

「カーリルの図書館 API やシステム連携で広がる可能性」

吉本 龍司 氏（株式会社カーリル代表取締役）

【講師紹介】

委員長：株式会社カーリルの代表取締役であります吉本様は、昭和 57 年に岐阜県今の中津川の市生まれで、慶応義塾大学の環境情報学部卒業です。小学生の頃からプログラミングを始め、高校時代の平成 12 年 8 月に有限会社アール・ワイ・システムを設立し、代表取締役を務めるということです。平成 24 年 6 月からは日本最大の図書館蔵書検索サイト「カーリル」を運営する株式会社カーリルの代表取締役ということで現在にいたっているという御経歴をお持ちです。今日は、「カーリルの図書館 API やシステム連携で広がる可能性」ということで御講演をお願いしています。それでは、よろしく申し上げます。



講師：カーリルの吉本と申します。最近ですね、図書館で話をさせていただくことが増えているのですが、今日は大変マニアックな話で、自分は非常にわくわくして聞いておりました。僕の興味とピッチャシで、僕のための会じゃないのかなと思っております。今日は本当に勉強させていただいてありがとうございます、という状況です。

カーリルというサービスを運営しております。カーリルとは、いくつかの図書館をまとめて検索することができるサービスで、全国 6,700 館以上の図書館に対応しています。公共図書館、図書館法に基づく図書館で計算すると 93 パーセントくらいの図書館に対応しているサービスです。民間企業である株式会社カーリルという組織で運営をさせていただいています。今から約 5 年前、2010 年 3 月にサービスを開始しました。なので、もうそろそろ 5 年になります。

最初からのコンセプトとして、全ての図書館をつなぐということを考えていました。とにかく、ウェブサービスということなので、日本の図書館が全部探せなかったらサービスとして成立しないというところからスタートいたしました。本を検索して、都道府県ごとのベースラインをこうまとめて簡単に、各図

書館でどういう貸出状況なのかっていうことをまとめて検索できます。もちろんそのまま OPAC に飛ぶこともできます。

カーリルのウェブサービスは月に 50 万人ぐらいの方にユニークとして使っていただいています。実際 API というもので、いろいろなアプリでもカーリルのデータを使っていただいています。単純にカーリルのウェブサービスっていうところで見るとこのくらいの方が使っていただいているというサービスです。

そもそも 2010 年の 1 月、最初にリリースする二か月前になるんですけど、新しいウェブサービスをやりたいねと言う話をしまして、どんなサービスをこれから作っていかうかというアイデアを出した中に「図書館」というものがありました。それからいろいろ考えまして、2010 年 3 月にカーリルというサービスを開始したわけです。大体 2 か月ぐらいの期間で、始めた当時が全国の、特に公共図書館中心に 4,300 館だと思うんですけど、これでもかなりの率に対応した、僕たちの出来るところの限界に対応したというのが 4,300 館でした。とにかくすべての図書館に対応したいというふうに考えたわけです。実は、これをこれからがつつりやってみようって話ではなくて、みんなどうやってこれを使ってくれるんだろうとか、こういうのが要るとか要らないとかわからないので、まずは出してみようというところからスタートしたわけです。

そうしたところ、実は一番最初に僕たちが感じたのは、IT 関係者の人からすごく大きな反響をいただいたということでした。何が反響だったかというのはすごくはっきりしていて、「図書館ってこんなに新しい本があるんだ」って話なわけですね。

幸いウェブとか IT とかやってる方たちは知識を仕事にしている、本ってものの自体は比較的親しみはある。ただなかなかアマゾンで本を買っていて図書館に行くことはないですという人が多い中で、「カーリルというサービスを試してみるか。やってみると意外と出てくるね」っていうところがすごく反響を呼びました。正直予想外で、なんとなくこう図書館と IT って、僕たちとしても面白いものをやっているんですけど、こんなに食い付きがいいんだってというのが最初の印象だったわけです。

その 2010 年の 4 月、API という形でカーリル自体、全国の図書館 OPAC から本の情報、所蔵データを集めて統合的に情報提供していくという仕組みを作ったのですが、このカーリルの後ろの仕組み、からくりの部分全てを無償で提供しますということを開始しました。カーリルをやっているだけじゃ面白くないと、もっと面白いことをみんなやってよというサービスを開始したわけです。現在スマートフォンのアプリを「図書館の検索」といった言葉で検索するとたくさん出てきます。図書館の本を検索しますみたいなアプリがいっぱい出てきて、カーリル以外もたくさんあるんですけど、こういったアプリのほとんどが

カーリルの API を使って図書館の本の情報を提供しているという形になっています。

その2年後ですね、2012年の6月に株式会社カーリルを設立しました。実はそれまでも別の会社の一事業、実験的プロジェクトとして取り組んでいたわけなんですけど、ウェブサイトを作る会社ではなくて図書館に取り組んでいく会社というのをとにかく作ろうということで、関わっていたメンバー全員で株式会社カーリルを設立しました。

【図書館 API の提供】

今日はこれから、図書館 API から共同保存みたいところで今カーリルが取り組んでいる話、先ほどいろいろと勉強させていただいた総合目録の話に繋がる話、カーリルが何かできることはないだろうかということをお話しさせていただきます。実は答えがあるわけではないですが、こんな技術、ネタがありませんという話をいろいろさせていただいて、一緒に考えていくきっかけになればいいなど、逆に思っていることがあれば教えていただけたら、もっと面白いことをいろいろできるのではないかと思います。

これも先ほど説明がありました但僕自身は岐阜県の中津川市というところに住んでいて、会社も岐阜県中津川市にあります。コンピュータと出会ってかれこれ20年以上経つわけですけども、図書館という分野においては5年生になったので、そろそろだいぶベテランになってきたと僕自身は勝手に思っています。初心者ですとは言えないだろうと、結構詳しいぞと思って、「除籍」とか言えるし、そんな感じです(笑)。

