

Web 情報を活用する蔵書管理システム ～パーソナライズへの適用～

塚本 哲也[†]西山 裕之[†]溝口 文雄[†][†]東京理科大学大学院理工学研究科

1 はじめに

本研究では実世界の蔵書を Web 上で一括管理する蔵書管理システムの提案を行う。一般に個人が持つ書籍は、保管場所が 1 箇所に集中していない、数が大量であるという理由から正確な所有状況の把握をするのは困難になっている。自分の所有状況を他人に伝えることや、他人の書籍の所有状況を知ることも容易ではない。

このシステムでは複数ユーザーの蔵書をそれぞれ電子的に管理する。システムに蔵書を登録することで、書籍の詳細な出版情報にユーザーが独自に定義した書籍所在地や貸出記録といった情報を付加して一括管理することができる。これによりユーザーは常に正確な書籍保有情報を得ることができ、他ユーザーの蔵書リストを閲覧することもできる。

また、蓄積した蔵書の出版情報(例。ジャンル、著者)からキーワードを集計し、個人の特性を明らかにすることでパーソナライズを実現するシステムを実装する。ユーザーの特性を用いて他のユーザーと比較することで、似た特性を持つユーザーを検索する機能を提案する。

2 蔵書管理システム

蔵書管理システムは従来、図書館や大型書店などで導入されているシステムである。個人で導入できる蔵書管理システムは少ない。また近年、書籍販売をする E-コマースサイトから書籍情報を取得できるサービスが公開されており、それを利用した蔵書管理システムもいくつが存在する。しかし、それらのシステムには書籍の登録方法が十分に用意されていないことや書籍を自由に分類できないことなどの問題点がある。

本論文では、既存の蔵書管理システムの問題点を解決するシステムを実装した。システム構成図を図 1 に示す。各ユーザーは蔵書管理の基本機能として書籍の登録、ユーザーが定義できる要素として書籍評価と本棚管理、登録した書籍を閲覧・検索することができる機能を使い蔵書を管理する。そうして得た複数の蔵書情報をパーソナライズに活用してゆく。

2.1 書籍登録

書籍を登録するには、ユーザーは書籍のバーコードを入力するかタイトルや著者といったキーワードを指定することでシステムに蔵書を追加できる。書籍登録

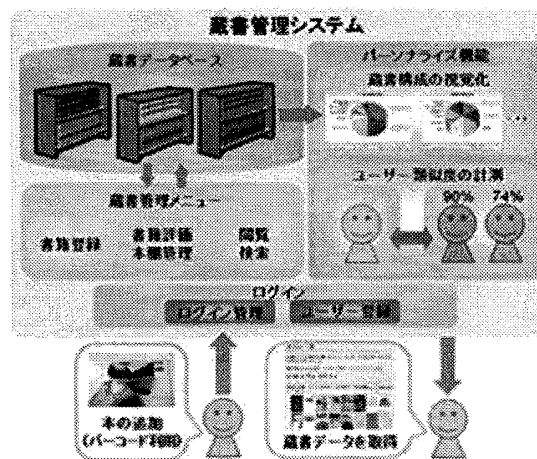


図 1: システム構成図

フローを図 2 に示す。書籍情報の取得には Amazon E-commerce Service 4.0 (Amazon ECS) のサービスを利用している。1 冊につき詳細な書籍情報(タイトル、著者、ジャンル、関連書籍など 13 項目)を取得し DB に蓄積する。システムはユーザーからバーコードを受け取ると ISBN コードを解析し、Amazon ECS から書籍情報を取得して、蔵書データベースに追加する。キーワード検索の場合、送信したキーワードを含む書籍の検索結果を Amazon ECS から取得し、目的の書籍を選択して蔵書データベースに追加する。

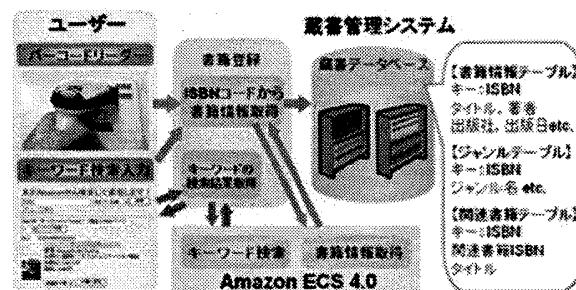


図 2: 書籍登録フロー

2.2 ユーザー定義情報

本システムでユーザーが独自に定義できる要素として、書籍の評価と本棚管理の 2 つがある。書籍の評価では 5 段階評価とユーザーのコメントを登録することができ、本棚管理ではタグを蔵書に付加することで本棚に整理するように管理できる。

Book management system using Web information applying personalizing.

