

# 職場利用のための図書管理システムの提案

西村 駿人<sup>1</sup> 野村 佳秀<sup>1</sup>

**概要：**本稿では、職場内での利用を想定した図書管理システムの提案を行う。多くの公立図書館や大学図書館で利用される Web-OPAC と呼ばれるインターネットを利用したオンライン蔵書目録検索システムではなく、大学の研究室や企業の部署内で利用するための小規模な図書管理システムを提案する。一般的な図書館との蔵書の規模や利用ユースケースの違いによる差分を列挙し、差分をもとに新たな図書管理システムを開発するために、要求獲得から、システムのプロトタイプの一部実装を行ったことを報告する。要求獲得では、以前より部署で利用していた図書管理システムを参考にしながら、改善点を挙げていき、また利用者にヒアリングを行うことで、要求を獲得する。システムのプロトタイプ実装では、図書管理システムを Web アプリケーションとして機能することを考慮し、また、社内のネットワークからの利用を想定し、実装を行う。特に、既存の図書分類法が機能しないという問題に着目し、カテゴリ分類と図書の配架の改善を実施し、評価を行った。

## Proposal of Book Management System in Use at Workplace

HAYATO NISHIMURA<sup>1</sup> YOSHIHIDE NOMURA<sup>1</sup>

### 1. はじめに

本稿では職場内での利用を想定した図書管理システムの提案を行う。職場の蔵書の特徴を踏まえて、新たな図書管理システムを開発するために、要求獲得からシステムのプロトタイプの一部実装を行ったことを報告する。

要求獲得では、以前より部署で利用していた図書管理システムを参考にしながら、改善点を挙げていき、また利用者にヒアリングを行うことで、要求を獲得した(2章)。システムのプロトタイプ実装では、図書管理システムを Web アプリケーションとして機能することを考慮し、また、社内のネットワークからの利用を想定し、実装を行った(3章)。特に、既存の図書分類法が機能しないという問題に着目し、カテゴリ分類と図書の配架の改善を実施し、カテゴリの適切さの評価を行った(4章)。利用ユースケースや図書管理システムの再構築の計画など、運用後のシステムのライフサイクルを今後の課題として述べる(5章)。関連技術・研究として、図書管理システムのオープンソース実装と、図書館情報学分野の分類問題に関する研究を挙げる

(6章)。

### 2. 想定するユースケースと要求

#### 2.1 システムを利用した蔵書管理

まず、システムを利用した蔵書管理について述べる。情報システムを利用して図書館の蔵書を検索・確認することは、2020年現在広く行われている。図書館資料の書誌情報や蔵書情報を電子化し、利用者がコンピュータを使って閲覧・編集できるシステムは、OPAC (Online Public Access Catalog) と呼ばれる。多くの公立図書館や大学図書館では Web-OPAC と呼ばれるインターネットを利用したオンライン蔵書目録検索システムが利用されており、このシステムを利用し、図書館内だけでなくインターネットを経由して、蔵書情報の検索や貸出依頼を実施することができる。日本図書館協会の2012年の調査[1]によると、都道府県立図書館、政令指定都市、特別区、県庁所在地の市の図書館はすべて Web 経由で蔵書検索でき、市区町村図書館でも9割近くで検索が可能であると報告されている。また公立図書館は地域の住民や通勤通学者の利用を想定しており、大人から子供まで多くの利用者に対して、幅広い分野の蔵書を提供するというユースケースがあると考えられる。

<sup>1</sup> (株)富士通研究所  
FUJITSU LABORATORIES LTD.

