

# 図書館リソースのリワイヤリングによる セマンティックなサービスの創出 —フィードとデータマイニングの活用—

たなべ      みのる  
田邊      稔

(メディアセンター本部課長代理)

## 1 はじめに

Web2.0の構成要素であるロングテールやフォークソノミーといった概念は、とりわけビジネス社会に、顧客のニーズや多様性に目を向け、受け入れることの重要性やきっかけを与えた。その結果、近年、企業ではクレームデータやサービス後のフォローアンケートなど「顧客の声」を活用した顧客重視型サービスにウエイトが置かれるようになった。大学図書館でも、Library2.0やOPAC2.0といった言葉を使って、利用者指向型の図書館へと変貌すべく、利用者の声に真剣に耳を傾けるようになった。しかし、利用者の声をリアルタイムに収集・分析し、経営トップ層が俯瞰できる形で共有され、そこから新しいサービスを創出し、利用者満足度（以下、CS）向上に直結できている図書館はどのくらいあるか。

慶應義塾大学メディアセンター（以下、慶應MC）では、昨年度に本格的な利用者調査の取り組みを始めた。本稿では、膨大な情報から利用者の心を読み解き、図書館内外のリソースを効果的に組み替え（リワイヤリング）、クールでセマンティックなサービスを創出するための、様々な視点や方法を模索する。

## 2 利用者ニーズ分析

慶應MC利用者調査ワーキンググループにより2007年7月に実施されたフォーカス・グループ・インタビュー調査&分析の結果、現状サービスに対するニーズが、(1) ストレスのない手続き、(2) 一目でわかるナビゲーション、(3) 安心・快適さの提供、(4) 利用者の文脈の中での提供、(5) 利用者との信頼関係づくり、の5つに大別された。ここでは、「(4) 利用者の文脈の中での提供」に注目する。掘り下げると、「授業内容と課題にマッチしたもの」、「必要な時にタイミングよく提供されるもの」などが要望されていることがわかる。これらが次の学習支援サービ

スのヒントとなる。学習行動に合わせたデータ提供とは、例えば「授業の年間カリキュラムに沿って入門的な情報から専門の情報へとレベルアップする」とか「テストやレポート期に課題やテーマに即した関連図書やネット情報をサマリー提供する」などのしくみが想定される。これを実現するには、図書館と資料、資料と利用者、資料と著者、著者と利用者、といった関係性を見出し、最適なサービスを提供する必要がある。いわば、コレクション（コンテンツ）からサービス（コンテキスト）へのパラダイムシフトであり、図書館員の「編集力」が問われる。それには、まず図書館に蓄積された膨大な事例やノウハウを生かした分析パターン・ルールを生成する。これを「セマンティック・フィルター（文脈を意識した高度なふるい）」として分析システムに設定し、ノイズの少ない洗練された情報を導き出せるようにする。編集力とは、「図書館員の匠の技」ともいえる。

実際、数名の卒業生に個別ヒアリングしたところ、「大学生活における、ゆりかご（入学時）から墓場（卒業後）までの情報をタイムラインでスパッと見通したい」という意見が出された。この類の欲求は、図書館という枠の中で情報収集するだけでは到底満たせるものではない。部署内外を縦横無尽に飛び回り、利用者が「今」必要としている情報をかき集め、利用者が使いやすい形に最適化（編集）し、コンパクトに提供できるしくみが必要だ。ここで、RSS (Really Simple Syndication) フィードと呼ばれるオープンなXML (Extensible Markup Language) とそれを効果的にハンドリングするしくみが効果を発揮する。

## 3 RSS フィードの特性と期待

RSS フィード（以下、フィード）について、ここで詳細には触れないが、これまでニュースサイトやブログの更新情報として普及してきたフィードは、

今や「広告メディア」や「マーケティングツール」としての役割を担う。フィードは、セマンティックWebの架け橋と期待の高い「マイクロフォーマット」というコンパクトなXMLで記述される。フィードは一般的に裏側で自動生成されるため、ユーザーがフォーマットを意識することは少ない。さらに、フィードアグリゲータやSaaS（Software as a Service）型のRSSリーダーの出現により、フィードは、より一層ブラックボックス化している。

