

# 国立国会図書館のウェブページを使い尽くそうアイデアソン

～NDL オープンデータ・ワークショップ～

[各班のアイデアメモ]

2015年2月21日（土）

## 目次

(1) 書誌・典拠データ.....	1
(1)a コンセプト・マインドログ.....	1
(1)b 時空間と概念空間2軸で見せる.....	1
(2) デジタルコレクション.....	3
(2)a 江戸時代人生ゲーム.....	3
(2)b カジュアル・デジタルコレクション.....	4
(2)c 古地名陣取りゲーム.....	4
(3) レファレンスデータ.....	5
(3)a レファレンス協同データベースとの出会いをすてきに演出.....	5
(3)b 感情情報で本をレコメンド.....	6

2015年3月16日に「(2)c 古地名陣取りゲーム」の記述を更新しました。

## (1) 書誌・典拠データ

### (1)a コンセプト・マインドログ

- ・メディア・芸術作品に内在する各種の関係性を、コンセプトマップのような形で、分かりやすくビジュアル化するというアイデア。ある事柄の関係資料や関係概念をもれなく効率的に知りたいという要望から思いついた。資料間関係性だけでなく、資料と人、概念と人、概念と資料などをフラットな形で図示する。
- ・使うデータは、典拠データ（国立国会図書館件名標目表（NDLSH）、名称典拠）、ウィキペディア、Amazonのデータ。協調フィルタリングのような手法を使うのであれば、利用ログも使えるかもしれない。
- ・データを用いて、関係性をグラフ化する。はじめは機械で、あとでソーシャル（集合知）による手作業でブラッシュアップする。
- ・例えば『図書館戦争』は、本もあれば映画もある。原作には作者がいるので、その典拠データから、その人の別の著作を知ることができる。また、映画には登場人物やそれを演じた役者がいるので、その典拠データからその役者が演じた別の作品を知ることができる。また、撮影した場所の地名や、撮影した監督、脚本家についても、それぞれの典拠データから、それぞれの人が関わった作品や、場所に関連する作品を知ることができる。映画であれば主題歌に至ることもできるし、作品が何らかの賞を受賞すれば、その賞を受賞した別の作家の作品を探すこともできる。このように、さまざまなつながりをたどることで、思わぬつながりに至ることが可能である。また、このようにつながりをたどる手法は他分野にも応用できるだろう。
- ・ただし、思わぬつながりをどんどんたどっていくと、どのような過程でそこに至ったのかが分からなくなってしまう可能性もある（迷子になる）。そこで、たどった過程を追えるように、たどったプロセスを記録（ログ）し、他者と共有する仕組みも提供する。その方法として、データをたどった順番や履歴をスニペット表示のような形で示し、たどったデータに戻ることもできるようにする。
- ・また、何を中心に関係性を見るかというフォーカスは、（固定化せずに）その都度選べるようにしたい。多くの人々がたどる関係性を抽出しサムネイル化すれば、サジェスト機能として働くこともできる。
- ・メディア・芸術作品だけでなく、学術研究分野などにも応用可能だろう。

### (1)b 時空間と概念空間2軸で見せる

- ・サービスを考える上で最も重要なのは「どのように見せるか」である。
- ・NDLのデータを用いる際には、時空間と概念空間の二つを軸として考えるとよさそう。
- ・レファレンス協同データベース（レファ協）の回答部分に含まれる書誌情報は、現在は構造化されていないが、これをNDLサーチのAPIを使ってNDL-OPACのデータと全件突合し、著者とタイトルの組み合わせがある程度合致すれば一致したものとするなどして書誌データ化することができれば、レファレンスデータを媒介として関連資料にナビゲートすることができるようになる。

