

ごぞんじですか？ カーリル

Nota Inc.
洛西一周 Isshu Rakusai

1. はじめに

GoogleやAmazonなどの大規模な検索エンジンが一般に普及する中で、図書館システムも大きな曲がり角を迎えている。GoogleやAmazonは検索対象となる情報量が膨大であるだけでなく、高速であることや、検索結果の並び方が秀逸であること、利用方法がシンプルであることなど、それまでの検索エンジンの常識を覆し、広く一般に普及し、検索という行為を「日常的な行為」にまで落とし込むことに成功した。これに対して、日本の公共図書館システムは、規模、速度、精度、操作性において劣っていると言わざるを得ない。公共的な資金を毎年何十億という規模でつぎ込んでいる日本の公共図書館システムはなぜ対抗できないのだろうか。

本稿では、我々が日本の公共図書館システムのレベルアップも視野に入れて2010年3月に公開した日本最大級の図書館蔵書検索サイト「カーリル」

(<http://calil.jp/>) について紹介する。カーリルは、現時点で唯一、日本の4300以上の図書館・図書室に対応した横断的検索システムである(図1)。利用者は、最初に地域を選ぶと、近くの図書館や相互貸借できる図書館など複数の図書館からわかりやすいデザインで蔵書を探することができる。探した蔵書は、「読みたいリスト」に登録し、印刷して持ち運ぶことができる。またAmazonと連携しており、蔵書がなければ購入も可能である。その他、既存の書籍や図書館に関するウェブサービスと積極的に連携し、利用者にとって総合的に書籍を楽しむことができるような工夫を盛り込んでいる。本稿では、カーリルのコンセプト、開発手法およびシステムの詳細について述べる。



図1 図書館蔵書検索サイト「カーリル」

2. 開発の背景

筆者が代表を務めるNota Inc. は、インターネット上でクリエイティブな体験を提供することをミッションに掲げるインターネットサービスの開発会社である。文字や画像、映像を簡単にレイアウトしてコミュニティで共有できるNOTA¹や、写真をその場でスライドショーにして公開できるPhotoPeach²などのサービスを開発し運営している。いずれもインターネット上で楽しく、簡単に使えることが特徴であり、年代を問わず幅広く利用されている。

メンバーの吉本は、もともと地元の行政システムの開発に関わっており、既存の図書館システムが利用者視点で作られておらず、普及していないことを改善したいという思いを持っていた。そのような中で、製品をブレインストーミングによって話し合う中で「楽しい図書館検索」というアイデアが生まれた。図書館検索は、本来、本を発見し出会う楽しみがあるはずであり、そのような図書館検索システムが実現できれば、利用者は積極的に図書館検索を利用し、結果として図書館に足を運ぶ機会も増えるのではないかと考えた。このような仮説を元にシステムの設計を開始した。

3. 楽しい図書館検索システムの提案

我々は図書館検索システムを開発するにあたり「本を発見する楽しみを提供する図書館検索システム」、「図書館の利用促進につながるシステム」を目標にした。この目標のため、具体的に以下に述べる5つのコンセプトを定めた。

3.1 本と出会う楽しみを提供

図書館は新しい本と出会う場所であると考え。本とは知識であり、図書館は新しい知識と出会う場である。最初に図書館を訪れるのは、子供の頃であることが多いと思うが、我々はこの図書館の原体験を重視した。子供の頃の図書館体験が本に対するポジティブな感情につながっていると考えた。すなわち、カラフルな本が本棚にずらっと並んでおり、一つずつ手にとるときの喜びをサイト上でなるべく再現することを目指した。そのためには、本の書影が並んでおり、ボタンを押すことで次々と本を眺めることができる必要がある。

3.2 検索を使って図書館を広報する

一般に図書館検索システムは、利用者がキーワードなどを入力するとそれに合致した結果を出力する。これは別の言葉で言い換えると、書誌をデータベースとして管理しているにもかかわらず、ユーザーが検索窓に入力しない限り、データベースの情報を出力していないということである。しかし、膨大な書誌データベースを利用者の入力があるときにしか表示しないのはもったいない。

