

掲載誌: 『薬学図書館』 第 567 第 2 号 (2012 年 4 月 30 日発行)

Vol. 57 No. 2 p. 101-115, 2012

(日本薬学図書館協議会より許可を得た上で転載)

アカデミックとリアルの谷を埋める道

－ 「マイニング探検会」 成果報告 －

The Way to Fill the Gap between the Academic World and the Real World - A Debriefing Session of *Mining Exploration Group* -

清田 陽司 (株式会社ネクスト リッテル研究所)

Yoji Kiyota (Littel Laboratory, NEXT Co., Ltd.)

KiyotaYoji@next-group.jp

三津石 智巳 (筑波大学 大学院図書館情報メディア研究科)

Tomomi Mitsuishi (Graduate School of Library, Information and Media Studies, University of Tsukuba)

tomomi@klis.tsukuba.ac.jp

関戸 麻衣 (国立情報学研究所)

Mai Sekido (National Institute of Informatics)

skdmai@nii.ac.jp

日高 崇 (有限会社スタジオ・ポット SD)

Takashi Hidaka (STUDIO POT SD Inc.)

hidaka@pot.co.jp

岡本 真 (アカデミック・リソース・ガイド株式会社)

Makoto Okamoto (Academic Resource Guide, Inc.)

mokamoto@arg-corp.jp

【抄録】 本稿では、2011 年 11 月の第 13 回図書館総合展にて開催されたフォーラム「アカデミックとリアルの谷を埋める道 - 知識情報社会でライブラリアンはどうあるべきか -」の内容を紹介する。本フォーラムは、ライブラリアンを中心した情報技術勉強会である「マイニング探検会」の成果報告会という趣旨で開催されたものである。本稿の前半では、マイニング探検会で行った開発合宿の成果物として、レファレンス協同データベースを用いたゲームアプリケー

ション、図書館 Web サイトの訪問者を支援するプラグインツール、機械学習アルゴリズムを利用した論文のフィルタリングツールの 3 種類のアプリケーションを紹介する。本稿の後半では、ライブラリアンが開発合宿のような活動に参加する意義を述べるとともに、「アカデミック」と「リアル」の間のギャップを埋めるため、これから図書館の未来を作っていくライブラリアンがどのような在り方を求められているのかについての考察を述べる。

[キーワード] 情報技術, 開発合宿, レファレンス協同データベース, 図書館 Web サイト, 機械学習, 学問領域, 実用領域

[Author Abstract] This paper describes the overview of the forum “The way to fill the gaps between the academic world and the real world”, held at *The 13th Library Fair & Forum* in November 2011. The forum was held as a debriefing session of *Mining Exploration Group*, a workshop for librarians, focused on Information Technology. The early part of this paper introduces three products of the development camp of *Mining Exploration Group*: a game application based on the Collaborative Reference Database (CRD), a plug-in tool for supporting visitors of web sites of libraries, and a filtering tool of papers based on machine learning techniques. The latter part of this paper describes the value of activities such as development camps by librarians, and addresses the way of librarians who aim to develop the future of libraries, focused on the gaps between the academic world and the real world.

[Keywords by Author] information technology, development camp, collaborative reference database, web sites of libraries, machine learning, academic world, real world

1. マイニング探検会とは

マイニング探検会（通称マイタン）は、図書館や出版業界など、知識情報産業に携わる参加者を中心として、情報マイニング技術にフォーカスして新たな図書館の未来を探ることを目的とした勉強会である。清田と岡本の両名を発起人として 2010 年 4 月に発足し、概ね月 1 回のペースで開催している。マイニング

探検会の開催にあたっては、不動産ポータル HOME'S を運営する株式会社ネクストの研究開発拠点であるリッテル研究所、および東京大学情報基盤センター学術情報研究部門の協力を得ている。

本稿では、2011 年 11 月に開催された図書館総合展において、マイニング探検会の成果発表会という趣旨で開催したフォーラム「アカデミックとリアルの谷を埋める道 - 知識情報社会でライブラリアンはどうあるべきか -」の内容を紹介する。マイニング探検会では、2011 年 4 月頃より成果発表に向けて準備を進めてきており、2011 年 7 月末に神奈川県真鶴町で 2 泊 3 日の開発合宿を開催し、4 チームに分かれて持ち寄ったアイデアをもとにアプリケーションの開発に取り組んだ。この開発合宿は、「共生プログラミング」¹⁾という開発手法の実践として行った。2 節では、3 チームがそれぞれ開発したアプリケーションの概要を紹介する。3 節では、フォーラムでの議論の内容を踏まえて、マイニング探検会という勉強会の意義、およびライブラリアンが開発合宿のような活動に参加する意義について整理する。4 節では、「アカデミックとリアル」「組織と個人」「ルールとガイドライン」という軸を中心に、図書館の未来を作っていく私たちがどのようなあり方を求められているのかについて考察していく。最後に、マイニング探検会が今後目指す方向性を示す。

2. マイニング探検会合宿@真鶴の成果報告

本節では、開発合宿に参加した 4 チームのうち、「Ref.Master チーム」「ブクリスチーム」「俺 CiNii チーム」の 3 チームの成果を紹介する。本開発合宿には、マイニング探検会のメンバーのうち 10 名が参加した。また、株式会社ネクストリッテル研究所のメンバー 5 名（清田を含む）がエンジニアとして参加した。

本開発合宿で実践した「共生プログラミング」という開発手法は、サービスの提供者・エンジニア・ユーザの立場の違いを超えて、課題を共有する人同士が集まり、互いの専門性を生かして相互に貢献し、具体的な解決策を創造し、実際に動くものを作るものであると定義されている¹⁾。今回は、主にマイニング探検会メンバーはサービスの提供者やユーザ、リッテル研究所のスタッフがエンジニアとしての役割を担当したが、メンバーによってはエンジニアリングのスキルを持っている方もいて、必ずしも役割は固定したものではない。また、2 泊 3 日の合宿の最終日にはいったん各チームからの成果発表会を行ったが、合宿終了後も各チームのメンバーが自主的に集まり、完成度を上げる取り組みを

行った。

以下では、3 チームの代表者それぞれからの発表内容の概要を紹介する。

2.1. Ref.Master (三津石)

Ref.Master は、レファレンスサービスを支援することを目的として制作したレファレンスのスキル向上ツールである。具体的には、レファレンス協同データベースの API をデータソースとして利用した Web ベースのゲームアプリとして開発しており、ユーザに「楽しみながら」レファレンスに必要なスキルを身につけてもらうことを狙いとしている。ユーザは、「遊びながら」レファレンスに必要なキーワードやレファレンス資料を手軽に学習することができる。また、ユーザのモチベーションを維持するための工夫としてレファレンスレベルという概念を持たせており、学習達成率に応じてレベルアップしていく仕組みを実装している。