図書館ネットワークの話はどうやっていこうかなといろいろと思いついていくと、最初にカーリルのサービスを始めた時、僕自身「相互貸借」という言葉を知りませんでした。言葉を知らなかったというだけではなくて、そもそも中津川の図書館に行ったら県立図書館の本が借りられるなんていうことは夢にも思わないわけです。何となくそういうサービスがあることは認識していて、それは県立図書館が本を送ってくれるサービスなんだという認識でした。図書館同士が繋がっているということはそもそもわかってなかった。そういう中でカーリルというサービスを作ったところが正直なところなんです。ただですね、中津川、車があればどこにでも行けるということなんですね。実際、周りの人にいろいろ聞いてみると、中津川の図書館は本がないので隣の恵那市まで行っていますよ、僕は名古屋市まで行っているんです、みたいな人が結構いる。車があればどこにでも行けるのだから、自分でどこに本があるか調べることがもっと便利になれば、図書館って活用の幅がもっと広がるんじゃないのかと、

そんなことを考えながらカーリルというサービスを作りました。だからカーリルというサービスで、「これは横断検索するサービスです」とは実はどこにも言っていない。図書館の本を探すということは、そもそも自分が使える図書館がいっぱいあって、それをまとめて探すということが図書館の本を探すということですよ、というようなそんな思いでサービスの設計をしたということです。

実際、カーリルを最初に作る時にみんながどうやって図書館の本を探しているんだろうかと考えました。本をどこかのウェブサイトで探してきて、ISBNとか本の書名とかを図書館のページにコピーしているという作業を、図書館の利用者もかなりやっているということがわかりました。自分の使える図書館がいっぱいあるとしたら、そこに順番に張り付けていって、ここにあったということを探る。これを一発で出来たらいいというのが本当に最初の、非常にシンプルなアイデアだったわけです。そういう意味では、自分たちがどんなウェブサービスがあったら図書館をもっと使えるだろうかというところからカーリルが出来てきたということになります。ただ、使い方とか言葉とか画面とかもばらばらで、図書館ごとに使い方を覚えないといけない。これが結構ストレスだなと、利用者目線で作られてないなっていうことを僕は感じたわけですね。

図書館では Web-OPAC に所蔵の有無、もしくは貸出中といった情報を公開していますので、もうデータはウェブに上がっている。だったらそのデータを解析して、統一的に、もうすべて同じように扱えるような後ろの仕組みがうまく作られれば、新しいウェブサービスでどう料理するかということだけだと考えております。カーリルを始めた当初はアマゾンの本のデータとリアルタイムに各図書館の Web-OPAC から集めた所蔵情報、これを組み合わせるというモデルでサービスをスタートしました。実は図書館の書誌は全く使っていないで、アマゾンの書誌を使ってそれにその ISBN があるか無いかということをつなぐことで、全体的には速く便利に見える仕掛けが作れているのではないかと考えました。実はカーリルを始めて最初に「TRC-MARC が」と言われて、MARC ってなんですか、MARC を買っているとかよくわからない、アマゾンでただで取れるじゃんみたいな話だったわけですが、そういうことをいろいろと図書館の人に勉強させていただきました。そういう中で、今ではアマゾンだけではなくて国立国会図書館とか、NACSIS とか、洋書で言うと WorldCat、TRC-MARC もブックポータルという形で借りていますので、アマゾンにない本を検索してもだいたい出てくるようになりました。これについてはまだまだカーリルにはやることはいっぱいあるのですが、一歩ずつ一歩ずつ進んでいるところになります。

各図書館、今日現在で 2200 の図書館システムが動いていて、そのうえでカーリルに対応しているもので 6700 の図書館の所蔵が管理されていると、これを「スクレイピング」と呼びますが、各図書館がそれぞれに作っているウェブページ

のデータを解析して同じ形に変換して持ってくる、こういう仕組みを作っています。これをカーリルではカーリルの図書館 API と呼んでいるのですが、仕組み自体はカーリルで使うのと同時に、無償で誰でも使えるように開放しているというサービスになっています。

【見えてきた次のステップ】

サービスの設計というところを説明していきたいのですが、大きく、カーリルを始めた時に目指したところとしては、「本を借りるためのパーマリンク」を目指したいと考えました。パーマリンクという言葉にはあまり馴染みがないかもしれませんが、固有の URL、アドレスのことです。一番イメージをしていたき易いのが、本を買うならアマゾンの URL を貼り付けてこの本だよと紹介すればいいじゃないですか。ただ、図書館の本を借りるときにこのアドレスを誰かに渡して、ここから本を借りたらとは言いにくい。それは利用者によって、使っている図書館が違うので、図書館によってアドレスが違ってしまふからです。ということで、本を買うならアマゾンなのだけど、借りるならどこっていうところは、サービスとして面白いのではと考えました。紀伊國屋とか hon.jp とかじゃないんですよ、世の中の的には。いくらがんばったって、現状ではアマゾンにリンクするんです。それは今後入れ替わる可能性はあるんですけど、同じメタに関してパーマリンクって複数いらんないわけです。この検索結果見てよってことならグーグルでいって話です。そのパーマリンクってすごく重要なことだと思います。ここについて、別にカーリルがなにか図書館のなかですごく主張していくっていうことではなくて、たとえば NDL がそれに成り得る可能性もあると思いますけども、今、現実的にはかなりそこになっているのかなと思っています。

例えばブックログのほうで下にいくと、カーリルで図書館の貸出状況を見る、みたいなリンクがついていたりとか、その下、国立情報学研究所で図書情報を見るということで、大学と公共みたいな形でも使ってもらっています。メディアマーカという、これも本棚サービスなんですけども、アマゾンでみる楽天で買う、カーリルで借りる、古書を見る、のように、こういう感じで図書館っていうものをいかにいろんなサービスの中に入れていくかということにカーリルでは取り組んできました。これも TRC ブックポータルなんかでいうと、右下に“図書館で検索”ボタンがもう付いていて、ここに本の情報がでます。これもカーリルの API を使って実現されています。これは本屋さんと言っているかわからないですけど、一応本屋さんが図書館の本を紹介するっていう、TRC だけからできるということもありますけど、こういうこともできるんだというふう