サイトにアクセスしたときに、最初からさまざまな並び方で書誌を表示することができれば、図書館の本をアピールでき、利用者が足を運ぶきっかけになる。例えば、人気の本、作家別の本、時事ニュースに関連する本、地域色のある本などを並べて紹介するのである。すべての書誌がこのような形でアクセスできるような状態になっていることが理想である。これによって書誌データベースをさらに活用できる。多くの図書館が館内の本の並びを工夫しているが、それをシステム上で実現できれば、利用者にと

って魅力が増す。それは図書館サイトにとって価値をストレートに伝える広報手段となる。

3.3 地域資源の中心地としての図書館

多くの図書館システムの検索結果のページでは、蔵書がある図書館の名称が、「○○図書館」「○○図書室」（もしくは単に地名のみ）などと表示されるが、具体的にすべての利用者がこれらの図書館の位置や特色を正確に把握しているとは言い難い。蔵書がどの図書館にあるかという情報は、その図書館の詳しい情報のページとリンクされていることが望ましい。このことは単にその図書館の詳しい情報を知ることができるというだけにとどまらず、利用者にもその地域への興味を持たせるきっかけとなる。公共図書館は、蔵書を検索できるだけでなく、地域の行政、文化、経済、観光などと密接に結びついた存在である。蔵書の所蔵館から地域を知り、また地域の他の資源にもアクセスできるようになることで、地域の知の貯蔵庫としての図書館の価値も高まる。

3.4 図書館にない本も検索

既存の図書館検索システムは、自館にある本しか検索できないことが多い。図書館にない本を検索するとき、利用者は、結果がゼロ件と表示されても、他のキーワードの可能性などを考え、なんどもキーワードを入力しなおすなどの行為を行う可能性がある。これは無駄な行為であり、検索システムへの信頼も落とすことになる。また、我々の司書へのヒアリングによれば、多くの司書が利用者から自館にない本に関する問い合わせを受けたときに、Amazonを第二の検索システムとして利用していた。

これらの問題は、図書館にない本も同時に検索結果に表示するようにすれば解消できる。司書は手間を軽減でき、利用者は発見した本を図書館にリクエストしたり、本屋で購入したりといった行為もとることができる。

3.5 外部サービスと連携して利便性を高める

インターネット上のサービスは日進月歩で発展しており、利用者が本のレビュー（書評）を投稿できるサイトや、本棚を作成して共有できるサイトなど

が多く存在する³。利用者は、これらのサイトを通して本を発見し、本の理解を深め、本を通じて友人とのコミュニケーションを楽しむ。図書館の検索システムは、これらと連携し、「本」に対する総合的な利便性を利用者に提供することが求められている。

以上が、カーリルのコンセプトである。この中で述べられている機能に関しては、カーリルのシステムでほぼ実現されている。システムの詳細に関しては第5章で述べる。

4. 開発手法

カーリルの開発と、既存のベンダーによる公共図書館システムの開発とは、その手法において異なる点がある。カーリルは、構想から公開までかかった期間が2ヶ月と短かったことが特徴である。本章では、カーリルの開発手法について述べる。

4.1 開発チームの構成

カーリルを開発するNota Inc. は、本社は米国San Joseだが、メンバーが住んでいる地域は京都、中津川、藤沢などに分散しており、それぞれにオフィス（または作業場）がある。普段はSkypeなどを使ってコミュニケーションをとっている。担当業務は「ディレクション」「検索エンジン・API開発」「インターフェイス開発」「アート・デザイン制作」「図書館データベース作成」「ユーザーサポート」となっている。カーリルに関わっているメンバーは全部で6人である。少人数で各地に分散しながら開発しているのが特徴である。