2.1.1. 開発の背景

図書館の現場において、個人のレファレンス能力をどのように向上させていくかは大きな課題となっている。レファレンス能力を継続的に向上していくには、新しいレファレンス資料を熟知することが求められるが、実際にレファレンス業務の現場で受け付ける質問と、新たに身につける知識は直接結びつかない。また、レファレンス質問に回答するには長い時間を要することが多く、レファレンス資料について学んでいくための時間がとれていないこともある。そのため、レファレンス資料に関する知識を身につけるためのモチベーションが続き、レファレンス能力がなかなか向上しないという現象が起きているのではないかと考えている。

そこで、様々な館種のレファレンス質問とその回答のデータベースを蓄積しているレファレンス協同データベースの API を活用することで、レファレンスに関する知識を学ぶモチベーションを維持できる仕組みを持ったレファレンス演習用のツールを作れないかという着想をもとに、マイニング探検会合宿での開発に取り組んだ。

2.1.2. 主な機能

合宿では、限られた時間でレファレンス演習用ツールを実現するため、可能な

限りシンプルな実装を心がけた。具体的には、本質的な以下の機能に絞り込むことで、開発に要する工数を最小限に抑えることと、ユーザにシステムのコンセプトが伝わることを両立させた。

- **OpenID による認証**

それぞれのユーザのレベルを管理する必要上、ユーザのログインの仕組みをもたせることは必須であるが、今回は独自のログイン機構は実装せず、OpenID による認証の仕組みを採用した。ユーザは、アカウント取得の煩雑な手続きを経ることなく、FaceBook、Google、Twitter などの既存の Web サイトのアカウントを利用することで、Ref.Master を継続的に利用することができる。

- **ユーザ用ポイントデータ管理**

Ref.Master には、他の多くのゲームアプリと同様に、ユーザがゲームを遊んだ実績をポイントとして管理している。ポイントがたまっていくことで、ユーザのレベルは「図書館採用」→「見習い司書」→...→「トップレファレンサー」と上がっていく。ちなみに最高レベルは「レジェンド」として

- **ジャンルに絞った事例選択**

レファレンス協同データベースの回答事例には日本十進分類（NDC）が付与されていることを利用し、ユーザに 10 ジャンル（NDC の最上位層の 10 カテゴリに相当）のうち遊びたい一つのジャンルを毎回選択させることで、ユーザは学びたいジャンルを絞り込んで遊ぶことができる。

2.1.3. 開発メンバー

Ref.Master の開発チームには、三津石のほか、外崎みゆき氏（関東学院大学図書館）、河村俊太郎氏（東洋大学／東洋英和女学院大学）がメンバーとして参加した。また、リッテル研究所のエンジニア 2 名（中塚寛幸氏、愛宕翔太氏）が技術支援スタッフとして参加した。

2.1.4. デモ

Ref.Master は、<http://refmaster.litteldev.net/> にアクセスすることで利用できる。Ref.Master にログインすると、図 1 に示すように、NDC 最上位層に相当する 10 カテゴリとそれぞれのカテゴリの現在のレベルが表示される。ここで、

例えば遊びたいジャンルとして「総記」を選択すると、総記分野についての質問とライブラリアンの回答にあわせて、ライブラリアンが検索に利用したキーワードが4択問題として出題される(図2)。一定の問題数を回答し、正解率が一定のレベルをクリアすると、上のレベルに昇格できる(図3)。

2.1.5. 今後の展望

今後は、Ref.Masterを多様な人々に実際に利用してもらい、そのログデータを大量に収集することがひとつの目標である。ログデータを収集し、解析することでデータソースとして利用しているレファレンス協同データベースに対して課題を提供し、フィードバックを行うことができる。例えば、ユーザレベルの高い複数のユーザが、同じ問題において誤回答を入力している場合、レファレンス協同データベースの提供している回答が不十分である可能性がある。これをフィードバックすることでデータベース品質向上に貢献できる。また、レファレンス共同データベースから取得したデータに対してマイニング技術を用いることにより、キーワードから分類番号、分類番号から参考資料などをナビゲートしたり、レファレンサー支援機能の機能を提供したりする方法も考えられる。

Refmaster

ようこそ (匿名) さん

[プロフィール変更](#) [ご意見箱](#)

問題カテゴリー一覧

000 総記 Lv.: 図書館採用(1)	100 哲学 Lv.: 図書館採用(1)
200 歴史 Lv.: 図書館採用(1)	300 社会科学 Lv.: 図書館採用(1)
400 自然科学 Lv.: 図書館採用(1)	500 技術・工学 Lv.: 図書館採用(1)
600 産業 Lv.: 図書館採用(1)	700 芸術美術 Lv.: 図書館採用(1)
800 言語 Lv.: 図書館採用(1)	900 文学 Lv.: 図書館採用(1)

図 1 Ref.Master のトップ画面

Refmaster

ようこそ 三津石智巳 さん

プロフィール変更

ご意見箱

総記 図書館採用

あと3問

問題:

埼玉県比企郡小川町の和紙に関する最近の統計(生産量、工場数等)を探している。

ライブラリアンの回答:

『統計おがわ 平成8年版』および小川町ホームページに掲載されている情報を提供し、「統計おがわ 平成13年版」(小川町図蔵)、及び類縁機関として埼玉伝統工業会館、小川和紙資料館、小川町役場を連絡する。

キーワード選択肢:

- 過疎対策・過疎問題
- フォークダンス
- Hearn Lafcadio (ハーン ラフカディオ)
- 伝統工芸-小川町-比企郡-埼玉県

解答チェック

図 2 Ref.Master の出題画面

Refmaster

ようこそ 三津石智巳 さん

プロフィール変更

ご意見箱

昇格

おめでとうございます！
総記で、レベルが上がりました。
次のレベルは、あと8問を正解する必要があります。
頑張ってください！

カテゴリ選択画面でレベルを確認する

図 3 Ref.Master のレベルアップ

2.2. ブクリス：ニュース×カーリル×OPAC+CiNii Books、『気になる本』と出会える図書館サイト（関戸）

ブクリスは、主に図書館の Web サイトを訪問するユーザを対象として、「旬の話題」のブックリストを自動的に生成・提示するシステムである。本との出会いを求めて図書館の Web サイトを訪れるユーザにとって、OPAC への検索キーワード入力を求めるのはハードルが高すぎるという問題意識が、開発の出発点である。多くの公共図書館で実際に使ってもらうことを目標に開発している。

2.2.1. 開発の動機

館種を問わず、多くの図書館の Web サイトでは、OPAC の検索キーワードを入力することができるフォームを設置し、蔵書を検索するというユーザのニーズに応えている。しかし、図書館 Web サイトを訪問するのは、必ずしも明確な情報要求をもったユーザだけではない。情報要求は曖昧であるものの、新しい本との出会いを求めて Web サイトを訪れるユーザも多いはずである。後者のタイプに属するユーザにとって、最初にキーワードの入力を求められるのはハードルが高い。すなわち、現在の標準的な図書館 Web サイトの作り方では、後者のタイプのユーザのニーズに十分応えられていない可能性がある。

