力調査を対象とした適用事例と評価
Sub Title	
Author	森, 薫(Mori, Kaoru)
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2015
Jtitle	リサーチメモ：CMR research memorandum 2014 ,p.67- 77
JaLC DOI	
Abstract	本稿では、拡張可能なマークシート処理システムを利用して、広域的な調査を実現するシステム構成を示す。本システムの適用事例として、宮城県で実施した平成24年度みやぎ学力状況調査におけるシステム運用について示す。この調査では、宮城県内公立高等学校の生徒(約30,700人)を対象として、マークシート形式の学力調査を実施した。回収した解答用紙は各学校で電子データ化され、そのデータを本システムに登録した。運用期間におけるデータの登録状況を評価し、本システムの実務での運用可能性について検討する。
Notes	
Genre	Technical Report
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO92001002-2015-001-0067

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

マークシート処理システムの拡張による調査方式の実現

—学力調査を対象とした適用事例と評価—

慶應義塾大学大学院政策・メディア研究科

森 薫

kmori@sfc.keio.ac.jp

要旨

本稿では、拡張可能なマークシート処理システムを利用して、広域的な調査を実現するシステム構成を示す。本システムの適用事例として、宮城県で実施した平成 24 年度みやぎ学力状況調査におけるシステム運用について示す。この調査では、宮城県内公立高等学校の生徒（約 30,700 人）を対象として、マークシート形式の学力調査を実施した。回収した解答用紙は各学校で電子データ化され、そのデータを本システムに登録した。運用期間におけるデータの登録状況を評価し、本システムの実務での運用可能性について検討する。

キーワード 学力調査, マークシート, データベース, データ分析

1. はじめに

学力調査や学校評価のアンケートなど、マークシート形式の調査が広く実施されている。慶應義塾大学 CMR プロジェクトにおいては、オープンソースのマークシート処理システムの開発が推進されてきた（久保ほか 2005）。学校評価におけるアンケート調査を支援することを主たる目的として、SQS は開発されてきた。SQS はオープンソースとして開発され、無償で利用可能であることから、学校評価以外にも利用されている。

東日本大震災の被災地である石巻市においては、被災者の健康状況や社会環境を調査するヘルスアセスメントが継続的に実施されている。看護師のボランティア団体によって、SQS を利用したマークシート形式の調査が実施された（宮川ほか 2013）。その他にも、スポーツの社会的価値向上を支援する評価手法（松橋ほか 2012）や、地方自治体における住民参加型行政のための住民アンケート（藤沢市 2011）に利用された。

本稿では、SQS を拡張して広域的な学力調査への適用について示す。学力調査を効果的に活用するためには、教育委員会や学校が多面的な分析を行うだけでなく、その調査結果が早期に提供されることが重要である（文部科学省 2008）。広域的な学力調査を実施し、学校や教育委員会において、役割に応じたデータ分析機能を提供するシステムが開発されてきた（木幡ほか 2007, 森ほか 2010）。このシステムを利用することにより、学力調査の採点と電子データ化を

教員が実施し、学校や教育委員会が役割に応じたデータ分析を実施することによって、調査結果の活用場面が拡大した。しかし、学力調査の採点結果を教員が手動でファイルに転記するなど、実務的なコストが課題となっていた。電子データ化を効率化するためには、マークシート形式の調査が有効である。本システムでは、オープンソースのマークシート処理システムである SQS (久保ほか 2005) を拡張して、自由度の高いマークシート調査票の作成を実現する。SQS は XML 技術を基盤としたマークシート処理システムであり、その拡張も XML によって可能である。さらに、マークシート処理システムを拡張するための XML を利用して、データ分析システムが動作する。データ分析システムは XML から学力調査の設問構造を認識し、その構造に応じたデータ分析を実行する。

普通紙とイメージスキャナーを利用したマークシート処理方式 (Deng et al. 2008, 河野ほか 2011) や、調査票におけるマーク位置の座標情報を設定することによる調査票の作成方式 (高橋・松田 2007) が提案されている。本システムの特徴は、XML 技術を基盤としたマークシート処理システム (久保ほか 2005) を拡張し、マークシート処理システムとデータ分析システムの連携を実現することにある。学力調査の実施において、マークシート処理システムとデータ分析システムが連携することによって、学校や教育委員会が調査結果を早期に電子データ化して、そのデータを分析することが可能となる。