になります。例えば、CiNii が大学図書館の本を紹介したあとに、下で公共図書館ならカーリルっていう形でリンクをしたり。実際に今、大学の図書館からカーリルに連携する事例っていうのはかなり増えてきています。

この辺をまとめていくと、カーリルのやりたいことというのは、そんなに大きなことは目指してなくて、図書館を使うきっかけを作りたいなと考えているわけです。だいたいの方はグーグルで検索していて、グーグルで検索する時に、何々図書館で検索する人はいないと、通常、何々に調べたいとか、この本っていうのをテレビでやってたから読みたいとか。そういった検索をしているわけですが、それだけではなくて、いろんなサービスだとか、ブログの記事でカーリルを、本を見つけたときに、アマゾンっていうのはだいたいリンク貼られていると思うんですが、ユーザーにとっても便利だから、それ読みたかったから買えるよねって。ただ、それを借りれるよというところとして図書館というところもひとつあるよということで、その本が、近くにあるよということをどう知ることができるのか、図書館というものをまずは発見してもらおうと考えていると、こういうことを目指していきたいなと思っています。

【図書館 API】

それで、カーリルの図書館 API っていうところに、もう一回戻って、いろんなデータをカーリルは提供しています。1つは、図書館の基本情報。どんな図書館が全国にあって、その図書館がどういう経度・緯度でどういう住所でどういう電話番号でホームページのアドレスどれですかという情報。それと、所蔵データですね。ISBN ごとの、これも実は ISBN しか対応してないわけなんですけども、所蔵データというものを提供しています。これは、全国の 6000 館以上、6700 館の本の所蔵データがすべて同じフォーマットで入手することができるという形になっています。実際 API を呼び出してみると、図書館のデータが API の形で要は機械可読なのでいろいろ処理できる形で入手できるという形になっています。これと同じように図書館の所蔵データも簡単に手に入るという仕組みになっています。大きい話は、実はここに書誌が含まれていない、ということなんです。カーリルの API というのは、書誌を提供していないということになります。実際、この API を使っているいろんなアプリが出ています。これは本当にごく一部いろんなアプリ、例えば国立国会図書館サーチとかもカーリル API を使ってもらっていて、ほとんど気づかれないですけど、右上に図書館の設定っていうのがあって、実は公共図書館を選択できます。これはカーリルの API が後ろで動いていて、実は所蔵データも公共図書館のデータにあわせてだせるという機能が、「ゆにか」に入っていないなくても出せる、そういう機能が付いてい

ます。このブログ記事とかを見ると、「本好き必見！図書館を利用する際に使える6つのアプリ！」とでていて、もはや図書館アプリは競争の時代を迎えていると言えます。カーリルがタダでデータをどんどん提供するので、100以上の、実際動いて提供されているアプリができています。これはスマホアプリだけに限らず、Webサービスだとか、いろんな分野で1,000人以上のひとに開発者の登録をして、カーリルのデータを使っていただいています。勝手な自画自賛ですけど、たぶん国立国会図書館よりよっぽど使われているAPIを提供していると考えています。こんなカーリルAPIをどうやって提供しているかということ、ちらっと見せてみます。これがカーリルのバックエンドで僕たちが普段カーリルの為に使っている画面です。どういうことを僕たちがやっているかというと、1つは、全国2,200のシステムが毎日メンテナンスをされていて、切り替わり、その切り替わった度にプログラムを書き直したりとか、データを修正したりとか、あるいはホームページのアドレスも毎日どんどん変わる。これを図書館からの申請とか通知とか一切なく、いかに正確な情報を提供できるかということに日々努力している状況です。もう1つはですね、ここで例えばレファ協のデータと郵便番号が違うとか、国立国会図書館が作っているISIL(アイシル)のデータで廃番扱いになっていますよとかでているんですが、こうやっていろんな機関が作っている他の図書館のデータベースと自動的にマッチングして、例えば、閉館したよとか移転したよとか、住所が違うよっていうことを自動的に検出するといったこともやっています。実際それだけではなくて、カーリルを使っていると「エラー」ってでることがあると思うのですが、このエラーの率が異常だっていうケースを検出して、何かシステムが変わったんじゃないかということも自動的に認知できて管理できるという仕組みを作っています。ここで、所蔵場所というものが図書館システムによって分館扱いだったりBMだったりいろいろと登録していると思うのですが、こういったものが日々登録名称を変えたり、短縮名だとわかりにくいので長くしようというって変わるわけですけども、変えた瞬間に、たぶん数分後にはこの画面に拾います。そうすると、うちのほうで、この何々分館っていうのはこの図書館IDでこのISILの番号を持っていてということをもう一度紐付けし直す、という作業を裏で行っているという仕組みになっています。これによって、検索してカーリルで検索したときに、ちゃんと図書館の情報に繋がって所蔵をみることができるということになっています。こういったいろんな管理をしているんですけど、こういったものも、今全部APIで結果提供されているというサービスです。

ここから先はですね、勝手に想い描いていることなんですけれど、カーリルの所蔵データというのはISBNに紐付いた所蔵がある・ないだけでなく、貸出中とか予約中とか、Web-OPACに載っているデータになります。これも全部同じ

形にしています。例えば図書館によって予約中の表現はみんな違います。あと、在庫ありという状況も「在庫あり」って書いてあったり、「在」であったりと、配架中の文字の標記すら5種類くらいある。そういったものをカーリルのなかでは、貸出中、予約中とか貸出可というように、同じ表現に全部変換をしています。こういったことによって、全国の貸出率を見るとかということが簡単にできるようになってきました。例えば出版社が、全国でどういうふうに自分たちの本が読まれているんだろうかということを確認するために使っていると聞きました。それによって全国で自社の本がどれくらい売れるのかであったりとか、こういうジャンルの本を出すとどれくらい売れるのかみたいなことを予測したりとか、あるいは、それをもって営業に行っているという話でした。こういったところに使うといいのではないかという話もしています。