一方、図書館によっては旬の話題や地域の話題にマッチしたブックリストをライブラリアンが定期的に作成し、本との出会いを求めているユーザのニーズに応える努力をしているところもある。しかし、定期的にブックリストを更新するにはそれなりの人手がかかるため、実現できているところはそれほど多くない。

そこで、新聞社のニュースサイトなどの情報源から「旬の話題」を抽出し、OPAC を検索することで自動的にブックリストを定期的に生成するシステムを実現し、本との出会いを求めるユーザのニーズに応えることを目標として開発を行った。開発にあたっては、多くの図書館で簡単に導入でき、難しい設定作業や高度な運用ノウハウを必要としない仕組みの実現を心がけた。

2.2.2. デモ

ブクリスのデモは、<http://maitan-a.litteldev.net/> で見ることができる。中津川市立図書館の Web サイトへの導入イメージを図 4 に示す。「こんな本はいかが

ですか」の欄に、新聞社ニュースサイトの RSS フィーダーを利用して抽出された「旬のキーワード」のうち、ランダムで選択されたキーワード（図中では「日本画家」）について、中津川市立図書館の OPAC での検索結果をブックリストとして表示している。



図 4 ブクリスのデモサイト（中津川市立図書館バージョン）

2.2.3. システム構成

ブクリスの基本的な機能は、ニュースサイトなどの情報源からキーワードのリストを作り、「何らかの処理」を行って各々のキーワードについての資料リストを生成することである。ここでいう「何らかの処理」をどのように実現するかが、ブクリス開発における重要なポイントであった。

合宿当初は、各図書館の蔵書検索結果をどのような方式で得るかが大きな課題となった。各図書館の OPAC を直接呼び出し、スクレイピングの手法で検索結果を抽出することで、技術的には実現可能であるものの、この方法では各々の図書館への対応に多大な工数がかかることが見込まれた。そこで、Nota, Inc が提供しているカーリルローカルの API を利用することで、どの図書館でも簡

単に導入が出来るような仕組みとした。カーリルローカルの API は、キーワードと対象図書館の ID を入力として与えると、検索結果(10 件)と各資料の OPAC の URL を出力として得る機能を有している。合宿中に実装したシステムの構成図を図 5 に示す。なお、各々の情報源からのキーワード抽出には言選 Web を利用している。

しかし、実際に図 5 のシステムを動かしてみたところ、以下に挙げる問題点が明らかになった。

- 一つの情報源から抽出されるキーワード候補すべてについて、各々の図書館の OPAC に検索リクエストが行われてしまう。具体的には、一つの情報源につき、100 回以上の検索が行われるため、各図書館の OPAC に高い負荷をかけてしまう。また、多数の図書館を対象とする場合には著しく処理効率が悪くなってしまう。
- 言選 Web の特性として、ある程度長く専門性の高いキーワードが抽出される傾向がある。このようなキーワードにヒットする資料が存在する場合はきわめて有用であるものの、このようなキーワードで OPAC を検索した場合はヒット件数が少なくなりがちであり、ゼロ件ヒットになってしまう場合も多い。一方、一般的な形態素解析ツールを使って短いキーワードを抽出すると、多くのヒット件数が得られるものの、そのようなキーワードの大部分は一般的すぎて面白い検索結果にならないことが多い。

そこで、合宿終了後にチームメンバーが何回か集まり、キーワード選択の手法についての再検討を行った。その結果、CiNii Books (NACSIS-CAT のデータ) の API を利用し、CiNii Books を検索して得られるヒット件数を共通フィルタとして利用する方法を採用することになった。見直し後のブクリスのシステム構成図を図 6 に示す。なお、キーワード選択の手法としては、件名などのシソーラスに出現するキーワードに絞る、Web 検索エンジンのヒット件数を利用する、シソーラス等によるキーワード展開を行うなどの手法も検討している。

また、どこの図書館でも利用できることを目指し、各々の図書館の担当者が簡単に「利用する情報源」「利用するキーワード」「設定するフィルタの種類および閾値」などを設定できる Web 管理画面 (ブクリス Manager) も実装している。ブクリス Manager のトップ画面、および利用するキーワードの選択画面のイメージを、図 7 および図 8 に示す。