Tetsuya Tsukamoto[†], Hiroyuki Nishiyama[†], Fumio Mizoguchi[†]

[†]Graduate School of Science and Technology Tokyo University of Science

2.3 閲覧検索

蔵書閲覧では、登録した蔵書を様々な条件でリスト表示する。ユーザーの蔵書一覧から、各書籍の詳細や他のユーザーの評価や書評までを閲覧できる。ユーザーが定義した本棚タグ、書籍表示方法（テキスト、イメージ、両方）、ソート（ABC順、追加日順）の3種の表示法を用意している。表示例を図3に示す。

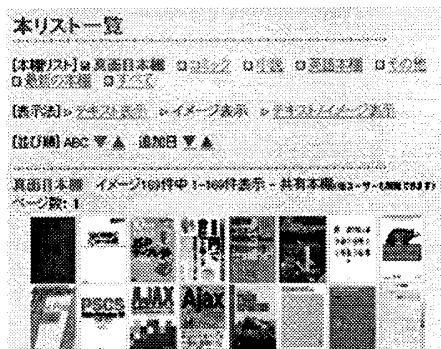


図3: 蔵書閲覧

2.4 パーソナライズへの適用

本システムはユーザーの蔵書に注目してパーソナライズを行う。これによりEコマースサイトに依存しないユーザーの書籍の購入傾向を読み取ることが可能になる。またユーザー間で蔵書の構成を比較することで、ユーザーを特徴付ける。

本論文ではパーソナライズ機能として、蔵書の構成の視覚化、ユーザー間類似度の算出について述べる。

2.4.1 蔵書構成の視覚化

蔵書構成を、書籍ジャンル、著者、出版社、金額の4つの属性について集計する。出力例が図4である。これがユーザーのパーソナライズの指標となる。本システム内でユーザーは蔵書の構成を数値、またはグラフで確認することができる。

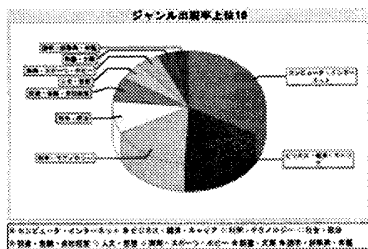


図4: 蔵書構成の視覚化例【ジャンル分布】

2.4.2 ユーザー間類似度の計測

ユーザー間の蔵書構成間の類似度を、ベクトル空間法[3]を適用して計算する。ベクトル空間法は、共通したベクトル空間の中で対象ベクトル間のコサインを算出することで類似度を計算する手法である。ユーザー間類似度を計算するために、各ユーザーの蔵書構成にある要素とその出現回数をベクトルとみなす。類似度は0.0から1.0の間の値を取り、1に近いほど似たジャンル構成のユーザーであるとわかる。

3 評価

3.1 従来システムとの比較

従来の蔵書管理システムは主に1ユーザーで利用するローカルタイプ（例、蔵書管理 Personal[1]）とWeb上で複数ユーザーが利用するWebタイプ（例、BOOKLOG[2]）の2つに分けられる。ローカルタイプには入力形式が基本的に手入力で蔵書を加える必要がある問題点があり、Webタイプにはユーザーの思い通りの蔵書のカテゴリ分けや、検索ができないという問題点がある。本論文での提案システムと従来のシステム（各タイプ2システムずつ）の機能の比較を行った結果が表1である。本システムは従来システムの問題点をすべて解決している。

表1: 従来の蔵書管理システムとの比較

	タイプ		本システム
	ローカル	Web	
バーコード入力	△	○	○
キーワード検索	×	△	○
書籍の分類	△	△	○
蔵書内検索	○	×	○

(○: 2システム可能, △: 1システム可能, ×: 両システム不可)

また、本システムとWebシステム型のサービス、BOOKLOG、と書籍追加テストを行った。このテストではシステムに異なる手法（バーコードリーダーを使う入力、ISBNの手入力、キーワードから検索して追加）で書籍を追加した経過時間の比較を行う。キーワードからの検索に関して本システムが3割ほど早いという結果が出た。

3.2 ユーザー間類似度の算出

システムを利用しているユーザー13名の蔵書構成分布（ジャンル）を利用して実際にユーザー間類似度を算出した。類似度の算出結果、0.9以上の値をとったユーザーペアが14通りあり、すべてのペアにて構成上位3位の項目が2つ以上一致している特徴があった。結果より、高い類似度をとったペアは構成上位の要素が類似しているといえる。構成上位3項目が含まれている場合と構成順位の1位が共通している場合で類似度が0.9を超えなかったペアの出現率は、それぞれ40%、30%であった。

4 おわりに

本論文では実世界での蔵書管理の問題点と従来の蔵書管理システムの問題点を解決する蔵書管理システムを実装した。Webから書籍情報を基にして、ユーザーの蔵書データの分析、ユーザー間類似度の算出を行った。これからユーザーの蔵書構成分布とユーザー間類似度を用いて、ユーザー間での書籍の推薦、既存の推薦書籍情報との融合を行い、蔵書を基にしたパーソナライズによる書籍推薦システムを実装する。

参考文献

- [1] 蔵書管理 Personal; <http://my.reset.jp/?triton/>
- [2] BOOKLOG; <http://booklog.jp/>
- [3] Salton, G.; McGill, M.J. Introduction to Modern Information Retrieval. McGraw-Hill.1983.