宮城県で実施した「平成 24 年度みやぎ学力状況調査」において本システムを運用した。この調査では、宮城県内公立高等学校の生徒 (約 30,700 人) を対象として、マークシート形式の学力調査を実施した。回収した解答用紙はマークシート処理システムによって各学校で電子データ化され、それをデータ分析システムに登録した。システム運用期間におけるデータの登録状況を評価し、本システムの実務での運用可能性について検討する。

本稿は教育工学会論文誌に掲載された論文 (森ほか 2013) に加筆して編集したものである。

2. マークシート処理システムの拡張

本システムは、学力調査に対応したマークシート形式の解答用紙を作成するために、オープンソースのマークシート処理システムである SQS を拡張する。SQS は普通紙によるマークシート調査を実現するソフトウェアであり、調査票を作成するための SQS SourceEditor と、イメージスキャナーで読み取った調査票の画像に対してマーク認識を実行する SQS MarkReader から構成される。SQS SourceEditor においては、XML エディタによって調査票を作成する。

SQS SourceEditor の XML エディタを利用することによって、設定した質問が自動的にレイアウトされる。SQS SourceEditor のレイアウト形式はアンケート調査を前提としているため、学力調査などへの適用が困難であった。本システムは、調査票のマーク位置を自由に設定可能とする拡張によって、SQS SourceEditor と互換性のある XML を生成する。これにより、SQS

MarkReader によってマーク認識可能な調査票を生成する。さらに、実施した調査結果を広域的に集計したデータ分析が可能となる。

2.1 データ構造

SQS SourceEditor で作成した調査票は PDF 形式で出力される。その PDF には SQS ソースファイルと SQS マークファイルが添付されている。これらのファイルは SQS MarkReader がマーク認識する際に利用する。

2.1.1 SQS ソースファイル

SQS ソースファイルは調査票の質問構造を格納する XML ファイルである。このファイルには質問文と選択肢の情報と調査のメタデータが記述されている。質問構造は XHTML、および、XForms に準拠した XML で記述されている（図 1）。

```

<?xml version="1.0" encoding="UTF-8"?>

<html xmlns="http://www.w3.org/2002/06/xhtml12" xmlns:sqs="http://sqs.cmr.sfc.keio.ac.jp/2004/sqs"
xmlns:xforms="http://www.w3.org/2002/xforms">
<head lang="ja">
  <title/>
  <meta content="" name="DC.Title"/>
  <meta content="kugua" name="DC.Title.Short"/>
  <meta content="kugua" name="DC.Creator"/>
</head>
<body>
  <h/>
  <p> </p>
  <p> </p>
  <h>ここから質問が始まります. </h>
  <p> </p>
  <xforms:select1 sqs:cols="3">
    <xforms:hint>学校種別</xforms:hint>
    <xforms:item sqs:colspan="1">
      <xforms:label>1</xforms:label>
      <xforms:value>1</xforms:value>
    </xforms:item>
    <xforms:item sqs:colspan="1">
      <xforms:label>2</xforms:label>
      <xforms:value>2</xforms:value>
    </xforms:item>
  </xforms>

```

図 1 SQS ソースファイル例

2.1.2 SQS マークファイル

調査票に描画されるオブジェクトは SQS マークファイルによって記述される。SQS マークファイルには、調査票上のマーク欄や自由記述欄に関する座標情報が記述されている (図 2)。

```

<svg:g id="mark1/0">
<svg:rect x="187" y="231" width="5" height="16">
<svg:metadata xmlns:svg="http://www.w3.org/2000/svg">
<xforms:select xmlns:xforms="http://www.w3.org/2002/xforms"
xmlns:sqs="http://sqs.cmr.sfc.keio.ac.jp/2004/sqs" sqs:qid="1" sqs:itemIndex="0" sqs:pxform-type="select1">
<xforms:label>(1)</xforms:label>
<xforms:hint>学校種別</xforms:hint>
<xforms:item>
<xforms:label>1</xforms:label>
<xforms:value>1</xforms:value>
</xforms:item>
</xforms:select>
</svg:metadata>
</svg:rect>
</svg:g>

```

図 2 SQS マークファイル例

SQS MarkReader は SQS マークファイルの情報に基づいてマーク認識処理を実行する。

2.2 システム構成

本システムは調査票作成システム、マークシート処理システム、データ分析システムから構成される。本システムにおけるマークシート処理システムは SQS MarkReader である。SQS MarkReader でマーク認識可能な調査票を作成するために、調査票作成システムにおいて SQS ソースファイルと SQS マークファイルの仕様に合致した XML ファイルを生成する。マーク位置を設定できることから、自由度の高いマークシート形式の調査票を作成することが可能である (図 3)。マークシート処理システムは XML の記述に基づいてマーク認識処理を実行して、その結果をファイルに出力する。