## 2.2 職場内で蔵書管理を行うユースケースと現状のシステム構成

次に、今回対象にする蔵書の概要について述べる。筆者らが所属する組織では、企業内の図書館とは別に部署で蔵書を保有しており、本稿では部署内で利用する図書管理システムを提案する。本稿で想定する職場内での蔵書管理は、公立図書館や大学図書館と比較するといくつか異なる点がある。例えば、職場内の蔵書は専門書など特定の分野に偏りがあることや、部署の規模に依存するが利用者は100名未満と考えることができる。このような特徴は、企業の部署に限らず、大学の研究室においても似た状況であると考えられる。職場内の図書管理システムを取り巻く状況を列挙すると、以下のような特徴がある。

- 厳密な貸出管理やユーザ認証を必要とせず、公立図書館で利用されているシステムと比較すると少機能でよい
- 司書の業務は、部署の構成員が本業の合間に実施する必要があり、過剰なコストをかけることが難しい
- 専門書が多いため書籍に分野の偏りがあり、配架において既存の図書分類法が機能しにくい

筆者らが所属する部署では約800冊の書籍があり、複数存在する物理的な本棚に書籍は格納されており、書籍の貸出はWebシステムを利用する。書誌情報として物理的な本棚ごとにテーブルを持ち、各レコードに対して、タイトル、著者、ISBN、貸出状況を記録しており、そのレコードについて、操作したユーザとその日時を貸出記録として追加・編集するという方法で管理をしている。Webシステム上では、書籍の一覧表示、貸出・返却、新規書籍の登録や編集、削除が可能となっている(図1)。現状のシステムは、ユーザがアクセスするWebサーバがあり、管理のためのアプリケーションサーバがあり、DBが存在するという3層のアーキテクチャで構成されている(図2)。新たな図書管理システムを作成するために、要求獲得を行った。現状の図書管理システムに関する問題点や改善点を募集したところ、以下の点が挙げられた。

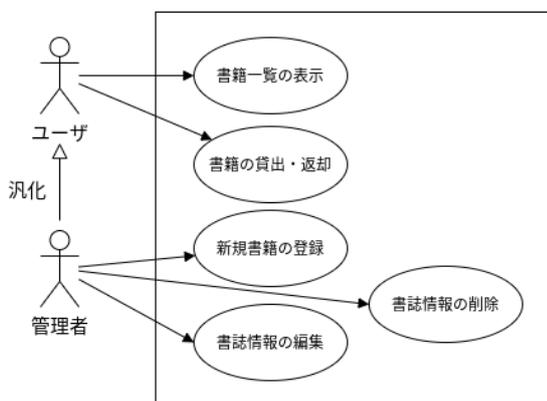


図1 Webシステムにおけるユースケース図

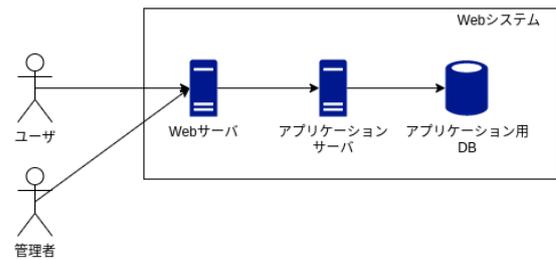


図2 Webシステムにおけるシステム構成図

- 図書管理システム利用方法の理解が容易ではないため、わかりやすくなってほしい
  - 部署内のWikiシステムの拡張として作成されているため、独立したシステムがほしい
  - 貸出方法がISBNの手入力のため、入力が容易になってほしい
  - 書誌情報の主キーとなるIDがISBNのため、重複する書籍の考慮がされていない
  - ユーザ認証の仕組みがあるとよい
  - 現在は最低限の検索のため、探しづらい。検索機能を改善してほしい
  - 物理的な棚ごとの配置情報はあるが、複数の本棚に同一ジャンルの書籍が配架されており、書籍のジャンルが統一されていないため、整理してほしい
- また管理運用者の立場としての、現行システムの問題点や改善点は以下である。
- 書誌情報登録のための独立したスクリプトは存在するので、アプリケーションに組み込んでほしい
  - 管理運用者が実施する蔵書の管理のための管理画面がほしい
  - 数年に一回ある書籍の棚卸しに対応する機能がほしい
  - 部署の異動に伴った図書の移動に対応する機能がほしい