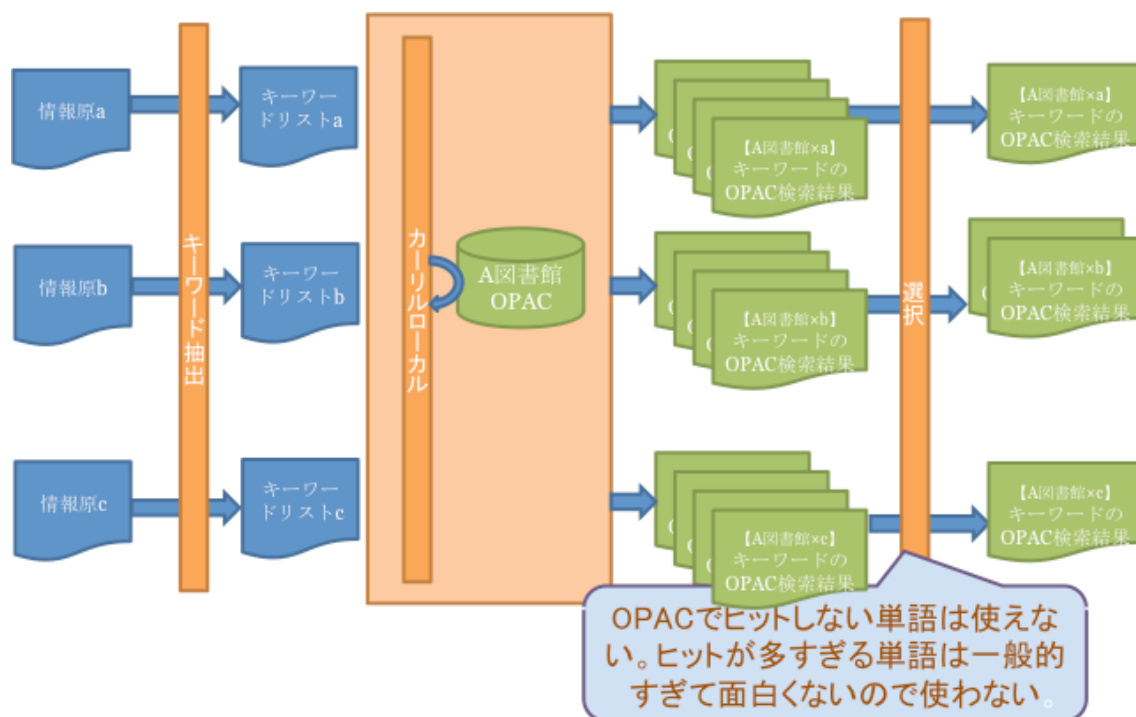


図 5 ブクリス 合宿終了時のシステム構成図

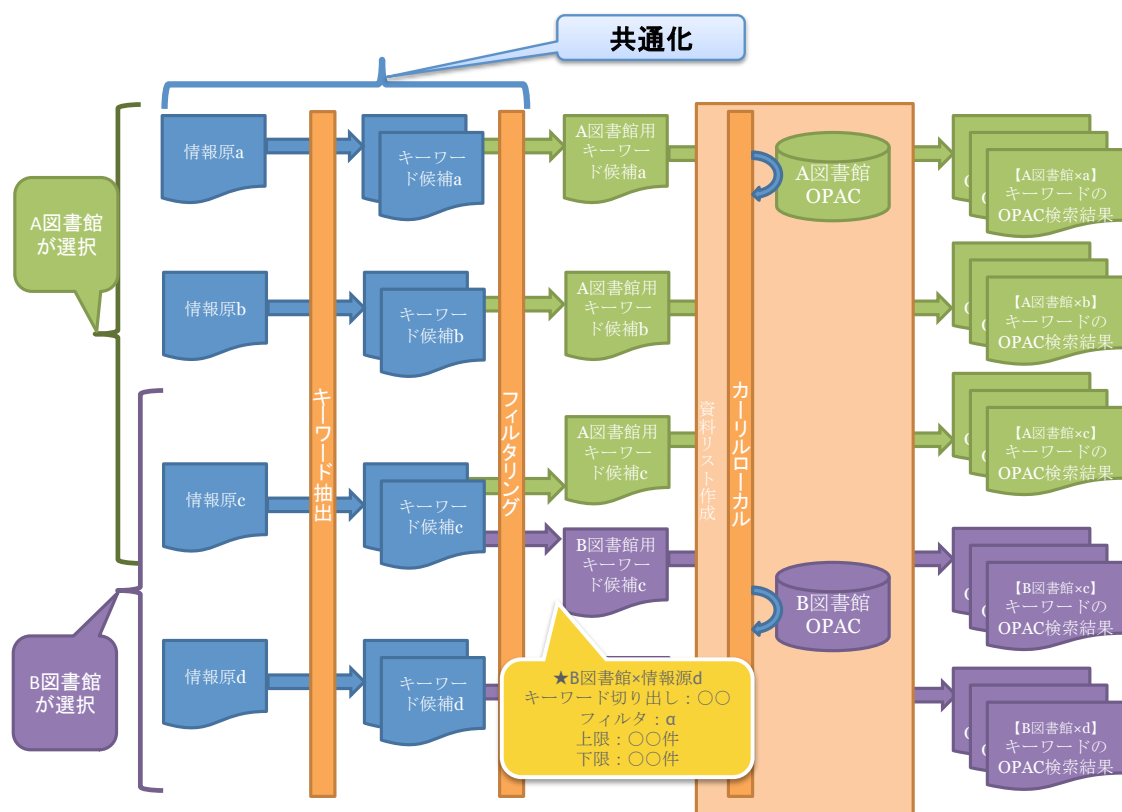


図 6 ブクリス キーワード抽出手法見直し後のシステム構成図

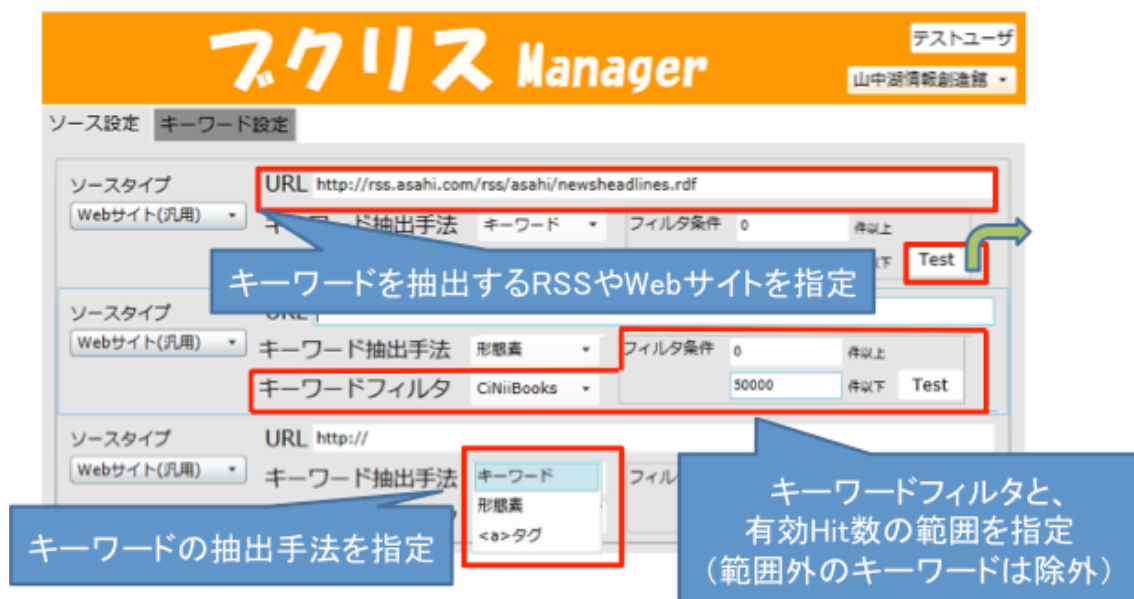


図 7 ブクリス Manager のトップ画面

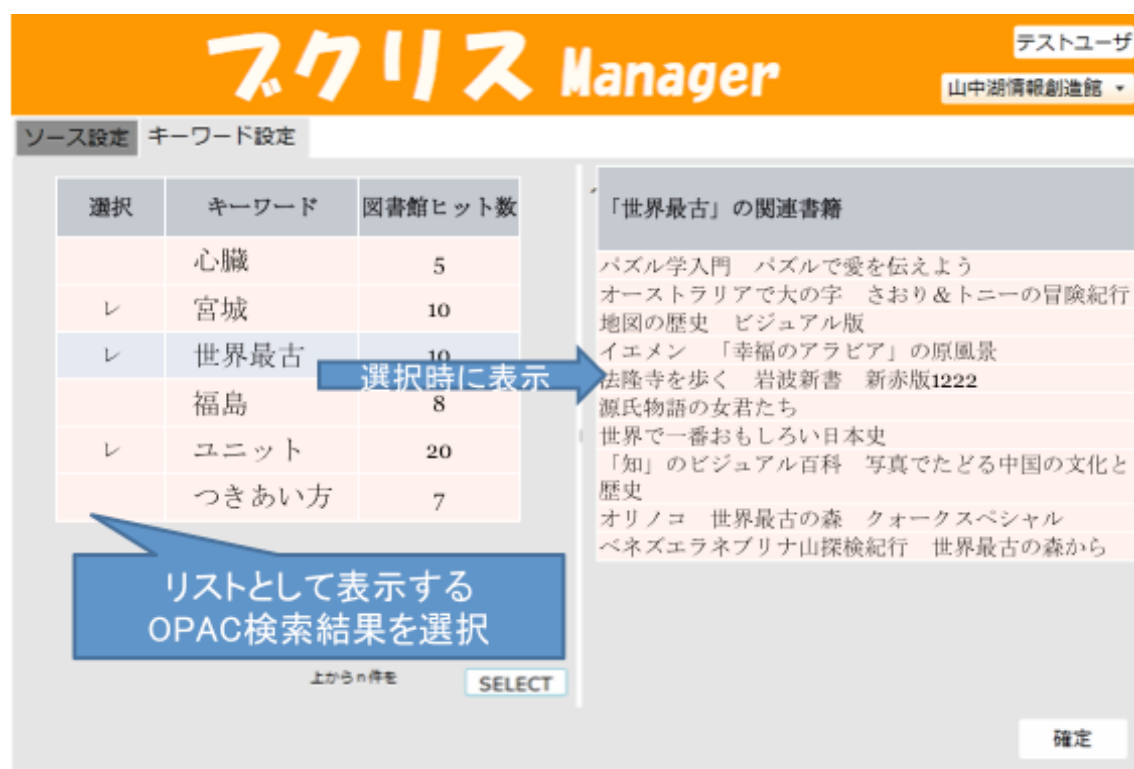


図 8 ブクリス Manager におけるキーワード設定画面

2.2.4. 開発メンバー

ブクリスの開発チームには、関戸のほか、嶋田綾子氏（千代田区立日比谷図書館文化館）、南雲知也氏（ブレインテック）、土井亮氏（農林水産研究情報総合センター）がメンバーとして参加した。また、リッテル研究所のエンジニア 1 名（太田飛鳥氏）が技術支援スタッフとして参加した。

2.2.5. 今後の展望

今後は実際に図書館に導入して使っていただくために環境等を整えていく段階となる。最初は少数の図書館に協力していただき、チームメンバーと密に連絡を取りながら導入を進め、その過程で得たフィードバックを反映し、また、ドキュメントや運用体制など、広く利用いただける形に整えていく。課題となるのは継続可能な運用体制とビジネスモデルである。サーバ環境や運用サポートを安定的に提供するためには現在のボランティア体制では不安がある。利用状況をみながら、外部からの協力を得られないかという点を含めて引き続き検討していく予定である。

2.3. 俺 CiNii（日高）

俺 CiNii は、「自分好みの論文探しをスパッと支援」することをコンセプトとし、CiNii の新着論文リストから自動的に自分のニーズに合った論文だけを抽出するフィルタを簡単に作成することができるシステムである。機械学習アルゴリズムの一種である Support Vector Machines (SVM) ²⁾を採用し、満足できるフィルタリング精度を実現している。

2.3.1. 開発の動機および議論

俺 CiNii 開発のそもそもの発端は、「自分が興味のある論文（だけ）を漏らさずチェックしたい」というチームメンバーの一人によるニーズであった。例えば、国文学系の最新論文だけを見ることができるサイトを作ることができれば、メンバー自身にとって有用なものになることが期待された。この考え方を発展させ、国文学に限らずあらゆる分野についての論文抽出フィルタのセットが作れて、自由に切り替え可能な仕組みが実現できれば、ニーズはあるのではないかと考えた。

あらゆる分野で汎用的に使える仕組みとしては、機械学習アルゴリズムの適用が必要となるが、今回は汎化能力が高く、多くのタスクで高い精度が実現で

きることから、SVM を採用することとなった。しかし、SVM を採用するにあたって以下のどちらに開発の重点を置くかで議論となった。

- A) 論文抽出タスクへの SVM の有用性の検証（精度評価）、および精度の向上を中心にすべきである。
- B) フィルタの作成と利用を簡単にできるようなアプリケーションとしての作り込みを中心にすべきである。

議論の結果、アプリケーションのコンセプトを明確に伝えることを重視して B) を重点におきつつも、精度に関する簡単な実証実験を行うこととした。

2.3.2. SVM の適用および実証実験