4.2 開発期間とモチベーション

カーリルの開発は、構想に1ヶ月、開発に1ヶ月、合計2ヶ月で行われた。開発者（エンジニア）の人数は4人である。公開後にいくつかのバグが見つかったが、数日で修正され、現在は安定して動作している。開発期間が短時間であるのは以下のような理由がある。

一つ目は開発手法である。まずプロトタイプを作り、利用者目線で使用して納得できない部分を修正するサイクルを短期間で繰り返すという手法をとる。最初に決めた仕様を最後まで固持するというような

ことはしない。仕様を最初に固めると、開発中に利用イメージと異なる部分が出てきたときに修正作業に予想以上の時間がかかるためである。

二つ目は開発者のモチベーションである。カーリルが短期間で開発できたのは、図書館を便利にし、図書館利用者を増やすという仕事に対して開発者が闘志を燃やしたことが大きい。ソフトウェア開発は、頭脳労働であり、一定の時間をかければ一定の結果が生み出されるという仕事ではない。開発期間は、開発者のやる気によって大きく変わる。また細部の仕上がりに関しては、仕様を満たす最低限のレベルのものを作るか、求められる以上のものを作るかどうかは、やる気や、労働環境などが影響する。筆者は、仕様を定めて、「人月」という単位で仕事を割り振るソフトウェア開発の手法は、開発者のやる気を引き出すという点において非効率な方法であると考えている。

4.3 ユーザーフィードバックに基づく修正

サイトは公開後も継続的に修正を行っている。カーリルはTwitter（アカウントは caliljp）においてユーザーからの要望を受け付け、投稿された要望のうち、必要と判断したものを順次開発し公開している。これまで、公開後2ヶ月半の間に30以上の要望を採用し、開発した。システムの更新を頻繁に行うことで、利用者のニーズにきめ細かく対応していくことが可能である。

5. カーリルのシステム

本章では、カーリルのシステムの詳細について述べる。情報技術の専門的な内容も含むが、技術開発に興味のある方は読んで欲しい。

5.1 マッシュアップの活用

カーリルは、他サービスが提供する機能やコンテンツを組み合わせて活用している。これをマッシュアップと呼ぶ。まず一つ目はAmazonの書誌情報の利用である。Amazonの書誌情報は、Amazon AWS⁴に登録することで利用できる。カーリルでは、キーワードを用いて書誌を検索するItemSearchと個別の本の詳細情報を取得するItemLookupを利用している。また

書誌のサムネイル画像もAmazonが提供するものを利用している。

二つ目はGoogle Mapsである。各図書館の場所を表示するために利用している。蔵書情報のページから、各図書館のページにリンクしているため、リンクを辿れば、蔵書がどこにあるか地図上でわかる。

三つ目はWikipediaである。キーワードによる検索が行われたときに、そのキーワードに関する記事をWikipediaから取得し、記事の概要を本の検索結果の上に表示している。これにより、利用者はキーワードへの理解を深めることができる。また、記事の概要に別のWikipediaのキーワードへのリンクがあった場合、クリックすることによってそのキーワードの検索を行うことができる。この機能によって利用者は、新しいキーワードに出会い、それらのキーワードに関連する本を発見することができる。このようにカーリルでは、利用者に対して、関連するキーワードを推薦している。

四つ目はYahoo Image Searchである。このAPIは、キーワードをもとに画像を検索できる。この機能を利用して、各図書館のページに図書館の画像を表示させている。

五つ目は、東京図書館制覇!⁵である。東京図書館制覇!は東京の図書館のうち246館を回り、その訪問記や、開館時間、蔵書の種類、設備などを詳しくまとめたサイトであり、図書館の情報ポータルサイトとしての価値が高い。竹内庸子さんが、個人で運営している。カーリルではこのサイトと連携し、カーリルの各図書館の詳細ページに東京図書館制覇!の訪問記の一部を掲載し、リンクを張ることで利用者の利便性を高めている。