もう1つはですね、図書館で使えないかということにもカーリル積極的に取り組んでいきたいと思っているわけです。これまでも、図書館システムの会社はたくさんあったと思うのですが、彼らは図書館を客として、基本的には図書館に向けてシステムを売る商売をしていたかと思います。ですが、少なくともカーリルにはそのつもりはさらさらないと考えています。図書館にシステムを売ることにはそんなに興味はなくて、カーリルというのは図書館サービスをやっているのだということを社内でもよく聞きます。本が見つかるということは図書館サービスの一部であって、本来もしかしたら図書館がやる仕事かもしれませんが、これは私たちが責任を持ってやっていこうじゃないかと、それぐらいの責任を持ってやっていきたいよねという話は常々社内です。なので、図書館はたぶんパートナーであって、図書館サービスを共にやっていく立場で、そのように考えると、例えば除籍の支援や、先程の共同保存のような話は実はカーリルにとってすごくメリットがある話なんですね。つまり、カーリルを使って見つかる可能性が上がるということは、図書館サービスそのものがとても良くなって、やはりユーザーが図書館を信頼してくれるひとつの大きなきっかけになるのではないかと期待をしているというところなんです。実際、出版社などでカーリルをどのように使えるかということで、今試作のシステムをいろいろと作っています。カーリルだと、今は都道府県ごとでしか見えないのですが、全国ではどうなっているのかという情報が約1~2分で見られるというところまでできています。例えば『1Q84』が全国で25,000冊あって、そのタイミングで58%が貸出中です、というような、いわゆる資料の状態を含めたものが見えるようになってきています。そして、どんなところに買われているのかということも、実はここを見ると、公民館図書室がかなり買っている、ということがわかってきます。これも勝手に言っているという段階ですが、選書づくり、分類みたいなのところにも興味があります。今、多摩地区で除籍を

最適化できないかというところも共同研究していますので、今から少し見てみたいと思います。データ分析で除籍を支援したいと考えています。これはまさに今話があった大先輩の埼玉の ISBN 総合目録などもかなり見させていただいている訳ですが、東京都においても 10 年ぐらい前までは東京都の目録をやっていたと伺っています。ただそれが無くなってから、基本的にはみんな勝手に本を捨てていて、実際には、やっているところはちゃんと周りの所蔵を見たりとか、都立図書館の所蔵を見て捨てているところもあれば、何もせずに捨てているところもありますというようなことが、いろいろアンケートなどでわかってきたりしています。

そういうなかで、なかなか前に進んでいかないと相談を受けまして、カーリルの技術を使ったら出来るのではないかとということもやってみました。やったことはすごく単純です。国立国会図書館と NII の和書、これを合計すると 180 万冊ぐらいになるようなのですが、それを多摩地区のすべての図書館 30 館、30 自治体ぐらいあると思うのですが、総当たりしようということを進めています。これができる、1990 年に出版された本のラスト 1, 2 になっている割合はどれくらいなのかみたいなことがわかるようになりました。これがひとつ特徴的なのは、図書館から一切データ提供を受けていないという点です。図書館はノータッチ、図書館は一切作業をしていない。カーリルをぽちっとやっただけで出来たという状況です。こういうことができると、データを自分たちで集めて図書館は公開だけをすればいいということになります。使う人は自由ということで、例えば、別の自治体の情報はいくらでも引っ張って来られる訳ですね。埼玉じゃない自治体の情報でも使うことができます。

これからどうやって自分たちで考えていくかということと、図書館がどれだけオープンに開放していくんだということを考えていくか次第でもありますが、今まで大きいところにまかせておけばいいという仕組みがいろいろあったのだと僕自身は感じているわけなのですが、ここまで来ると、もはや誰でもできるようになっていく。みんなで寄らなくても、1 館がやりたいと思えばできるぐらいのレベルの話になるんじゃないかと期待しています。

先程のツールのところで今議論を進めていて、例えばバーコードを読んで除籍のただラスト 1, 2 を見る目的で話をしているんですけど、バーコードを読んで調べたいとか、除籍候補のエクセルデータに他館の所蔵数を追加したようなデータに変換してくれないとか、システム会社とデータ形式について調整したくないとか、こういう話をいろいろと議論しています。そういうなかで、多摩地区のトライアルというか共同研究のなかで図書館に提供していこうというところできているのは、とにかく、どんなデータ形式でも Web でアップロードすれば他館でこれを何冊持っているのかというのがついて返ってくる

みたいなことができないかというような話をしています。

この辺はまとめになるのかもしれませんが、オープンデータというのは結構重要と考えていて、今まで共同利用していきましようという話はいろいろあったのですが、なかなか進みませんでした。ニーズもそれぞれで、足並み揃えるのもなかなか難しいということでした。ただ、もう図書館ではWeb-OPACでデータを公開しているので、オープンにするということは実はやっているというところで、みんなでいろいろ考えていこうよというのは、みんなでもできるし、ひとりでもできるし、どんな組み合わせでもできるのではないかと、そんなことを考えていったりするわけです。

【ビッグデータ】

こういう話の基盤になる部分というのが、最近よく聞くかと思いますが、ビッグデータという話があります。例えば、カーリルが提供する所蔵情報、本があるかないかという情報は、1分間に約1000冊分カーリルを通過しています。これは、図書館に取りに行くと、ユーザーに渡している、あり、ない、ないという情報を冊数でみていると、1分間に1000冊流れてきます。この1000冊のうちの、実は7割ぐらいがカーリルのWebからの利用、残りの3割ぐらいがユーザーの作ったアプリの利用です。いろいろな使われ方をしていることを考えると、1日に150万冊ぐらいのデータが動いています。年間にすると5億冊ぐらいのデータが動いていると。これがデイリーの規模です。それに比べて、日本の公共図書館の蔵書ってというのは、日図協のデータを見ると4億冊ぐらいあるらしいです。一方で、カーリルの通過している150万冊のデータをユニークのISBNにすると何冊ぐらいあるのだろうかというのを調べてみました。そうしたら、ニーズがあった本というのは10万冊ぐらい。一方で、世の中にある本っていうのが和書だけで、180万冊。何となくデータの的にいろいろ見えてきたぞと、こういう状況があります。カーリル対応システムが2000あると仮定して、180万冊とすると、36億件。あるいはこれを図書館の数になおして、6700館の180万冊分、あるない、総組み合わせ、総当たりの情報ですので120億件のデータをもってくれば、ざっくり日本の全国の総合目録ができるということになります。