SVM は教師あり学習アルゴリズムであり、人手で作った教師用データ集合が必要となる。また、教師用データは「正例」と「負例」の 2 クラスのどちらかに分類されていること、教師用データおよび判定対象データは多次元ベクトル化されている必要がある。今回は、「自分にとって興味のある論文」を「正例」クラス、そうでない論文を「負例」クラスに割り当てた。また、多次元ベクトル化するための素性（機械学習の手がかり）として、論文の「著者名」そのもの、および「タイトル」を MeCab¹によって形態素解析して抽出される自立語を利用することとした。

SVM は非常に有名な機械学習アルゴリズムであり、多くの実装がすでに存在する。今回は、Cornell 大学の Thorsten Joachims 氏による SVM^{light} ²を利用した。限られた時間内で実装および実証実験を完了させるため、開発言語としては MeCab および SVM^{light} のバインディングモジュールが存在する Perl を採用した。また、あえて高度な汎用性は持たせず、CiNii の文献のみを対象とし、学習および判定に機能を絞った Perl モジュールを作成した。

こうして実装された Perl モジュールの実証実験を行うため、まずメンバーの一人が CiNii から国文学分野に属する論文 1000 件を選択した。この 1000 件の論文を学習用の「正例」データとし、CiNii の新着論文リストからランダムに選択した論文 1000 件を「負例」データとして³学習を行い、学習モデルを構築した。判定対象データとしては、CiNii の新着論文リストからランダムに選択した

¹ オープンソースの日本語形態素解析エンジン。 <http://mecab.sourceforge.net/>

² <http://svmlight.joachims.org/>

³ 「負例」データの中に国文学分野に属する論文が混ざっている可能性はあるが、その割合は少ないと見込まれることから、実証実験ではこの問題は無視した。

別の 1000 件を利用した。判定対象データに学習モデルを適用して得られた判定結果について、それぞれの論文が国文学分野に属するか否かを再び人手で判定することとした。

SVM による判定結果は、符号付きの実数値のスコアとして得ることができる。ゼロよりも大きければ大きいほど正例である可能性が高く、ゼロよりも小さければ小さいほど負例である可能性が大きい。そこで、スコアの大きい順にデータをソートし、上位 N 件以内に国文学分野の論文が何件あるかを評価することとした。その結果を表 1 に示す。上位 100 件以内に、国文学分野の 8 割以上が含まれるという結果が得られた。判定対象となった 1000 件中に含まれている国文学分野の論文 31 件（含有率 3.1%）のうち、25 件が上位 100 件に含まれていた。また、上位 100 件における国文学分野の論文の含有率は 25.0%であった。

この結果は、「自分の興味のある文献だけを漏らさずチェックする」という目的に SVM が利用できる可能性を示唆している。また、SVM に与える多次元ベクトルの作り方を工夫することで、さらに精度が向上することが期待できる。

表 1 国文学分野の論文抽出に SVM を適用する評価実験の結果

適合率の順位	「正解」精度	再現率
10 位以内	50.0%	16.1%
20 位以内	35.0%	22.6%
30 位以内	26.7%	25.8%
40 位以内	32.5%	41.9%
50 位以内	34.0%	54.8%
60 位以内	31.7%	61.3%
70 位以内	30.0%	67.7%
80 位以内	27.5%	71.0%
90 位以内	24.4%	71.0%
100 位以内	25.0%	80.6%