上記の点や既存の図書管理システム実装をふまえ、今回はシステムのプロトタイプ実装をWebアプリケーションとして実施した。

## 3. プロトタイプの実装とユーザのフィードバック

システムのプロトタイプ実装では、図書管理システムをWebアプリケーションとして機能することを考慮し、また、社内のネットワークからの利用を想定し、実装を行った。以下では、プロトタイプの画面と機能を一部紹介し、そこから得られたフィードバックを記載する。

図3は、プロトタイプ実装として作成したWebアプリケーションの書籍一覧の画面である。書籍一覧は表の形式で提供されており、ユーザは自身が利用したい書籍を確認し、行の右に配置された「借りる」「返す」のボタンをク

リックすることで、貸出画面や返却画面へ遷移する。

図4は、書籍貸出の画面であり、ユーザは自身の名前を入力して、“submit” ボタンをクリックすることで、貸出処理を実施することができる。

従来の実装との変更点を述べる。まず、他アプリケーションの機能ではなく、独立したアプリケーションとして実装を行った。また、主キーがISBNから、ユーザ画面からは不可視であるIDを主キーとしたことにより、同じISBNを持つ同名の書籍を複数管理することができるようになった。これは、問題点として挙げられていた項目を改善して実装したものとなる。しかし、同じISBNを持つ書籍がある場合は、物理的な書籍側にもIDに対応する追番や別途IDを記載して、書籍とレコードの同一性を保つ必要がある。

プロトタイプの利用から得られたフィードバックを以下に記載する。貸出画面では以前の実装を参照して、ユーザ名はテキストの自由入力としたものの、利用者の意見として、社内の認証基盤との連携の可能性を指摘された。また、ユーザごとのページがほしいという意見や、ユーザ自身が貸出中の書籍一覧を確認したいという意見も挙がった。また、ハンディ型のバーコードリーダーを利用した、容易な貸出・返却のフローを実施したいという希望もあり、その場合にはスキャンして得られたISBNから書籍の主キーを絞り込み、貸出画面に遷移するという機能実装の必要がある。

## 4. 機能検討：本棚配置のためのカテゴリ分類

本章では、既存の図書分類法が部署や研究室の専門書が多いという状態では機能しないという問題に着目し、適切な本棚の配置を実施するために、自動カテゴリ分類器の実装をしたことを報告する。

### 4.1 カテゴリ分類機能検討の背景

現行のシステムではいくつか改善したいポイントが挙げられていたが、そのうち、物理的な本棚において、関連する蔵書が近くに配架されていないため、ある書籍に関連する書籍が簡単に探せないという問題や、新着図書を適切な



ISBN	Title	Authors	Shelf ID	Last Borrower	Last Updated	Status	Edit
9784764903241	要求分析	本位田真一, 田原康之	4	野村	2020-01-01	貸出中	送す
9784764903258	VDM-SLLによる形式仕様記述	田口研治, 柴田啓伸	4	西村	2020-02-17	貸出可能	送す
9784764903265	アジャイル開発テキスト	安藤洋, 田中秀樹	4	高島	2020-05-14	貸出中	送す
9784764903272	コンポーネントベース開発テキスト	廣城弘章, 丸山勝久, 山本聖枝子, 本位田真一	4	佐伯	2020-05-14	貸出中	送す

図3 プロトタイプ実装の書籍一覧画面の抜粋



ISBN	Title	Authors	Shelf ID	Last Borrower	Last Updated	Status
9784764903258	VDM-SLLによる形式仕様記述	田口研治, 柴田啓伸	4	西村	2020-02-17	貸出可能

山田

SUBMIT キャンセル

図4 プロトタイプ実装の書籍貸出画面の抜粋

棚に配架する判断が難しいという問題が、優先順位の高い問題として、利用者と管理運用者に認識されていた。カテゴリ分類は専門知識が必要ではあるが、低コストに配架業務を行うことを目指すことは、部署や研究室内の図書管理においては重要であると考えられる。そこで、本章では書籍に対して既存の分類をもとに、自動で分類を付与できる分類器を作成し、書籍への分類を自動付与できる仕組みを作成したことを報告する。