Web2.0のバイブル「Web2.0 BOOK」の著者で株式会社Modiphi代表の小川浩氏は、フィードを「インターネット上を駆け巡る血液」と定義し、「ウェブサイトは星、RSSフィードは光」と表現する。従来フィードの役割は、ウェブサイトに来てもらうための「釣り餌」に過ぎなかったが、タイトル以外にも画像付きの本文などが生成され、RSSリーダーで受信できるようになると、フィードだけで事足りるため、わざわざウェブサイトを訪問する必要はない。

図書館でも、フィードの特性を生かし、ニュースや新着資料の配信だけでなく、フィード情報を使った検索・収集・発信・共有ができるフロントエンドのコミュニティ・プラットフォームを提供し、そこで語られる話題や論文の評価、書評レビュー等を効果的に配信し、利用者からリアルタイムにフィードバックを得るといったモデルも考えられる。利用者とのコミュニケーションが活性化すれば、膨大なテキスト情報が集まり、それらを図書館が保有する情報と組み合わせて効果的にマイニングすることで、次のサービス戦略も立てやすくなる。

#### 4 データマイニングの必要性

図書館では、従来から様々な統計手法によるデータの収集や分析が行われてきた。しかし、ユーザー数や質問事項に限られたアンケートデータや年次統計などであれば、腰を据えてデータを集計すればよいが、利用者の声をリアルタイムに集め、分析し、即座にサービス改善につなげて行くためには、データマイニング（以下、DM）が必要となる。

いうまでもなく、インターネット上の情報は玉石混交である。企業のマーケティングであれ、個人的な情報収集であれ、膨大なデータから「石（ノイズ）」を排除し、「玉」のみ拾う方法やしゅみ（DM）を模索している。Googleの「ページランク」はDMの好例

である。Googleのおかげで、有用な情報が上位にランキングされるため、1ページ目か2ページ目を見れば大半事足りる状況が実現された。但し、必ずしも万人にとって有用なランク付けとは言えない。

ネットワークの高速化やインターネットテクノロジーの進化は目覚しく、情報収集経路は多岐に渡り、膨大なデータを蓄積するハードコストや、データ処理コストも低格化してきた。そのため、利用者からせつかく「声」情報を集めても、使うかどうかわからないデータや使い終わったデータを即座に捨てる必要がなくなった。数年間アーカイブしておき、経年変化や推移を追いかけることもできる。また、コンピュータの性能も向上し、PCクラスのマシンでも十分に分析処理できるレベルに達している。

DMとは、蓄積された膨大な量のデータから、意味のあるパターンやルールを発見すること。中でも、テキストマイニングは、大量のテキスト（文字）データを対象とし、文字情報から新たな事実や傾向を発見することを支援する技術である。いずれも、コンピュータが意味のあるパターンやルールを勝手に解釈して自動的に見つけてくれるわけではない。分析者である図書館員が、その蓄積されたデータから何を見出したいか考えた上で分析パターンやルール（セマンティック・フィルター）を設定する必要がある。このフィルターは一定ではなく、様々な切り口を用意しておき、その都度設定を変更する。これにより、有用な情報が得られる。DMにおいて重要なことは、膨大な情報を効率良くデータベースに蓄積することと、蓄積された膨大な量のデータをいかに高速に処理するかということ。情報を蓄積するにあたり、必要な情報は全て記録しておくことが分析を行うためにも重要である。

#### 5 データマイニングの参考事例

ここで、独自のデータマイニング手法を使い、利用者の声を効果的に分析・可視化し、企業などに提供してサービス効果分析や新商品の企画支援などを実践しているホットなベンチャーを3社紹介する。

##### (1) 関心空間

「関心空間」はユーザーがホスト役をつとめる「空間」をベースとしたコミュニティサイト。ひとりひとりの関心事がキーワードとして登録され、他のユーザーの“共感”を得て自由につながりあい、話題の