2.3.3. デモとシステム構成

俺 CiNii のシステムは、<http://sd.pot.co.jp/orecinii/> にアクセスして利用することができる。アカウントを作成してログインすると、図 9 に示すように CiNii の新着論文が表示される。ここで、「正例」「負例」にあたる情報を、「気に入った!」「いまひとつ」というボタンのいずれかをクリックすることでシステムに入力することができる。その後、「訓練開始」ボタンをクリックすることで、SVM による学習が行われる。なお、正例、負例の論文の NAID (NII 論文 ID) のリストを TSV で作成し、一括アップロードすることもできる。判定結果は、同じく図 9 の画面において、右側の青い棒グラフによって表示されている。グラフが長いほど、正例である可能性が高いことを示している。

俺 CiNii のシステム構成図を図 10 に示す。SVM の学習・判定にかかる機能は、「Kajiki」という Perl パッケージとしてコンポーネント化されていて、俺 CiNii のサイト上でオープンソースとして公開している。

日高崇 | ログアウト

俺CiNii

自分好みの論文探しをスパッと支援

powered by KAJIKI

新着論文 一括登録・判定 使い方

熟練奏者のバイオリン演奏時における左手指応力特性について (第 31 回バイオメカニズム学術講演会 SOBIM2010 予稿集)

小幡 哲史, 平野 剛, 木下 博
バイオメカニズム学術講演会予稿集 31 [2010-11-06]

熟練奏者のバイオリン演奏時における左手指応力特性について (第 31 回バイオメカニズム学術講演会 SOBIM2010 予稿集)

小幡 哲史, 平野 剛, 木下 博
バイオメカニズム学術講演会予稿集 31 [2010-11-06]

熟練奏者のバイオリン演奏時における左手指応力特性について (第 31 回バイオメカニズム学術講演会 SOBIM2010 予稿集)

小幡 哲史, 平野 剛, 木下 博
バイオメカニズム学術講演会予稿集 31 [2010-11-06]

熟練奏者のバイオリン演奏時における左手指応力特性について (第 31 回バイオメカニズム学術講演会 SOBIM2010 予稿集)

小幡 哲史, 平野 剛, 木下 博
バイオメカニズム学術講演会予稿集 31 [2010-11-06]

図 9 俺 CiNii のスクリーンショット

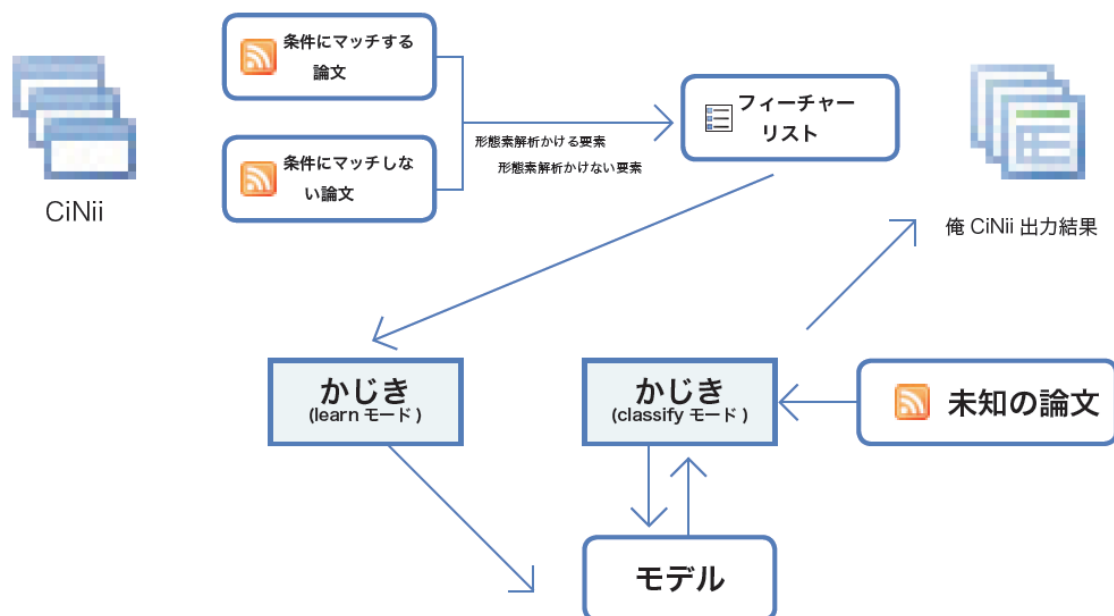


図 10 俺 CiNii のシステム構成図

2.3.4. 開発メンバー

俺 CiNii の開発チームには、日高のほか、青木隆平氏（連想出版）、岡田圭介氏（笠間書院）、山田俊幸氏（明治大学図書館）がメンバーとして参加した。SVM の適用手法に関するアドバイスおよび実装作業は清田が技術支援スタッフとして担当した。

2.3.5. 今後の展望

今回の試みは、SVM という、学術研究の場においてきわめてありふれた手法であっても、「論文探し支援」「フィルタリング自体をプラグブルなものに」といった、現実の課題に即した実装を行うことで、十分実用的なサービスとして提案できる、ということを示唆している。IT 技術と出版業の現場の間に身を置くチーム代表としては、今後もこのような両者を取りもつ翻訳者として、新たなビジョンを示すことを念頭に、現場の課題解決に取り組んでいければ幸いである。

3. 合宿の成果総括（清田、岡本）

今回のマイニング探検会合宿において共生プログラミングを実践した主な目的は、ライブラリアン自身が、自分とは立場の違うエンジニアに対して「解決し

たい課題」や「作りたいもの」を明確に伝えるための能力を磨くことにある。図書館の現場では、蔵書管理システムや OPAC、Web サイトなどの多様な情報システムが運用されているが、システムの導入は多くの場合ベンダーに発注することになる³⁾。ベンダーに発注する場合にはシステム仕様書を定める必要があるが、解決したい課題を明確に言語化できなければ、満足できるシステムを実現することは難しい。個人としてエンジニアと向き合い、コミュニケーションを行うことは、システム仕様書を書く上でも格好のトレーニングになるのではないかと考えている。

「エンジニアに明確に伝える」という観点に照らすと、合宿という特殊な環境は、参加者にとって貴重な経験を得る機会になっていたという印象を受けた。また、エンジニア側にとっても、日常の開発現場という「枠」を超えて、エンジニアにとっての専門用語を使わずに説明したり、プレゼンテーションを意識したコミュニケーションをしたりという形で、大いに得るところがあるように感じた。とくに、今回参加したリッテル研究所のエンジニアは新入社員が大半であったが、得た経験が合宿終了後に業務に生かされていることは、勉強会への協賛を行っている立場からは貴重な収穫であった。