今まで、こういう大規模なデータを処理するためには、それを処理するだけのコンピュータが必要でした。それが、なかなか始められない要因だったわけなんですけど、いろいろ技術が進歩して、みんながクラウドに載るようになったことで、例えば、3000台のコンピュータを10秒借りるみたいなレンタルが今可能になってきました。この技術を使うと、単純に1台だと24時間かかる処理っていうのを10秒でできるようになって、何かできることが広がったぞという

のが今のビッグデータの基本的な議論です。ではこれにどれくらいお金がかかるのというと、1台のコンピュータを8時間借りるのと同じなんですね。秒数に直して考えると、正しいかわかりませんが、3000台のコンピュータを10秒借りるってことと1台のコンピュータを8時間借りるのは一緒なので、サーバを一か月借りると1,000円とかいっているの、だいたいこれにかかるコストが見えてきて、基本的にはとても安いという訳です。こういうことができるようになったことで、この3000台を3万台にしようがコストは変わらない状況になりました。

今どのようなビッグデータの技術になっているかというと、やる処理の量に合わせて台数が決まって、秒数は一緒という合わせ方がされているシステムが増えてきています。どんな難しいことをいっても、とにかく2秒で終わると。ただ、コストが違うだけだよと。というのは、台数が変わるので、大変処理量の多いことをやるために10万台のコンピュータを3秒借りると、そういうことができるようになってきたということです。今はこういうところで、カーリルのデータをいろいろと解析で使っています。

まだまだこういったものは始まったばかりでごくごく一部ですが、この多摩地区のデータ収集というのが入っているの、さっき言ったようにかなり量が多いです。昨日1日でカーリルを流れた冊数っていうのが470万件。その前の日が390万件というような形で、このデータがずっと整理されて、蓄積されるということが技術的に可能になってきました。これの活用はカーリルのサービス上では全然できていません。唯一使っているのは運用の管理のためです。過去にどういう所蔵があったのかということの変化を見ることで、そのデータがちゃんと出なくなったとか、例えば、ここで佐用町三日月文化センター図書室が30日で232冊あったのに過去5日間で1冊もないといったことを検出して、これはおかしい、この図書館は閉館になっているのではないかとといったことを検出したりと、このレベルですが、今活用が始まっています。こういったちょっとしたことですけど、いままでこういうことをやろうとすると結構大変でした。データ量が多すぎて、その中から、こういう小さな話を見つけ出すのがなかなか大変でした。こういった処理をするために、1回あたり2円とか3円とか払えばいいという話になったので、全然問題なく使えてしまう。そういう時代になってきました。ちなみに、グーグルのこういった分析サービスの値段が1テラバイトで5ドルという単価になっていて、NDLのデータは2ギガにもならないので、0.0003円くらいで処理できるということですね。そもそも図書館のデータなんてビッグデータではないということになってきているという現状があります。ちょっとこういういろいろとサーチをしながら、妄想していたということになります。

今ある、総合目録、ISBN 総合目録とカーリルの API を組み合わせたら何ができるんだろうかというのは、そこで書いていたのでは十分ではないかもしれないですけども、何となく流れが作られていたような気がするので、こう考えてみました。1つはですね、先程紹介のあった総合目録のデメリットが解決できるということです。さっきの総当たりのときに実はすごく有効だなと感じました。実は総合目録があれば本当にすべてを総当たりする必要がなくて、まず、どんなデータでもあるものだけ総当たりすることができます。これによって、最新の状態っていうのに塗り替えることができるということが、実はかなりリアルタイムにできるのではないかとことを考え、感じました。そういう意味では、事前作業っていうのを無くすというところまではいかなくても、最小限にする、もしくは2年に1回データが来ればよいといったことができる可能性もあるかもしれません。なぜかというと、1回動き始めたらかなり自動化できるということです。例えば、所蔵が少なくなっているものに関しては、一か月に1回 OPAC の上で見ましょう、あるいは、潤沢にあるものに関しては、そんなに重要ではないので、三か月に1回見ましょうといったことで、図書館にデータを見に行く回数を戦略的に減らすことによって、どんどんコストダウンすることが出来そうだと考えられます。

いろいろ出来そうなことはありますが、一番いいのは、総当たりしなくていいということで、その後の差分というのは実は簡単なんですね。新しい本のデータを年間に出ている本の数8万冊とか多くて10万冊なので、それをもう1回入れなおせばいいだけで、実は最初の200万冊を作るのがなかなか大変なんだけど、そこは実はすぐできてしまうのではないかと考えられるのです。ということで、組み合わせで適度に ISBN もちゃんと集める仕組みができると相乗効果でコストダウンできて最新でという、ある程度制約はあるんですけども、そのなかでどうしたらベターかという話がいろいろ考えられるんじゃないかということも思ったりしました。

ただ、問題になっているのは、ISBN が無い本への対応だろうなということも思っています。実はカーリルも始めた当初から ISBN の無い本が探せないということを常に言われていて、いろいろ努力はして、カーリルローカルみたいな仕組みも作ったりはしているのですが、まだ探せていないものがあり、実は全然満足していない訳です。そこに横断検索モデルの限界というものがかかなりあるんだろうなと感じます。

一方で、大半のものは ISBN があってみんなが探したい情報というのは完結してしまっているというところもあると感じます。それで、なかなか進まないというところがあります。ここに関して、いろいろな方向からアプローチはしているのですが、最近始めた別のプロジェクトでちょっと近いものがありまし