るだろう。

- 特許に含まれる文献情報を NDL の書誌全件と突合してデータ化すれば、企業での研究開発のテーマの検討や研究の際に役立つツールができるだろう。また、特許の主題データと書誌の主題データをマイニングして文献やテーマをレコメンドするツールを作っても役立つだろう。
- 大学のシラバスは書誌の書き方もキーワードも教員によって異なるが、これを書誌データ化することができれば、学術分野と参考文献のデータベースができる。そのために必要なデータ整備を NDL のデータで支援する。例えば、タイトルやキーワードの一部を入れると NDL の書誌データや NDLSH の統制語彙から自動的に著者やタイトル、キーワードの入力候補が提示されてワンクリックで選べるツールがあれば、データ整備が容易に可能になる。
- NDL の II 種資料の購入のためにやっている選定作業は、複写や閲覧の回数から、購入候補資料として推薦するツールがあれば、業務が省力化されるのではないか。
- 全国の公立図書館での利用履歴（検索履歴、貸出履歴、閲覧請求回数）があれば、ある時期におけるあるキーワードでの検索回数とそのキーワードに関連する出版物の量の関係から、出版と需要の関係を見せることができるのではないか。
- 現在は、絶版情報があいまいである。絶版の判定基準を、Books.or.jp、honto.jp、紀伊國屋ウェブストアなど複数の書籍販売情報検索サイトで横断検索した結果入手不可能であることとし、一定期間ごとに絶版判定を行い、出版年と絶版年から本の生存期間を算出する。同時に、出版年ごとの日本十進分類法 (NDC) または国立国会図書館分類表 (NDLC) の分類で出版件数を算出する。両方を掛け合わせて、何年にどのような分類の本が何件出版されていたかを、グラフィックに見せる。NDC と NDLC は NDL サーチの API を使って取得できる。見せ方としては、まず横軸に年代を、縦軸に件数を取り、何らかの閾値を設定する。そして NDC、NDLC を NDLSH と対応づけた LOD を用いて、NDLSH の言葉をタグクラウドのイメージでシャボン玉のように表示させる。検索すべきキーワードを持たない人でも情報を得ることができ、説明がなくても何が示されているかが直感的に分かる見せ方が重要である。
- 地名典拠を持つ出版物の示す他の件名・NDC・NDLC も件数を算出し、ある地域については、どういった主題の本が出版されているかといった関係性を見せるサービスもあるとよい。
- Web NDL Authorities の NDLSH は米国議会図書館件名標目表 (LCSH) とリンクしており、LCSH はドイツやイタリアで作成された件名データとリンクしている。このため、LCSH を媒介として、海外資料を日本語で検索できるようにしたい。日本語のキーワードを入れると入力候補として NDLSH が提示され、そこから検索キーワードと言語を選ぶとドイツ語やイタリア語のそのテーマの資料が検索できる、というツールがあるとよい。発展して、アラビア語のように日本人になじみのない言語で書かれた資料も検索できるようになってほしい。
- 作品のコンテキストのうち、描写されないものを提示できるとよい。たとえばドラマの時代背景にあるはずなのに描かれない同時代の風俗、習慣、先住民族などの情報を、そのドラマについて検索している人に提示できれば、文化的背景への理解が広がり深まるだけでなく、なぜ存在したはずのものに関する情報がドラマでは省かれているのか、といった、現代に生きる我々の意識の深層にあるものについて考えるきっかけにもなるだろう。

- ・小説や漫画については、Wikipedia に非常に詳細な情報があり作品のコンテキストとなる。そこでまず Wikipedia の登場人物の一覧と『架空人名辞典』を典拠データ化する。同時に、コミックマーケットで販売される二次創作作品の情報（原作、原作の作者、二次創作作品、その作者、登場人物など）も典拠データ化し、図書館のデータよりも詳細な分類である書店の C コードを介してそれぞれと書誌データをリンクする。これによって作品のコンテキスト情報に導かれれば、オリジナル作品や二次創作作品を読みたいという気持ちになり、作品購入につながることもできるだろう。と考えるとデータ化の予算を出版社から取れるかもしれない。
- ・その場所でしか体験できない、という見せ方もあるのではないか。例えば大河ドラマの旅行案内のように、本に含まれる地名を地理データ化して地図とマッピングする。さらに、その場所で起きた出来事をデータ化する。美術館で、あるコーナーに行くとヘッドフォンから解説が聞こえてくる、というのと同じ仕組みで、実際に現地に行くとその場所で起きた出来事に関する情報が得られたり関連するデジタル画像が見られたりするようなサービスにする。