次に、各チームの成果についての簡潔な評価を述べる。

Ref.Master は、レファレンス協同データベース API の全く新しい利用法を示している点が積極的に評価できる。また、このシステムは GWAP⁴⁾ (Games with a purpose, 人がゲームを遊んだ副産物を有益な処理につなげる手法) 研究への発展が期待できる。現在の課題は、Ref.Master が自動的に生成する選択肢の品質である。現時点では開発期間の制約上、アドホックな手法で生成しているが、今後は GWAP や自然言語処理技術の要素を取り入れて選択肢の品質をブラッシュアップしていくことが期待される。

ブクリスのチームは、合宿開始時は何を作るかについて模索していたものの、スケッチブックや付箋などを活用して具体化するプロセスを経て明確なコンセプトを提示できた点と、開発中に発生した技術的課題を参加者とエンジニアの間の積極的なコミュニケーションによって解決した点が評価できる。共生プログラミングの実践としてもよい見本であると感じた。課題は、サービスの継続性、すなわちビジネスモデルの確立にある。そのためには、新たなユーザの開拓に向けての努力が必要とされるだろう。

俺 CiNii は、あえて完璧なものを求めず、80%の完成度でも動くものを時間

内に作りきるといふ進め方が功を奏したと感じている。また、フィルタ自体を誰もが作れて共有できるというコンセプトはこれまでになかったものであり、新規性も評価できる。ただ、Web アプリケーションに仕上げていく上で特定のメンバー一人（日高氏）に開発の負担が集中してしまった点が、共生プログラミングの実践という面では課題として挙げられる。今後は、過度に誰かだけに負担が集中しないよう、興味をもつメンバーを巻き込んでいくコミュニケーションをとることが重要かもしれない。

4. アカデミックとリアルの谷を埋めるとはどういうことか（清田）

マイニング探検会の取り組みは、互いに異なるバックグラウンドをもつ参加者同士が、組織や立場の違いを超えて共通の課題に取り組むことによって、「伝える」能力を磨くことを主眼に置いてきている。参加者は、日常においてそれぞれの専門性を持って業務にあたっている方がほとんどであるが、お互いの専門性を生かしながら現実存在する課題に取り組むためには、埋めるべき「谷」がいくつもあると感じている。私自身も、もともとは大学教員として情報検索や自然言語処理技術に関する研究を行っていて、そのアイデアを生かして実用に供するシステムを作っていく過程で、さまざまな局面で「谷」を埋める必要を感じてきた。この節では、その「谷」の本質と、その「谷」を埋めるための方法論を、私自身の経験を踏まえて整理することを試みる。

私自身も含めて多くの職業人は、それぞれの持つ専門性を生かした仕事に日々取り組んでいるが、専門性の根源は、理論によって体系化された知識と方法の集合体、すなわちアカデミックの世界に存在する学問体系にある。例えば、図書館に関わるスタッフに必要とされる知識や技術は、図書館情報学という学問領域によって体系化され、正統性の裏付けがなされている。専門性を生かした仕事をしている限りは、私たちはアカデミックの世界と無縁でいることは難しい。

一方で、専門性を生かした仕事には、現実世界との接点も必ず存在する。図書館に関わるスタッフであれば、来館してくるさまざまな人々のニーズに応えていく必要があるし、場合によっては日々発生するトラブルへの対処も求められる。大学に在籍する研究者であっても、研究を遂行するためには、予算の獲得やスタッフ・学生の募集など、さまざまな現実の課題をクリアしていくことが求められる。もちろん、対処しなければならない現実の課題の大半は、組織

内で共有されている知識やシステムによって対処可能であり、それが専門性をもつ職業人の多くが組織に帰属する理由となっている。

しかし、職業人が現実世界に影響を与える方向に専門性を生かしていくことを目指したときに、多くの人は壁を感じるのではないだろうか？ マイニング探検会のメンバーの方からもそのような声をよく聞くのは、マイニング探検会で目指している「伝える」という能力の向上は、メンバーひとりひとりの影響力の向上と強く結びついているからではないかと考えている。私自身も、自分が関わっている研究の知見を社会に還元していきたいという想いがあって、大学教員でありながらベンチャーの起業に関わるという経験をしているが、その過程ではさまざまな壁に直面している。その過程で埋めるべき谷は、「学問領域」と「実用領域」、「組織」と「個人」、「ルール」と「ガイドライン」という 3 つの対立軸の上にそれぞれ存在していると私は感じている。以下に、その 3 つの対立軸について整理する。

- 「学問領域」と「実用領域」

学問体系を構成していく上では、その学問領域に関わる人々の間で通用する「評価軸」を作るプロセスが欠かせない。ここで、「評価軸」を設定するために、学問体系の一貫性を崩す要素はあえて排除されることがある。例えば、情報検索システムの研究分野では、「入力」と「出力」をもつブラックボックスとして情報検索システムをとらえ、「入力」と「出力」の間の関係を定義することで、客観的な性能測定ができるようにし、工学的なアプローチによる研究を可能としている（図 11）。しかし、現実の情報検索システムには必ず「人間」が利用者として介在しており、システムの「出力」が「人間」の次の「入力」に影響を与えるというモデルを考えなくてはならない（図 12）。また、現実の情報検索システムの入力は、必ず「人間」のリアルなニーズに根差している。私自身は、システムと人のインタラクションという側面で情報検索システムに興味を持っており、アカデミックの世界で共有されている情報検索システムの評価軸だけでは本質を明らかにすることが難しいと感じていたことが、「実用」に供するシステムを新たに作ることで研究するというアプローチを選択する大きな動機となった。一方で、「実用」という領域に踏み込むためには、周囲を納得させることができる新たな「評価軸」を創り出すという努力が求められる。すでに共有されている「評価軸」に根差して活動するアプローチと比較すると、遠回

りをしなければならない場合も多い。ここには、現実世界との接点を増やすためのコストと、現実世界への影響力というベネフィットの間のトレードオフが存在するのではないだろうか？