た。それは電子書籍に関してで、最近全国的にいろいろなところで図書館が電子書籍サービスというのを始めていますけれども、このサービスは書誌データがばらばらになってしまっていて、底本の情報がちゃんと無かったり、あるいは、図書館の検索とうまく連動できていなかったりと、いろいろな問題があります。カーリルで検索したときにも、紙の本はちゃんと出るんだけど、電子書籍については ISBN もなく、なかなかうまく探せないという状況があります。でするので、まだ正直数が少ないこともあって、すべての書誌をカーリルで解析して同定をするという作業を今やっています。どういう仕組みになっているかというところ、ここに、各図書館の電子書籍サービスごとの所蔵数と種類ってというのがでています。副産物としてどういう出版社が入っているかとか、そういったデータも明らかになってきたわけなんですけども、同定作業というのをしますと、上にある書名に対してこれが ISBN じゃないのかという候補を出して、紐付けていくという作業になってきます。実はこの仕組みも一切図書館とのやりとりはしていなくて、僕たちが勝手に図書館の作った書誌の上に目録をさらに作っているという状況です。こういった仕組み、ISBN の無い各図書館の書誌をやっていく仕組みというのが出来ると ISBN の無い本についても割と正確な同定が行えるのではないかと思います。

こういったものを進めていく、こういった作業を図書館でやるというのはたぶん困難だろうなというふうを感じる訳ですが、こういったところを、例えば図書館の経験がある方など、いろんな立場の人で分担してやっていくことで、面白い仕組みってできるんじゃないかなと感じます。この紐付先って例えば国立国会図書館の書誌であったり、そういうものに張り付けていけば、おのずと国会図書館に無い本っていうものが見えてくる。そういったものを、図書館の外側、図書館システムと離れたところで作っていくことで、図書館にとってもシステム更新をしたりとか、そういったときもあまり考えなくてよくなると思います。いろいろ難しい仕組みを作ってしまうと、システムリニューアルなどのときに、どんどん大変になってしまおうと。そういうのを、なるべく粗い結合をしながら作っていく、つながっていく仕組みというのが、これから求められていくんじゃないかなということを感じています。この辺をカーリルとしては是非全国のユニークな本、ここにしかない本っていうのがちゃんと検索で引っかかるようにしたいというところを考えていまして、このあたりについては、図書館と一緒にやっていくということが必要不可欠なんだろうとも感じています。ただ、図書館に一方的に負担を求めるという話ではなくて、みんなで分担しながらいい方向に、ベストな方向に、ベターな方向に行けるやり方があるのではないかなということを感じています。

【カーリルの目指すもの】

カーリルのやりたいこととはどういうものかというところで、決してですね、便利にしたいと思っていない、むしろ図書館をもっともっと面倒なものにしていきたいと思っています。これはなぜかという、先程のビッグデータの話というところとすごく関連していて、実はもうある意味では、カーリルが今持っているデータ、入手できるデータ、もしくはカーリルでなくてもカーリルのAPIを使った誰かが頑張れば、いい図書館を作ることができるようになっていくというところが思い至るわけです。どういうことかという、例えば、本の回転率が高いとか、貸出率が多いとかいうところを目標に設定してあげれば、たぶんコンピュータが自動的に今日朝から並べるべき本とか買うべき本とかは全自動でやっていけます。その図書館がいいかどうかは別として、確実に数字が伸びることは自動化できるということは解ってきたと。じゃあ、それでいいのと言われたときに、カーリルとしては、そんなのはもう当たり前と言って、その先に行くことを考えているということです。それは何かという、正にこういった目録みたいなものを多分図書館が積極的に活用して使っていくというところができないとコンピュータの方がいいという話になる。どうやって運用していくかということについて、多分図書館と一緒にできることがたくさんあるのではないかと考えています。そういう意味でも図書館をもっともっと面倒なものにして、何か図書館をやるということはいろいろ考えなきゃいけないと、本を一つ買うにも本を一つ捨てるにもいろいろ考えなきゃいけないというところを理解し、それを図書館それぞれで考えるのではなくて、こういったデータを使うことで、他の図書館がどう考えたのかとか、他の図書館の利用者はどういう傾向なのかということや小さい図書館でも分析ができるようになる、つまり、大きな図書館と同じ知見を持てるようになるんじゃないかなと考えています。こういうところにカーリルのAPIを使ったデータというのがうまく使われる仕組みというのを目指していきたいなと思っています。

カーリルは、「図書館をもっと楽しく」ということを、今まで会社設立以来テーマに掲げてきました。カーリルを最初に始めたときに、第一印象ですね、案外図書館の人が暗いと感じました。何かこういういろいろ出来ないということや言うので、カーリルとしては、出来るんじゃないかということや、ちょっとずつちょっとずつやってきたと考えています。ただ、まだ出来ていないことがいっぱいあります。実際、ISBNを持っていない地域資料はカーリルはちゃんと探せないし、そんな当たり前のことがまだ5年経っても出来ていないということがあります。ただ、元々僕自身もWebの技術、新しいそのベンチャーを立ち上げるみたいな仕事をやってきたので、そういったところ生かしながらですね、

そういうことをやっていきたいなと思っています。そういう意味では、多分これから図書館といっしょにやっていくことがいっぱいあるなということを考えています。用意してきた話はここまでで、時間的にもいい時間になったと思いますので、一回話を締めたいと思います。

《質疑》

講師：図書館でデータ活用していきたいなと思う人はどれくらいいますか？

・・・会場参加者の挙手・・・

講師：半分ぐらい。カーリル、正直この先、どういうビジネスになるかなんてほとんど考えていなくて、とにかく図書館が10倍、100倍使われるようになってほしいと思っているのですが、うちは実際の現場を持っていないくてですね、こんなこと出来るんじゃないかと言っているだけになってしまうので、是非フィールドがあるよという方は声を掛けていただければ、どんどん一緒にやっていきたいなと思っていますので、よろしくお願いします。

質問：鶴ヶ島市の中央図書館藤田と申します。質問なのですが、前半のISBN総合目録の間部さんとWebの方のISBN総合目録の話をお願いした上で、現在のカーリルで総合目録に替えちゃえばいいのではという感じには思われませんでしたか。