日本十進分類法 (Nippon Decimal Classification ; NDC) は、国立国会図書館や公立図書館で広く利用されている図書分類法である。国立国会図書館は日本で刊行された書籍をもれなく収集しており、その書籍に対してNDCを付与しているため、国立国会図書館にWeb経由で問い合わせることで、分類を得ることが可能である。公立図書館などの大規模で利用者層が幅広い図書館においては、この分類は有効に機能すると考えるが、専門書の分類、特に情報科学の書籍の分類においては、更に細かい独自の分類が必要になると考える。本稿では、専門書に特化した分類について、筆者らの所属する組織の企業内図書館で利用されている既存の細分類 (17分類) を、正解データとして使用して実験と評価を実施したことを報告する。

### 4.2 カテゴリ分類機能の実験に利用したデータセットと評価尺度

本実験で利用した加工前のデータの内訳は、部署内で収集し管理している書籍である829件と、企業内図書館に登録されている専門書籍のうち、近年一定の年数分を抽出した3,315件である。そのうち、国立国会図書館サーチ<sup>\*1</sup>を利用して、国立国会図書館のデータベースに登録されている出版社情報やキーワードなどの書誌情報を補完した。最終的に利用したデータは、企業内独自の分類と国立国会図書館内でキーワードが付与されていた、1,347件 (企業内図書館にのみ存在) と367件 (部署内と企業内図書館に存在) である。企業内図書館にのみ存在するデータで分類器の検証を行い、部署内に存在するデータで評価を実施した。

本実験に利用した評価指標は、全分類結果中、正しく判定した数である正解率と、各カテゴリにおける適合率、再現率、適合率と再現率の調和平均であるF1値を利用した。

### 4.3 カテゴリ分類機能の実験

分類器を作成するために利用したアルゴリズムは線形SVMを利用し、タイトルの中の名詞、キーワード、出版社を特徴量として利用した。名詞の抽出方法として、形態素解析器MeCab<sup>\*2</sup>を利用し、辞書としてipadicとmecab-ipadic-NEologd<sup>\*3</sup>を利用した。

\*1 <https://iss.ndl.go.jp/>

\*2 <https://taku910.github.io/mecab/>

\*3 <https://github.com/neologd/mecab-ipadic-neologd>

まず分類器の妥当性を確認するために、内部データでの検証として、企業内図書館にのみ存在する書籍を利用し、5分割交差検定を実施したところ、平均正解率は81.1%であった。次に分類器の評価として、学習は企業内図書館にのみ存在する書籍全件、評価を部署内にも存在する書籍で実施したところ、正解率は65.7%であった。個別のカテゴリの結果については、次節で述べる。

#### 4.4 カテゴリ分類機能における実験の考察と今後の課題

個別のカテゴリの分類結果を表1に示す。まず、システム開発とソフトウェア工学を混同した誤分類が多かったことが挙げられる。頻出したキーワードを調査すると、どちらのカテゴリにも、「システム開発」や「システム設計」といったキーワードが頻出していることが原因として考えられる。次に認知科学とプログラミングに関するカテゴリについては誤分類が少なかった。これも同様に頻出したキーワードを調査すると、認知科学には、「機械学習」や「深層学習」、プログラミング言語には「プログラミング」や「コンピュータ」といったキーワードが挙げられており、どちらも各カテゴリに独立したキーワードであったため、誤分類が少なかったと考えられる。

今回の実験ではテスト対象のデータセットに対しては、正解率は65.7%と決して高い数値ではなかった。今後の課題としては、精度向上のために、学習に利用するデータ数の増加や、複数の特徴量の検討、複数の分類器の選択などの実験の試行錯誤の実施、既存のカテゴリの再検討や、さらなる教師データの検討などを実施したい。