マークシート処理システムから出力されたファイルをデータ分析システムに登録することで、データ分析が可能となる。データ分析システムにおいては、調査票作成システムから生成された XML を利用して、データ処理が実行される。すなわち、調査票作成システムから生成された XML を介して、マークシート処理システムとデータ分析システムが連携する。

国語A 小学校第6学年
解答用紙

5

4

ウ	イ	ア
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5

3

1 ○
 2 ○
 3 ○
 4 ○

2

イ
花が

↓

1 ○
 2 ○
 3 ○
 4 ○

ア
わたしは

↓

1 ○
 2 ○
 3 ○
 4 ○

1

ニ			一		
(3)	(2)	(1)	(3)	(2)	(1)
			()		

《ここから下には解答を書いてはいけません。》

○ 正
○ 誤

○ 正 ○ 正 ○ 正 ○ 正 ○ 正

○ 誤 ○ 誤 ○ 誤 ○ 誤 ○ 誤

性別	出席 番号	組
男	0	0
女	1	1
	2	2
	3	3
	4	4
	5	5
	6	6
	7	7
	8	8
	9	9

図 3 調査票の例

3. データ分析システム

SQS MarkReader によるマーク認識を実行すると、その結果はエクセル形式と TSV(タブ区切り)形式のファイルに出力される。また、SQS MarkReader は単体のソフトウェアとして、マーク認識結果を円グラフや棒グラフで表示する機能を有している(図 4)。しかし、広域的な調査におけるデータ分析を実現するためには、各学校のデータを集約して集計することが必要となる。各学校においてデータを登録し、そのデータを分析可能とするデータ分析システムを実装した。

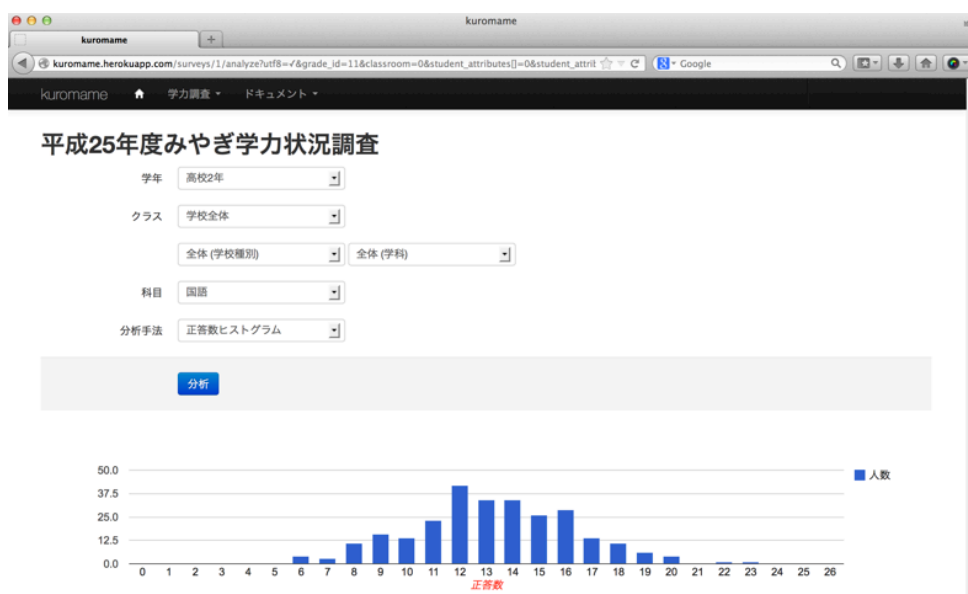


図 4 データ分析画面例

3.1 データ登録機能

広域的なデータの集約を実現するために、SQS MarkReader で解答用紙をマーク認識した結果として出力されたファイルを、データ分析システムに登録する。データ分析システムにおいては、登録されたデータの妥当性が検証される。調査票作成システムから生成された XML を利用して、調査の設問数と各設問に対応した数値範囲を取得する。それらの情報から、登録されたデータが調査票の仕様に合致しているか検証することが可能となる。