講師：総合目録に替えるというか、何となくですが、カーリルもいろいろやっている結果からそういう総合目録になってきている部分はあるんだろうなと感じています。ただ、たぶんデータをどう使っていくかという切り口によって、たぶん求められるところの思いが違うんだろうなというのがあって、元のデータは一緒にいいのかもしれないと考えています。

うちは外に総合目録があるからそのデータ引っ張ってきて使いましょうと考えていますし、逆にそれがなければうちのデータをどうぞといったところがあって、分散的に管理される状況というのがカーリルとしてはすごく望ましいと思っています。

というのは、国立国会図書館が大きい何かを作って運用するモデルというのは有り得ないというふうに思うわけですね。たぶんそれって運用されないし、進化が止まると考えています。カーリルなんかいらんじゃんといって、新しいものが生まれる余地も残したいところがあって、大きな1個になるというよりは、カーリルというのは、カーリル自体をどんどん使ってもらおうと

ということと同時に、要は巨大なものになるというよりは小さなものがいっぱいできるというモデルが私としては面白いなと思っています。そういう中でたぶん被る部分もいろいろあって、共通化して効率化できる部分もたくさんあると思うのですが、いろいろな組織でいろいろなものが運用されて、それぞれにいろいろなトライアルができるっていう状況をどうやって担保していくかというところが私としてはすごく興味があるところかなと思っています。

カーリルとしては全てオープンでやりたいし、どんな使い方をされてもいいと思っているわけですが、いろいろな思いがあるところがあって、いろいろな運用があっただけいいのではないかなというところは感じています。ちょっと曖昧な答えにはなってしまうのですが、いろいろな取り組みがされるっていうことがすごく重要で、たぶんすごくいいものがあるので、みんなそれに乗ってしまっただけじゃないというモデルは、私としてはあまり面白くないなと考えます。いろいろな挑戦をする人が出ていける仕組みを作っていきたいなと思っています。

これはどういうことを意味するかというと、是非ですね、埼玉県の ISBN 総合目録がオープンなデータになってくると全然意味合いが違くと、カーリルはこのデータを使って全国の図書館の本の検索の精度を上げるためにももしかしたら使えるかもしれないということになってくるかなと思っています。そういう意味では、すごく期待をしています。なんとなくカーリルとの間でそのデータがただ融通されて、馴れ合的に使われるのは面白くないじゃないかと。これは、カーリルと同じような第三者機関が何か新しいことをしようとしたときに、阻害要因になるというところを僕自身は思っています。なんとなく公共とカーリルとの関係性というのは常にオープンなものとしてやっていきたいなというふうには思っています。

質問: 埼玉県立浦和図書館の湯川と申します。今日はありがとうございました。今日のお話とは直接関係ないかもしれませんが、いただいた資料の中にあるハイカのプロジェクトについて、お伺いしたいのですが、現在の進捗状況と我々が日夜カウンターをやっていて困ってしまう、お客様は検索機に行って自分で調べて自分で資料の位置にたどり着こうと努力をされているんですが、あちこち探して、結局わからなくて、最初から聞けばよかったよねと言いながらカウンターに来て下さる。あちこち回ってしまっているから検索機から見た位置がここだって表示されても、もうわからなくなってしまうので、自分がいる位置と資料の位置関係っていうのが常に把握できれば最初からあちこちうろする必要もないというのが、こういったシステムのいいところかなと思っているのですが、現在のそういった取り組みの状況とかについて教えてい

ただければと思います。

講師：大変すどい突っ込みありがとうございます。まさにいい質問で、カーリルでやっぱり問い合わせの多いことの一つとして、図書館に行ったけど無かったという話があるんです。図書館に聞いてほしい話なのですが、なぜかカーリルに問い合わせが来ることってよくあって、カーリルの情報が間違っているのではないかという問い合わせがあるわけです。意外と聞けないんだなというところと、丁度このハイカは、大学と一緒にやらせていただいているのですが、大学はいろいろ切実みたいで、カウンターでここにあるよと言っても、辿りつけない状況が多いという話をよく聞きます。それは図書館にとって楽というのがあるのですが、案外、利用者にとっては、探しても見つからないから行くのがいやだという話になってくることが多いそうです。グーグルで検索するとそれで Web に出てくるんですけど、図書館に行くと Web では出ていたにも関わらず無くって、期待を裏切られるというところにめげてしまう。そういうのがあって、なんとかしていきたいなということを取り組んでいます。それをやる為にですね、今もたぶん OPAC で配架図を出したりする機能はいろいろあるとは思いますが、カーリルが今取り組んでいるのは、とにかく図書館の配架図のデータを抽象化して全国で同じようなデータ形式で扱えるようにしていきたいなということをやっています。さらには、それをいわゆる通常のアプリのマップの中で工夫していけば、図書館の配架図が出るという状況をなんとか実現していきたいということで、いろいろなところといま話を進めています。

今の進捗ですが、技術的にはですね、図書館の中で自分の位置がわかる技術というのをかなり進んできていて、この領域はもうアップルやグーグルが最先端で研究を進めている分野になります。こういう技術を使って実証実験を、名古屋大学で進めています。今カーリルがやっているのは、いかに図書館で管理が楽になるかということで、一回こうマップを作るのは簡単でいいですね。現状図書館では、例えば配架図をエクセルで作ってエクセル配架図で管理されているのがすごく多いと思うのですが、それを Web に飛ばしたりとか Web-OPAC で見たりとか、PDF にしたりということをいかに効率化できるかという、エディタの開発を進めています。

これは、いちおう今名古屋大学では動いているのですが、このモデルはまだ非常に手のかかるレベルで、今もう何館か公共図書館で実証実験をやらせていただく予定ですが、そこでの一番メインの取り組みというのは、いかにデータの管理が継続して図書館に負担なくできるかということです。それがそこで配架を変えた時にちゃんと地図会社のサーバとかに同期されて、当たり前図書館の本をスマホで探せるという状況を作っていくというところを目指しています。今年、来年ぐらいには自信を持ってこれ使ってみてくださいと言えるこ