### 5. 今後の課題

本章では利用ユースケースや図書管理システムの再構築の計画など、運用後のシステムのライフサイクルを今後の課題として述べる。まず利用ユースケースの説明のため、カテゴリ自動付与機能を実装した後のシステム構成図を図5に示す。図2の現状のシステムに加えて、今回実施した実験を行う分析環境を追加した。また実験の出力として得られる分類器を機能の一部として提供するカテゴリ分類サーバを、Webシステムの中に新たに組み込むことで、アプリケーションからカテゴリの分類結果を得ることができる。この機能追加により、例えば書籍を新たに購入し、蔵書一覧に追加する際にカテゴリ情報を付与することが可能となる。

表1 カテゴリ別の結果(抜粋)

カテゴリ	適合率	再現率	F1値
システム開発	0.60	0.57	0.58
ソフトウェア工学	0.68	0.59	0.64
認知科学	0.87	0.84	0.80
プログラミング言語	0.77	0.84	0.80

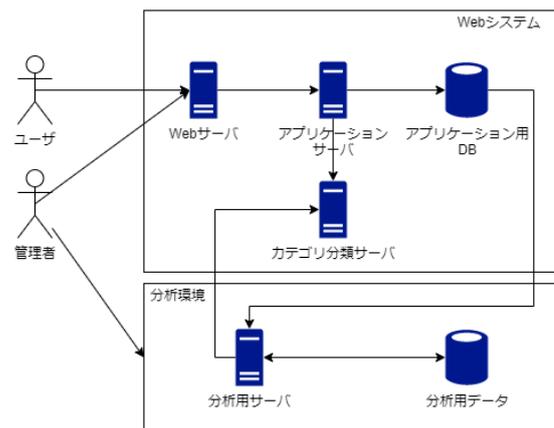


図5 Webシステムにおけるシステム構成図

今回実装の検討を行わなかった機能としては、大きくユーザ認証、蔵書検索、管理者向け機能がある。蔵書の検索は、利用者の利便性に大きく寄与すると考えるため、実装したい。また、棚卸しのための機能や蔵書の追加や削除といった管理者向けは、実装が容易であるが、利用者からフィードバックが広く得られにくいいため、今回のプロトタイプ実装では実施しなかった。ユーザ認証については、アプリケーションの利便性に関係してくるため、実装にあたっては、考慮する必要があると考えている。

今回は図書のカテゴリ分類の機能に着目して機能検討を実施したが、今後はこの機能の精度向上だけでなく、他に挙げられた機能の実装も行っていきたい。

### 6. 関連技術・研究

図書を管理するシステムは、大規模なものから小規模なものまで多数の実装・サービスが存在する。現在の公立図書館においては、ITベンダー各社が開発したOPACシステムをパッケージとして導入しているが、本稿では、部署内や大学研究室内で利用する場合に着目しているため、そのような環境でも利用可能かつ著名であると思われるオープンソースソフトウェアを2件紹介する。原田は、Next-L Enju<sup>\*4</sup>が日本初のオープンソース統合図書管理システムであることを報告している[2]。Next-L Enjuは、ITベンダーではなく図書館に関わる人々が中心となって2008年から開発を開始し、その実装はオープンソースソフトウェアとして公開され、国立国会図書館や物質・材料研究機構での導入実績があると報告している。Blacklight<sup>\*5</sup>は、バージニア大学で開発が開始した、オープンソースソフトウェアのApache Solrのインタフェースであり、コーネル大学やブラウン大学など、米国の大学図書館でOPACとして導入されている。

図書館情報学分野では、図書に関するカテゴリの自動

<sup>\*4</sup> <https://www.next-l.jp/>

<sup>\*5</sup> <http://projectblacklight.org/>

付与問題に取り組んでいる。以下では、図書やレファレンス事例を、図書分類法である日本十進分類法 (Nippon Decimal Classification ; NDC) の分類カテゴリに自動分類した研究を紹介する。