## (2) デジタルコレクション

### (2)a 江戸時代人生ゲーム

- ・宗家文書のテキスト化が完了しているという前提で考える。
- ・宗家文書の膨大な記録を基に、宗家の藩士の人生を追体験できるゲームにしたい。例えば参勤交代の場合は、ルートの検索ができ、当時の経路・日数と現在の最短ルート・時間の比較が行え、更に観光情報などを教えてくれるようにする。
- ・人生ゲーム内でのイベントごとに、関連情報を教えてくれるようにする。
  - ・例えば、関連する人やモノ、事件、出世に関する情報や、どのような習慣を持っていたかなど、様々な情報の解説としてリサーチ・ナビや辞書データ（ジャパナレッジなど）、典拠データ、近代デジタルライブラリーのデータへの参照を提供する。典拠データからは、NDL-OPAC を検索できるようにし、関連資料の所蔵状況が分かるようにする。
  - ・土地に関する情報として、近代デジタルライブラリーやデジタルコレクションの画像、他機関が所蔵・提供している画像（NDL サーチ経由）などの古地図への参照を付ける。
  - ・モノに関する情報としては、例えば宗家文書にある「人參」の記述から、「漢方」へとつながり、さらに同音異義語の「官報」につなげて、デジタルコレクションの官報の画像や官報情報検索サービスに飛ばすことで思わぬ情報に出くわさせるというのも面白いかもしれない（シソーラス展開はよくあるが同音異義語への展開は聞かないので）。
- ・実現のために、機械可読なデータや、リンク付けするデータが必要だが、次のようなアイデアがあった。
  - ・テキスト化には、翻デジ (<http://lab.kn.ndl.go.jp/dhii/omk2/>) などを使ったクラウドでの共同テキスト校正、また、崩し字のテキスト化技術も適用できるのではないか。
  - ・NDL が持つ主題情報である「調べ案内、参考図書の内容情報（解題、目次、独自分類、掲

載人名等)、レファレンス事例、調べ方マニュアル、」も活用できるのではないか。

- ・特に、リサーチ・ナビの調べ方案内の宗家文書のページ

([http://rnavi.ndl.go.jp/research\\_guide/entry/theme-honbun-304002.php](http://rnavi.ndl.go.jp/research_guide/entry/theme-honbun-304002.php)) の情報は使えるのではないか。

- ・画像へのメタデータ付与では、近デジなどの一枚ごとの画像の URL を参照して、はてなブックマークのようなものでソーシャルタギングができないか (会議後のアイデア)
- ・成果にまとまる前段のアイデアとして、次のものがあつた。
  - ・特定資料の所在情報 (NDL+国内他機関+海外) が一度に分かると非常によい。(デジタル情報まで得られないのは仕方がないとして) / 特定人物についての資料がデジタル・紙の媒体を問わず、一度に分かるとよい。情報の串刺し。見せ方も工夫して。/ 思いもよらない発見があると面白い。
  - ・(具体的な実現性が思い浮かばず最終成果では発表しなかったが、) ゲーム内の登場人物にその土地の方言を話させるというのも面白い。

## (2)b カジュアル・デジタルコレクション

- ・デジタルコレクションをウェブ上でもっと身近なものにしたい。ソーシャルメディア (twitter、facebook、tumblr など) で簡単に使える、暇つぶしになるようなものにしたい。
- ・名言 bot ならぬ近デジ bot のようなものがあるとよい。例えば Twitter で自分がつぶやいている内容に応じて本を勝手に推薦してくれれば、恋愛についてツイートしたら、昔の恋愛に関する本が出てきて、昔も変わらなかったんだなあ…としみじみする、といったように、デジタルコレクションを普段の文脈につなげていけるとよい。
- ・デジタルコレクションがシェアして広がっていくことを支援したい。例えば本の一部の面白い部分の切り抜き、音源の一部といったように、jpeg や mp3 で切り出して提供できるようにする API があるとよい。これにより、コンテンツを簡単に貼り付けることができ、元をたどることもできるようになるため、マッシュアップの可能性が高まる。
- ・歴史的音源を流しながら、同じ時代の本やそれに関係する本を自動で表示するアプリがあるとよい。例えばその音源の楽譜や実演者などを、音源の終了後に表示する。また、雰囲気味わえるよう、関連しそうな挿絵や文をスクリーンセーバー的に表示したい。
- ・学習支援もしたい。何かを学びたい時の定番の参考書をオススメする機能があるとよい。参考書は版を変えて売られ続けているものが多いため、そういった本にまつわる小ネタを教えてくれる機能があれば、「この本はこんな昔からあつたんだ」という気持ちになり、勉強も捗るだろう。
- ・古写真と現代の様子を重ね合わせることもできれば学習支援になるだろう。

## (2)c 古地名陣取りゲーム

- ・既存の資料から古地名を収集し、現在の位置情報と結びつける作業を、多くの人が楽しみながら行なえる陣取りゲーム。
- ・年代ごとの古地名と、現在の地名・緯度経度との対応が簡単に分かると、過去の災害記録をハザ