ろまでもっていきたいと思っています。

実はこの配架図に関しては、なるべくタダに近い、もしくはタダで提供したいと考えています。つまりカーリルのなかで、本の場所まで案内するところを、図書館の人がちゃんと協力してくれたら、Web-OPACとか、そういうところでもカーリルの開発した技術をどんどん使ってくださいねというところを進めていきたいと思っていますので、その辺を是非、また1,2年後になるかもしれませんが、協力いただけると幸いです。もしくはもっと、早い段階で手間がかかってもいいけどやってみたいというところがあれば、お声掛けいただけるととても助かります。がんばっていきたいと思います。



講師：もう一点だけいいですか。補足なのですが、さっきの配架のところでは実は一番注目しているのがNDCです。図書館ではNDCの分類で本が置かれていて、このNDCの情報と位置情報がちゃんと紐付くことができるので、例えば本屋さんがアプリをもって図書館に行けば、その図書館、本屋さんアプリで本が買うことなどできるよう

にもなる。つまり利用者はNDC、自分がいる棚がどういうところかという情報がオープンになっていけば、それをどう使うかは自由で、いろんな使い方ができるんじゃないかなと考えています。これによって、図書館という空間自体は、たぶんビジネスという形でも、ユーザー側がなにかしようという形が実現できるのではないかなと感じます。電子書籍も全く一緒ですね。ここに関連した電子書籍がありますという話もNDCというところでどうやって落とし込んでいくか。これができるので、世の中のいろんな情報、売っている本もそうですし、電子書籍もそうなんですけども、こういったものをNDCに分類したら便利だよというところの動機付けができるというところで、NDCは今度版が変わるんですが、是非、版が変わったらオープン化してほしいと思っています。今のNDCはライセンスが明確ではなくて、自由に使えるのかよくわからないので、これは是非オープンにしていきたいと思いますという話をしています。こういうちょっとの広がりというものも、カーリルでやっていきたいことのひとつにあるところです。

質問（司会者）：私から質問なのですが、図書館によって特殊な配架というか、NDCだけじゃない配架があると思うのですが、そういったものは、図書館ひとつひとつにあって量がものすごいことになっていると思いますが、それにはどう対処しているのでしょうか。

講師：まさにですね、もう6館ぐらい、試しにデータを作ったりしているのですが、いろいろありますね。例えば、上が大人で下が子どもとか、ここは別分類ですとか、いろいろあるので、そういったものをいろいろやらせていただいて、要は毎回壊しながらモデル化しているという状況です。それで、もう一つ重要なのは、今を吸収するだけではなくて、未来にどう変化していくかということデータを形式に盛り込んでいきたいなと考えて進めています。この次のステップというのは、そもそも配架図ってどういうものなのか、ちゃんとみんな議論しながら共通の形式作っていかうというものだと思います。多分、図書館においてこれはなかなか難しいのだと思いますが、とりあえずカーリルがこんな形でやっていきませんかというのは提示するので、そこから先は一緒にいろいろやっていきたいなと思っています。まさにその多様性というところがカギになるなと思っています。

カーリルとしては、システムがこうなっているからこういう配架できませんとは絶対言わせたくないと思っています。どういう配架にも、逆に変な配架に挑戦するときにも、挑戦するっておかしいかもしれませんが、うちはNDCやめるんだという時に、多分ハイカミみたいな技術がより重要になるだろうなと考えています。それがNDCの分類だとかだと、ちゃんと別の代替手段が提示できますみたいなことができるのではないかと考えています。

質問（講師・間部氏）：すみません、間部です。いまの配架の話、面白かったのですが、つまり、図書館側が配架のルール、配列規則とかどういう順序で流れているのかということをもしかしたら完全には承知していない図書館もあるかもしれないのですが、どのように別置をして、どのように並べているのかというルールがカーリル側に伝われば、その完全な配架情報をコンピュータ上で再現できるということによろしいのでしょうか。

講師：そういうことです。やっぱり面白くって、名古屋大学で実際やって動かしてみると、案外できないものが出てくる訳ですね。OPACのデータから場所がわからないってケースがやっぱりかなり出てくる。これがいままで図書館側が聞いてほしいと言っていた部分なんです。これをいかに聞けという部分を無くすかということがカーリルとしてはすごく重要なんだろうなと感じています。それが図書館としては世の中のデータ、資料というものをちゃんと体系化して整理していくことがミッションだとすると、聞けということはいかに無くすかということがすごく重要なテーマにあるのだろうなということはやってみて思ったところです。

質問（講師・間部氏）：もう一点いいですか。そこまでデータを持てば、例えば、

本の並び順が再現できるわけですから、例えば画像的な電子図書館のような書架を再現して表示するような、そんなシステムを作れるのではないのでしょうか。

講師：面白い話を、そうだと思います。これもまだちょっと実現していないのですが、社内のプロジェクトとしてあるのは、背をスキャンしたいという話があって、背が本を同定するための最低限の情報なんです。さっきのこの ISBN の無い本の同定についても書影、実際にまあ 3 面ですね、3 面のデータを各図書館がユニークなものだけ、うまく撮ってもらえると同定もできるし、多分利用も伸びるし、特に閉架にあるものがなかなか動かないところに関して、書影だけあるだけでも違ってくるのではないかということを感じている訳です。これがやはり面倒なので、これを自動化できる機械を作りたいなと思っていて、是非そんな研究なら一緒にやってもいいという図書館があるととてもうれしいと考えています。もう一個は、たぶんこういう話はカーリルがやるのではなくて、カーリルはこうやってフォーマットなど地味なことをやるので、その華々しい部分は、それが得意なエンジニアが行ってくれればと思っています。そういう人は世の中にはいっぱいいて、使えるデータをちゃんと提供してあげれば絶対取り組んでくれる人たちです。それは、カーリルのアプリでもいろんなエンジニアが関わって作ってくれていますけども、それはデータが、使えるデータがあるからであって、そこを整理していくことが重要なんだと思っています。