石田らは図書の自動分類を行う際に、複数の手法を比較することでどの手法が有効かを検討し、また分類に利用する情報として、それまで利用されていた書名だけではなく、目次や帯情報を利用することの有用性を検討した [3]。また、NDC の 0 類から 9 類までの類 (10 分類) それぞれについて実験を行うことで、どの情報が各類において有用かを検討している。宮田らは、NDC の階層構造を利用した図書の自動分類に取り組んだ [4]。NDC のカテゴリは、さらに細かい分類であるサブカテゴリを持つという性質に着目し、0 類から 9 類までの類 (10 分類) で自動分類した後に、更に細かい分類である細分類までの自動分類を実施するという手法 (二段階の手法) を実施した結果、一度に細分類までを自動分類するという従来手法と同程度の結果を得ている。ただし 10 分類の結果においては、一度に細分類まで行ってしまう従来手法よりも、二段階の手法が高再現率であることを報告している。

荒井・辻は、国立国会図書館が提供しているレファレンス協同データベースに登録されているレファレンス事例に対して、機械学習を用いて NDC を自動付与するというタスクを実施した [5]。レファレンス事例とは、図書館員が図書館利用者から問い合わせを受けた質問とその回答をまとめたものである。3 種類の手法 (質問文から特徴ベクトルを作成する手法、参考資料に挙げられた情報から図書検索をし、その図書の NDC を用いる手法、両者を組み合わせた手法) を実施したところ、質問文から特徴ベクトルを作成する手法が、最も高い適合率・再現率であった。また質問文以外に、レファレンス事例のキーワード、レファレンス事例のタイトルから得られる特徴量を使って特徴ベクトルを作成し実験したところ、質問文、キーワード、タイトルの順に適合率・再現率が高かったことを報告している。

## 7. おわりに

本稿では職場内での利用を想定した図書管理システムの提案を行った。

職場の蔵書の特徴を踏まえて、新たな図書管理システムを開発するために、要求獲得からシステムのプロトタイプの一部実装を行ったことを報告した。要求獲得では、以前より部署で利用していた図書管理システムを参考にしながら、改善点を挙げていき、また利用者にヒアリングを行うことで、要求を獲得した。システムのプロトタイプ実装では、図書管理システムを Web アプリケーションとして機能することを考慮し、また、社内のネットワークからの利用を想定し、実装を行った。プロトタイプの実装を踏まえて、さらなる改善点に関する意見の列挙した。特に新たな

機能について、既存の図書分類法が機能しないという問題に着目し、カテゴリ分類と図書の配架の改善を実施し、カテゴリの適切さの評価を行った。利用ユースケースや図書管理システムの再構築の計画など、運用後のシステムのライフサイクルを今後の課題として述べた。

## 参考文献

- [1] 日本図書館協会: 公共図書館 Web サイトのサービス, 入手先 (<http://www.jla.or.jp/link/link/tabid/167/Default.aspx>) (参照 2020-03-16).
- [2] 原田隆史: Project Next-L と Next-L Enju 日本初のオープンソース統合図書館システムの開発と現状, 情報管理, Vol.53, No.11, pp.725-737 (2012).
- [3] 石田栄美, 宮田洋輔, 神門典子, 上田修一: 目次と帯を用いた図書の自動分類, 情報処理学会研究報告, DD, [デジタル・ドキュメント] 2006(33), pp.85-92 (2006).
- [4] 宮田洋輔, 石田栄美, 神門典子, 上田修一: NDC の階層構造を利用した図書の自動分類の試み, 2006 年度日本図書館情報学会春季研究集会発表要綱, pp.51-54 (2006).
- [5] 荒井俊介, 辻慶太: 機械学習を用いたレファレンスデータへの NDC の自動付与, 情報知識学会誌, Vol.25, No.1, pp.23-40 (2015)