5.2 開発言語

カーリルのシステムはすべてプログラミング言語のPython⁶によって実装されている。Python2.5および2.6を利用しており、Pythonの標準ライブラリとオープンソースライブラリのみで構成されている。サーバーへのセットアップはなるべく簡単に行えるように設計されている。標準的なLinuxサーバーに

プログラムファイルを入れて立ち上げるだけでタスク処理が開始される。

5.3 分散処理

カーリルでは、低コストで大量のアクセスを処理するため、分散処理を工夫している。現在は、チューニングも完了し一日12万アクセス、ピークトラフィック毎秒30アクセスを安定的に処理している。各図書館システムへのアクセスに関しては、共有メモリ(memcache)を利用したキューによって管理している。キューは、各図書館システムへのアクセスが高頻度になりすぎないようにするために調整する役割を果たしている。具体的には、各図書館システムへのアクセスは最大1~10セッションまでに設定されており、その頻度を超えないようにキューがアクセスを制御する。利用者が一時的にこれ以上の頻度でカーリルの検索を行った場合、限度を超えたキューが実行待ちの状態となり、順次、前の処理が完了するまで利用者は待つことになる。

また、カーリルのインフラは、クラウドサーバー環境であるGoogle App Engine⁷と通常のサーバーのハイブリッドクラウドによって構成されている。利用者からのアクセスは、まずGoogle App Engineサーバーが受信し、キューの処理は、Google App Engineと通常のサーバーに分散して処理する。すなわちGoogle App Engineを司令塔として利用している。Google App Engineは、カーリルへのアクセスが一時的に増大したときにも、多少の速度低下は見られたものの、問題なく処理することができた。しかし、時間のかかる重い処理はコストがかかるため、Google App Engineに振り分けず通常のサーバーに振り分けるようにしている。このようにクラウドサーバーと通常サーバーのメリット・デメリットを考慮して、コストやアクセス量に応じて処理を振り分けている。

5.4 図書館クラス

カーリルは各図書館の蔵書情報をスクレイピングによって取得している。スクレイピングとはウェブサイト上のHTMLを取得し、解析する技術のことである。

スクレイピングを利用することで、APIなどの共通インターフェイスを用意していない図書館システムであっても蔵書情報を取得することが可能になる。カーリルでは、1200以上の図書館システムに対応するため、蔵書の検索を行う処理を、オブジェクト指向のクラスの「継承」機能を利用して階層的に実装している（図2）。一番上の階層は独自に開発したLibrarySearchクラスである。このクラスがHTMLの取得や、解析などすべての図書館システムに共通する処理を行う。ISBNの処理や、蔵書の状態（「貸出中」「貸出可」「予約中」）などの管理もこのクラスで行う。この下に各ベンダーのクラスがあり、その下にベンダーのシステムのバージョンごとの対応クラスがある。これらのクラスで各ベンダーのシステムへ対応するための処理が記述されている。ベンダーのクラスは合計で約10個ある。この下に各図書館のクラスがある。図書館のクラスは合計で約1200個ある。各図書館クラスは図書館毎の個別のカスタマイズに対応する。図書館毎の違いはそれほど大きくない。実際には、システムクラスの関数やパラメーターをオーバーライドして対応を行うが、例えば岐阜県中津川市のクラスではコードの記述量は4行である。このため、既知のベンダーの図書館システムを一つ追加することは短期間で完了できる。また新しいベンダーのシステムに対応する場合もLibrarySearchという基本クラスができているために一からコーディングするのに比べて少ないコード量で対応できる。