ードマップ作成に参考にできるなど有益。

- NLD が所蔵・提供している資料には古地名が書かれた物が多いはずだが、網羅的に調べるのは人手がかかる。
- マイクロタスク化して陣取りゲームの皮を被せることで、プレイヤーが遊んでいると、知らずにこの作業に貢献してしまう環境をつくる。
- 具体的には、プレイヤーはいくつかの陣営に分けられ、自分が読んだ本等に古地名を発見したら、古地名・出典情報とセットでその時空的な位置座標(緯度経度と年代)を送信。ひとつの陣営で4つの点を取ると「時空四面体」を構成してその陣営の陣地とする。以降は点を追加することで陣地を拡大。
- 別の陣営の見つけた点を囲うような四面体を構成すれば、囲われた点を使った陣地は無効となり、陣地を奪える。
- NDL デジタルコレクションのページへのリンクに含めることで、情報の信頼性(検証可能性)を確保。発見された点の信頼性を失わせる資料の発見によっても、相手陣営の点を無効化できる。
- ゲームの結果構築された地名・位置・年代・出典のデータ組は誰でも入手できる情報源となる。
- 成果にまとまる前段のアイデアとして、次のものがあつた。
  - NDL デジタルコレクションから、表紙や挿絵などのデータを取得し、簡単に個人スクラップやTシャツにできるサービス。画像を「こんな感じ」という雰囲気を探せるUIや、権利処理が課題。
  - NDL デジタルコレクションのページ欠けや書誌情報の誤り等を共有できるサービス。特に別の図書館等の資料から欠けたページについて補足できたりすると良い。

### (3) レファレンスデータ

#### (3)a レファレンス協同データベースとの出会いをすてきに演出

- レファ協のデータには、回答プロセス・NDC(分類情報)・回答に用いた参考資料などが含まれており、レファレンスの信ぴょう性を高めている。利用者を質の高い情報へとナビゲートするためには、レファ協のように図書館員が生み出した質の高い情報の集合が役に立つと思われる。また、利用者がレファレンスの機能を知りレファレンスを利用することで、さらにナレッジが蓄積されより有益なデータベースとなりうる。
- しかし、実際には、一般の利用者がレファ協そのものを検索することはあまりなく、レファレンス情報と出会う機会は少ない。
- 現在は、NDL 検索でレファ協のデータ(レファレンス事例そのもの)を検索することができる。しかし検索結果は見づらく、レファレンス情報の本来の有用性を表現できていない。
- 現在は、レファ協の「参考資料」の項目からはNDL 検索の書誌データへのリンクが張られているデータもある。しかし、NDL 検索の書誌データからはレファ協のデータにリンクしていない。

- そこで、レファ協との出会いの場として、NDLサーチの書誌データからレファ協のデータへのリンクを実現させたい。
- レファ協のデータは「当該資料からどのような知識が得られるのか」の記録であり、当該資料の内容説明として捉えることもできる。このリンクによりNDLサーチの書誌データがより豊かなものとなる。
- また、現在のレファ協はキーワードでの検索がメインとなっているが、利用者が有用なレファレンス情報を探す方法としては改善の余地がある、という意見があった。「利用者にとって有用なレファレンス」の判定やそれに基づいたランキングができると、利用者が求めるレファレンス事例へのナビゲーションがしやすくなるのではないか、というアイデアも出た。

### (3)b 感情情報で本をレコメンド

- 抽象的な情報から具体化された情報を導くレファレンスサービスの特長を生かし、Twitter等と連携して気分などの情報を具体化して本をレコメンドするシステム。
- Twitter等でつぶやいた気分をタグに変換する。例えば「眠い」という気分を「眠」というタグにする。ブックログ、Yahoo!知恵袋、レファ協などの事例に同様のタグを自動付与しておき、マッチしたタグから本を紹介する。たとえば、タグ「眠」から、バクの本、枕の本、布団の本などを紹介する。
- 本だけでなく、音楽などにも拡張が可能。
- 「眠い」という気分を表す言葉ひとつにも、「眠いけれど眠れない」、「眠りたくないのに眠ってしまいそう」など、様々な状態がある。そこで、精度を高めるために、紹介した本が気分に合わせていたかどうかを評価してもらう仕組みを組み込む。