解答用紙をスキャンする際の歪みやマーク欄の記入状況が原因となり、マーク認識が正常に実行できなかった場合、マーク認識結果のファイルを編集してデータを修正することも想定される。手動で修正したデータには記入ミスが想定されるため、データ登録時にファイルの妥当性を検証することにより、利用者がファイルを修正する作業を支援する。

3.2 データ分析機能

登録されたデータに対して分析処理を実行する際、分析対象となるデータの規模に応じて計算時間は増大する。利用者からのリクエストに応じて分析処理を起動した場合、多数の利用者からのリクエストの集中によって計算資源が枯渇することが想定される。データ分析システムにおいては、データ分析結果はデータ登録の際に計算してデータベースに蓄積される。結果が事前に計算されているため、対象となるデータの規模が増大したとしても、それによる性能低下を回避することが可能である。この処理はデータ登録をトリガーとして非同期に実行される。

データ分析システムは Web アプリケーションとして構築した。データ分析システムを利用するためには、ID とパスワードによるユーザ認証が必要である。Web アプリケーションは Web

ブラウザを経由してアクセスするため、クライアント環境への依存性が少ない。広域的な調査を実施する際には、多数のクライアントに専用のソフトウェアを導入して利用環境を構築することは現実的ではないため、Web アプリケーションが適していると考えられる。Web アプリケーションは Ruby on Rails で構築し、データベースには MySQL を利用した。また、データ分析のための非同期処理は DelayedJob で制御した。

利用者にデータ分析の結果を表示するデータ分析機能を実装した。データ分析の種類として、設問別正答率、観点・領域別正答率、正答数ヒストグラム、正答率箱ひげ図、クロス集計がある。これらは SQS で提供されていない機能であり、本システムにおける拡張によって実現された。データ分析のひとつの画面例として、正答数ヒストグラムの分析画面を図 2 に示す。データ分析においては、分析するデータの範囲とデータ分析の種類を指定することによって、データ分析の結果が表示される。各学校がデータ登録した時点でデータ分析が可能であり、学校は分析データを早期に活用することが可能である。

4. システム運用

4.1. 学力調査の実施とデータ登録

平成 24 年度みやぎ学力状況調査において本システムを運用した。調査は平成 24 年 7 月 5 日から 7 月 13 日の間に宮城県内公立高等学校の生徒（約 30,700 人）を対象として実施した。回収したマークシート形式の解答用紙は各学校でスキャンされ、マークシート処理システムから出力されたファイルをデータ分析システムに登録した。

マーク認識の際に発生したエラーについては、学校が結果ファイルを編集して修正することとした。各学校がデータ分析システムを利用し、7 月 27 日までに 82.4%、8 月 3 日までに 96.4% のデータ登録が完了した。データ登録の状況を図 5 に示す。