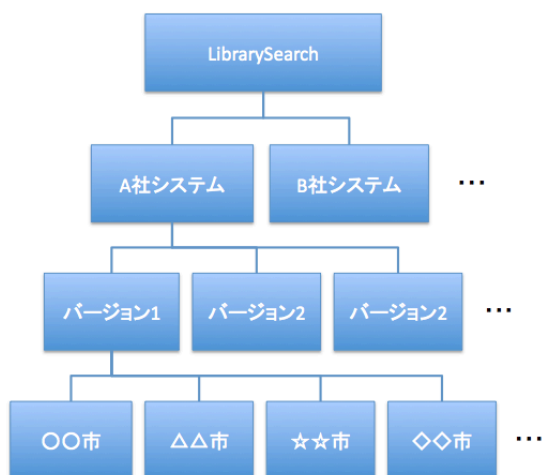


図2 図書館クラスの構成

5.5 OPACの管理

カーリルでは、図書館システムの入替、市町村の合併、図書館の新設、休館、閉館などにより図書館システムに変更が生じたときにすばやく対応できるような工夫を行っている。カーリルは2010年4月だけで約10の図書館システムの変更に対応し、約10の図書館の情報更新に対応した。全国の公共図書館システムの数は1200以上あり、常に何ヶ所かは入替や更新が行われていると考えなければならない。このため図書館システムの変更にすばやく対応することが重要になる。我々は、これに対し二つの工夫を行っている。一つ目は、図書館システムの変更追跡システムを開発し、運用している。このシステムでは、「図書館システム」を外部検索エンジンを利用して検索し、カーリルが対応していないものを自動で検出する。またすでに対応している図書館システムの対応図書館に変更が加えられた場合に検出を行う。検出されると、必要な情報が記載されたメールが開発者に自動で送信される。

二つ目は、Twitterなどを利用して、図書館の情報を追っている。カーリルのTwitterアカウントである@caliljpに対してのつぶやきや「図書館」が含まれるつぶやきを監視し、図書館システムの変更などの情報を得ている。

独自に調べたところでは、いくつかの都道府県横断検索システムは、検索対象となる図書館システムの変更が反映されておらず（反映を行っている最中の場合も含む）、虫食い状態になっているものが見受けられる。カーリルは、少ないコーディング量で図書館システムに対応でき、かつシステム変更を自動検出するシステムを導入することで、虫食いが起きないように工夫している。

5.6 図書館APIによる図書館システムの拡張

カーリルは、横断検索機能を外部の開発者も利用できるように「図書館API」として公開している。APIとはApplication Programming Interfaceの略で、共通の規格で他のシステムと連携するための仕組みのことである。これによって、外部開発者は、新たに図書館の横断検索システムを開発しなくてもカー

リルの検索機能を利用できる。図書館APIは (1) カーリルに対応した図書館情報を取得する機能、(2) 蔵書のISBNと図書館を指定することで蔵書の有無と貸出状態（「貸出中」「貸出可」「館内のみ」など）を取得する機能を提供する。蔵書の検索機能では、複数の図書館と複数の蔵書を指定できるため、横断検索のような用途にも利用できる。詳しい開発の方法は、ウェブサイトの図書館API仕様書⁸を参照してほしい。

5.7 図書館APIの利用例

現在、カーリルの図書館APIは、原稿執筆時点で法人と個人をあわせて120の利用申請がある。いくつかの利用例を紹介する。「かーりるん⁹」は、仮想的な司書の役割を持つTwitterボットであり、Twitterを利用して、かーりるんアカウント宛に「〇〇県△△市で〇〇（著書名）借りられる？」とつぶやくと蔵書の有無をつぶやきで返してくれるシステムである。図書館の蔵書をバーチャルなキャラクターとのやりとりを通して調べることができる点が斬新である。

「Libron¹⁰」はAmazonのサイト上で、図書館の蔵書の有無を調べることができるブラウザ拡張である。

「Libreq¹¹」は、Twitterを通して、図書館への蔵書のリクエストが行われたつぶやきを集計するシステムである。現時点では公式に図書館とは連携していないので、図書館へのリクエストを独自に集計している勝手サイトであるが、登録された本が図書館に所蔵されるとウェブサイト上に表示され、利用者に通知する機能がついている。「Next-L Enju 紹介状発行プラグイン」は「Next-L Enju」というオープンソースの図書館システムを拡張し、自動で近隣の図書館の蔵書を確認し、PDFファイル形式の紹介状を生成するシステムである。この他、スマートフォン（多機能携帯電話）のiPhoneやAndroid上で、GPS情報を利用して利用者の現在地点を取得し、近隣の図書館を地図上に表示したり、自動的に最寄りの図書館から蔵書を検索したりできるシステムもある¹²。