図 11 工学的アプローチによる情報検索システム研究のモデル

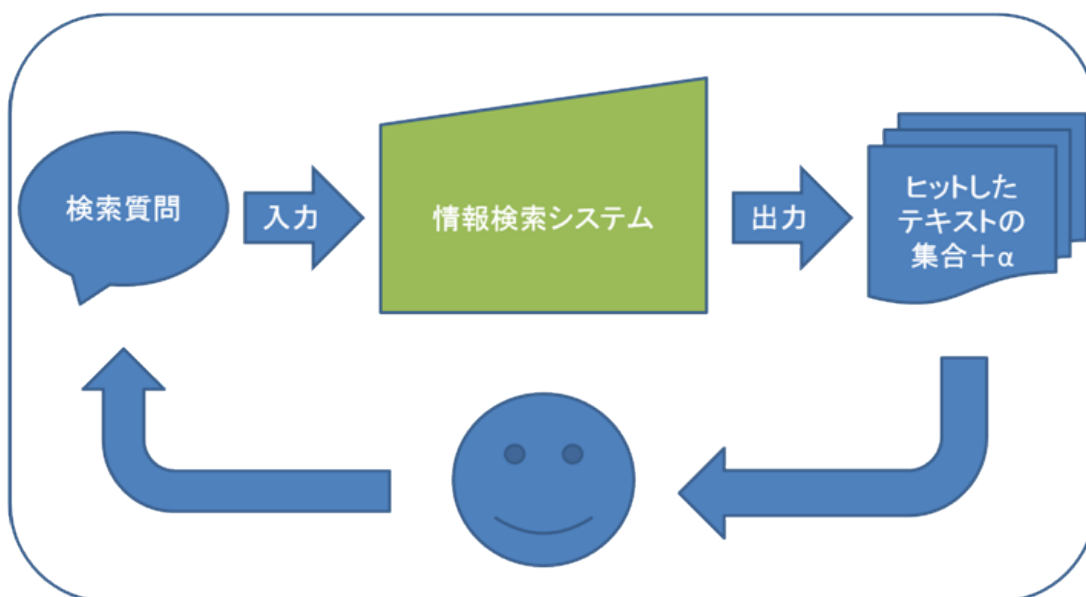


図 12 現実世界の情報検索システムのモデル

- 「組織」と「個人」

上述したように、専門性を持つ職業人にとっては組織に属することの利点は極めて大きい。組織に属することで、現実世界で起きる課題の多くは解決できるため、専門性の発揮にリソースを集中することが可能になる。また、専門性の発揮の積み重ねによって社会的信用が蓄積されたり、収入が

安定したりするというメリットは無視できない。

しかし、組織に属する人間が外部の世界に働きかけを行うときには、組織内でのコンセンサスを得るという過程が必要となる。多くの場合において、コンセンサスを得るためにはある程度の時間が必要となり、スピードという要素が失われてしまう。組織から離れて個人として行動すると、組織のメリットは享受できないものの、自分が決めればすぐに動くことができるため、変化の速い状況ではスピードという要素が何よりの強みとなる。

もう一つ、組織に属することの見えにくい危険についても指摘しておきたい。専門性を持つ職業人には、専門性を選択した段階で何らかの動機が存在するはずである。例えば、図書館員になる、研究者になるという選択の裏には、それぞれの人生の目的が必ず隠れている。組織のロジックにどっぷり浸かってしまうと、人生の目的を見失ってしまう傾向が私たちにはあるのではないだろうか？

- 「ルール」と「ガイドライン」

専門性のバックグラウンドとして存在する学問領域は、専門性の正統性を裏付けるとともに、専門性の規範としての「ルール」も導出する。例えば、図書館員には「利用者の秘密を守る」、研究者には「先人の業績を踏まえる」という「ルール」が存在する。また、すべての組織は共通の「ルール」なくしては成り立たない。「ルール」は私たちの行動を束縛する存在であるとともに、私たちを万が一のリスクから守ってくれる存在でもあり、決断のためのコストを下げる、すなわち「楽をさせてくれる」存在でもある。

一方で、学問領域が規定している枠の外（たとえば実用システムを作るとか）で活動したり、組織から離れて個人として活動したりする際には、必ずしもすべての事柄は「ルール」では規定されていない。そこに存在するのは、「何のために活動するのか」という目的であり、その目的に達するためにアドホックに共有されている「大まかな決まり事」でしかない場合がほとんどである。「大まかな決まり事」は、現実の課題に対処していく過程で頻繁に改訂されていく。このような「大まかな決まり事」のことは、「ガイドライン」と一般に呼ばれている。「ガイドライン」は、現実の課題に合わせて行動することのある程度許容するという緩さをもつ。しかし、「ルール」が機械的な行動の適否判断を可能にするのに対して、「ガイドライン」への行動の適否判断には、「大まかな決まり事」であるがゆえのコストがか

かる。また、行動の結果起こることについては、行動した人間自身が説明責任をもつことになる。

Wikipedia というシステムは、「ガイドライン」の考え方を理解するための格好の教材であると私は考えている。Wikipedia プロジェクトの最終的な目的は「信頼されるフリーの百科事典を作る」ことであり、Wikipedia コミュニティではその目的が最重要視されている。その目的を達成するため、「出典を明記する」「著作権を尊重する」といった「ガイドライン」の数々が共有されている。一方で、Wikipedia は現実世界で多数の編集者によって盛んに編集されている媒体でもあり、すべての記事が「ガイドライン」に沿っている状態の実現は、理想ではあるものの極めて難しい。それでも、目的を達成するための秩序維持の努力は、多数のボランティアによって継続してなされている。「ガイドライン」という考え方は、新たな価値を創造するためには不可欠である一方、「ガイドライン」を維持するためにはビジネスモデルの構築も含めた不断の努力が求められる。

5. おわりに

本稿では、マイニング探検会という新たな試みについて、合宿の成果を紹介することによって未来の図書館の可能性を示すとともに、マイニング探検会という試みの意図である「アカデミックとリアルの谷を埋める」ために解決すべき課題を示した。上に述べたように、「谷を埋める」というプロセスにはさまざまな「リスク」が存在する。より大きな努力を求められることや、「組織」や「ルール」という「私たちを守ってくれる存在」から離れて動くことで起きる事柄は、「リスク」ととらえてよいだろう。一方で、「谷を埋める」というプロセスには、「現実世界への影響力を大きくする」「本当の目的は何かを問い直す」という側面があることも提示した。

人間には「リスク」を避けて「安定」を求める傾向がある。「リスク」を取るか否かは個人の自由意志に委ねられているし、やみくもに「リスク」を取ることが正解ではないと考えている。しかし、現在の情報技術の急速な発達で、図書館という社会的機能にも大きな変化をもたらしている環境を踏まえると、Facebook の創業者 Mark Zuckerberg の「最大のリスクは全くリスクを取らないことである」という言葉には重みがあるのではないだろうか。図書館の未来は、図書館に関わる私たち一人ひとりの今この瞬間の行動によって規定される。

一人でも多くの方々が、図書館の未来を作るために今この瞬間から枠を超えて活動されることを期待している。

マイニング探検会の活動に興味を持たれた方は、清田または岡本まで個人的に連絡をいただきたい。

参考文献

- 1) 財団法人ハイパーネットワーク研究所. ハイパーネットワーク 2011 ワークショップ. <http://www.hyper.or.jp/staticpages/index.php/ws2011detail> (accessed: Jan. 31st, 2012).
- 2) Cortes, Crinna; Vapnik, Vladimir. *Support-vector networks*. Machine Learning. 1995, Vol. 20, Issue 3, p. 273-297.
- 3) 黒澤公人. 日本の図書館システムの現状. <特集>多様化する図書館システム. 情報の科学と技術. 2011, Vol. 61, No. 5, p. 194-199.
- 4) von Ahn, Luis. Games with a Purpose. Computer. 2006, Vol. 39, Issue 6, p. 92-94.