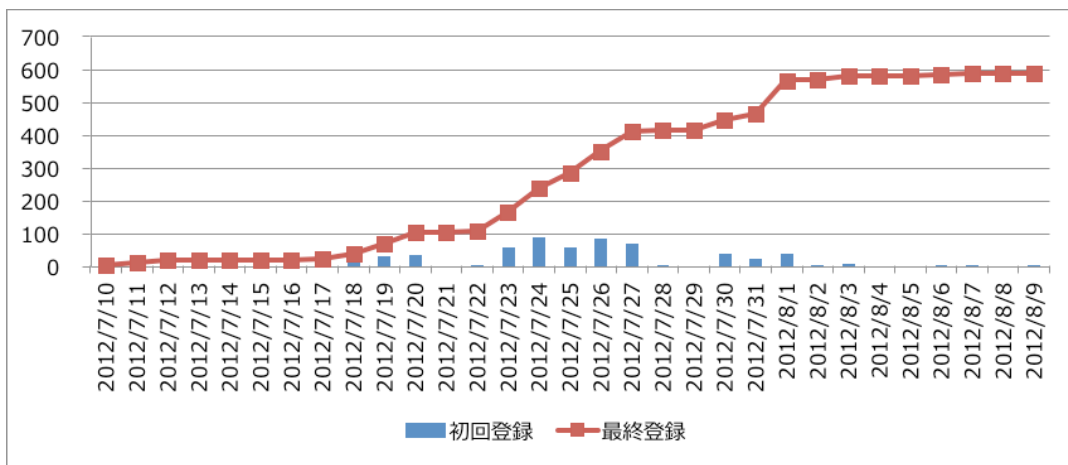


図 5 データ登録状況

4.2. 結果と考察

スキャンした解答用紙の画像をマークシート処理システムでマーク認識する際には、イメージスキャナーの設定に起因するエラーが発生した。具体的には、スキャンした画像をグレースケールとして処理した場合、画像の色濃度が低いためにSQS MarkReaderがマークを認識できないケースがあった。この問題に対処するため、SQS MarkReaderではモノクロ2階調の画像を利用することが推奨されている。また、イメージスキャナーの画質の設定によっては、生成された画像にノイズが発生しやすくなる。システム運用において実際に生成された画像を図6に示す。ただし、画像には個人情報が含まれているため、画像の一部を編集した。

図6 マーク認識エラーの発生した解答用紙例

イメージスキャナーの設定を修正することで障害が解決できたケースが大半であり、学校におけるイメージスキャナーの設定が本システムの利用率を左右したと考えられる。学校で解決することが不可能だったエラーについては、宮城県教育委員会と宮城県教育研修センターがサポートして、最終的には全データをデータ分析システムに登録した。

5. 結論

本稿では、拡張可能なマークシート処理システムを利用して、広域的な学力調査を実現するシステム構成を示した。本システムではマークシート処理システムとデータ分析システムが連携することによって、マークシート形式の広域的な学力調査を早期にデータ分析することが可能となった。

宮城県内公立高等学校の生徒を対象として実施した平成24年度みやぎ学力状況調査において本システムを運用した。その全データがデータ分析システムに登録され、学校と教育委員会にデータ分析機能を提供した。

今後の課題として、データ分析機能の拡充とマーク認識エラーの修正インターフェースの実現が挙げられる。より詳細なデータ分析をするために、複数の設問間で誤答パターンを分析する方式を実現する予定である。また、回収した解答用紙の状態に起因するマーク認識が不可能なケースは、今後も継続的に発生することが想定される。適切なデータ修正が簡単に実施できるインターフェースについて実現する予定である。

参考文献

- H. Deng, F. Wang and B. Liang: “A Low-Cost OMR Solution for Educational Applications”,
Proceedings of International Symposium on Parallel and Distributed Processing with
Applications, pp. 967-970 (2008)
- 藤沢市, “藤沢市新総合計画基本構想”,
<https://www.city.fujisawa.kanagawa.jp/content/000338776.pdf> (2011).
- 河野広貴, 船曳信生, 中西透, 天野憲: “Web を用いたマークシート方式テスト支援機能の実装”,
電子情報通信学会技術研究報告教育工学(ET), Vol. 110, No. 453, pp. 193-198 (2011)
- 木幡敬史, 岡子泰三, 森薫, 玉村雅敏, 金子郁容: “都道府県レベルのための学力テスト分析システム：デザイン・開発・実施運用”, 日本教育工学会論文誌, Vol. 31, Suppl., pp. 169-172 (2007)
- 久保裕也, 玉村雅敏, 木幡敬史, 金子郁容: “カスタマイズ可能な調査スキーマの共有による学校評価支援”, 情報処理学会論文誌, Vol. 46, No. 1, pp. 172-186 (2005)
- 松橋崇史, 玉村雅敏, 岩月基洋, 小西宏: スポーツの社会的価値マネジメントを支援する評価手法の研究開発, スポーツ産業学研究, Vol.22, No.1, pp.117-130 (2012).
- 文部科学省: “平成21年度全国学力・学習状況調査の実施方法等の改善について” (2008)
- 宮川祥子, 岩月基洋, 森薫, 佐々木あかね, 吉田絵里: “東日本大震災被災地におけるITを活用したヘルスアセスメント集計の効率化”, 第72回日本公衆衛生学会総会抄録集, p.529 (2013).

森薫, 田中佳代子, 玉村雅敏, 金子郁容: “マークシート処理システムの拡張による学力調査データ分析方式”, 日本教育工学会論文誌, Vol.37, Suppl., pp.157-160 (2013)

森薫, 木幡敬史, 田中佳代子, 玉村雅敏, 金子郁容: “学校評価を対象とした調査問の透過的データ結合機能を有するデータ分析システムの構築”, 教育システム情報学会第 35 回全国大会 (2010)

高橋等, 松田稔樹: “マーク位置を任意に設定可能なマークシート読み取りソフトウェアの開発と簡易図形判定への適用”, 日本教育工学会論文誌, Vol. 31, Suppl., pp. 97-100 (2007)