これらはいずれもカーリルの図書館APIをシステムの一部に利用している。図書館APIは現在のところ、

法人よりも個人のほうが開発者が多い。公共図書館の利用者の中で開発技術のある市民が、カーリルAPIを利用して便利なものを自主的に開発している。このような自主的な活動によって開発、公開されたシステムが、他の市民にも便利に利用されている。また一部の開発者は、アプリケーションを少額で販売したり、サイトからアフィリエイト収入や広告収入を得たりしている。

ITを利用したサービスは、広範囲にわたり、一つの開発事業者のみですべての利用者のニーズを満たすことは難しくなっている。新しい携帯端末への対応や、図書館以外の地域行政サービスとの連携、お年寄りや子共、障害者、外国人向けの特殊なシステムの開発などさまざまな用途へ柔軟に対応するために、図書館APIは有効に活用できる。

6. 課題

6.1 技術的課題

現在、カーリルのスクレイピングは、ISBNを持つ書籍しか検索できないという制約がある。このため、ISBNのない郷土史や行政史料を検索することができない。この点に関しては多くの要望が寄せられており、現在解決のための開発を行っている。

6.2 ビジネスモデルの模索

現在のカーリルの収入源は、Amazonのアフィリエイトであるが、今後、書店や出版社、図書館との提携を通じて新たなビジネスモデルを模索していく。

6.3 対応図書館の拡大

カーリルは、専門図書館からの要望を受け、対応館を増やしている。現在、東京国立近代美術館、日本点字図書館、国立科学博物館 自然教育園、国立新美術館 アートライブラリー、東京国立近代美術館フィルムセンターなど15以上の専門図書館に対応しているが、これらの図書館は、ISBNのない特殊な資料が多いため、まずそれらを検索対象とした上で、どのように検索結果を表示するのが望ましいか検討しなくてはならない。

7. おわりに

本稿では、カーリルが「利用者が楽しく本を探せること」及び「図書館の利用促進につながること」を目標に開発されたことを紹介し、その実現のための5つのコンセプトを提案した。また開発手法とシステムの詳細について説明した。

カーリルは、システムとしてすでに安定して稼働し、多くの利用者および図書館関係者から利用されている。今後は、読者の皆様のご意見やご批判を聞きながら、一層システムの改善に努めていく予定である。

参考文献

¹ NOTA <http://nota.jp/>

² PhotoPeach <http://photopeach.com/>

³ ブクログ <http://booklog.jp/> や本棚.org

<http://hondana.org> など

⁴ Amazon AWS <http://aws.amazon.com/>

⁵ 東京図書館制覇！ <http://www.tokyo-toshokan.net/>

⁶ Mark Lutz (著), 夏目 大 (翻訳) : 初めてのPython 第3版 (大型本)、オライリージャパン、2009

⁷ Google App Engine

<http://code.google.com/intl/ja/appengine/>

⁸ 図書館API仕様書 <http://calil.jp/doc/api.html>

⁹ かーりるん <http://twitter.com/calilun>

¹⁰ Libron- 図書館をもっと身近に <http://libron.net/>

¹¹ Libreq- 読みたい本を図書館にリクエスト

<http://libreq.net/>

¹² iPhone用は、「すぐカーリル」と「CALITZ」、Android用は、「図書館検索Beta」と「ナビリブ」が公開されている

<http://blog.calil.jp/2010/05/api.html>

本稿は専門図書館協議会発行機関誌「専門図書館」242号(2010年7月号)に掲載されたものです。
--