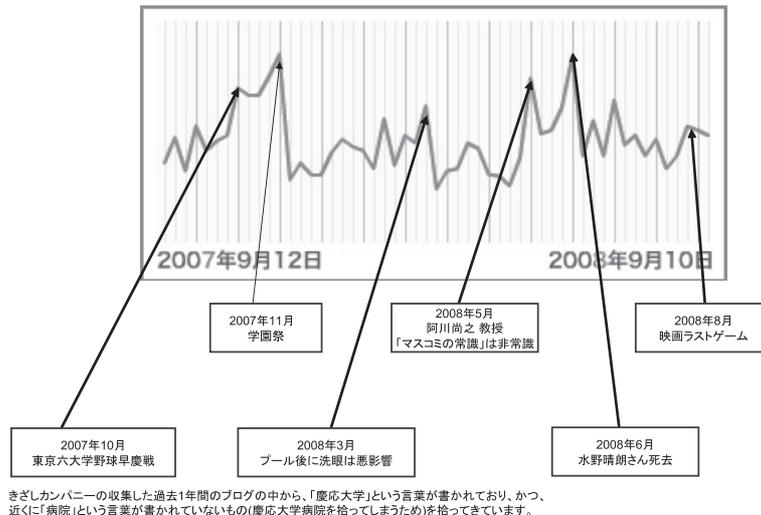


図1. kizasi.jpにて「慶応大学—(病院以外)」の条件で調査

広がりや意外な発見の連続を楽しむことができる。コンテンツ価値と異なり、コンテキストの価値は相対的で、瞬時に変化しやすい。そこで個別適合性の高いコンテキストとコンテンツを抽出するためのデータマイニングのアプローチを行っている。

### (2) きざしカンパニー

きざしカンパニーは、CGM/UGM<sup>1)</sup>領域における「話題」を切り口にした検索エンジン事業の開発を行っている。「kizasi検索エンジン」は、ブログは多数の人が世の中の変化に応じて、一斉に情報発信を行うという即時性を特徴とする。この検索エンジンの基本コンセプトは、ブログを「時間」と「人」の2軸に沿って言葉を解析すること。具体的には、話題を質と量の両面からとらえ、その変化をしめすことが可能。話題の量がどのように時間推移していくのか、また、コトバが表現する話題内容の推移を、共起語の推移から捉える。代表的なサービスに「kizasi.jp」がある(図1のサンプルを参照)。

### (3) グランドデザイン&カンパニー

グランドデザイン&カンパニーは、モバイルを中心としたソリューションの企画・開発・運営を行っている。独自のモバイル解析ツールを元に、ユーザーの行動分析等を行い、サイト内でのコンバージョンアップ策などを編出する。要素としては、ユーザー属性に基づく行動分析や、アクセス解析はもとより「配信したメールの内容に対する反響反応」をマイニングしていく手法など、複合的な分析・解析を行う。これらの解析ツールと多数のサイト運営を通じて蓄

積したデータベースを元に、有効なアウトプットを導いている。代表的なサービスに「コスメテスター」がある。

## 6 リワイアリングとは?

「今後社会のありとあらゆる企業のシステム、経済のシステムが1つにつながっていく」とマイクロソフトのビルゲイツ氏は主張する。例えば、POSシステムやGPS、Suica定期券、Fericaなどを駆使すれば、どの地域でどの商品が売れているのかがリアルタイムで把握でき、ユーザーの位置情報を集めることができ、電子マネーでどんな商品を買ったかという記録も収集可能である。あらゆるデータをコンピュータが自動的に分析し、消費者一人一人に合った広告や情報を配信していく時代になる。「広告という概念を超えたコミュニケーションがあらゆる電子機器を通じて、企業と消費者との間で始まろうとしている。そうしたコミュニケーションを通じて消費者の信頼や愛着を勝ち取った企業が、21世紀の成長企業となるであろう。」と湯川鶴章氏(時事通信社)は予測する<sup>2)</sup>。広告から販売まで1つの配線につながっていた時代から、無数の機器が複雑につながる時代へ。ビルゲイツ氏は「今、経済のリワイアリング(配線組み換え)が行われている」と表現している。

また、ダンカン・ワッツらが提唱した「スモールワールド・ネットワーク<sup>3)</sup>」は、一見、全然つながっていない人同士でも、大体5人を介すると必ずつな

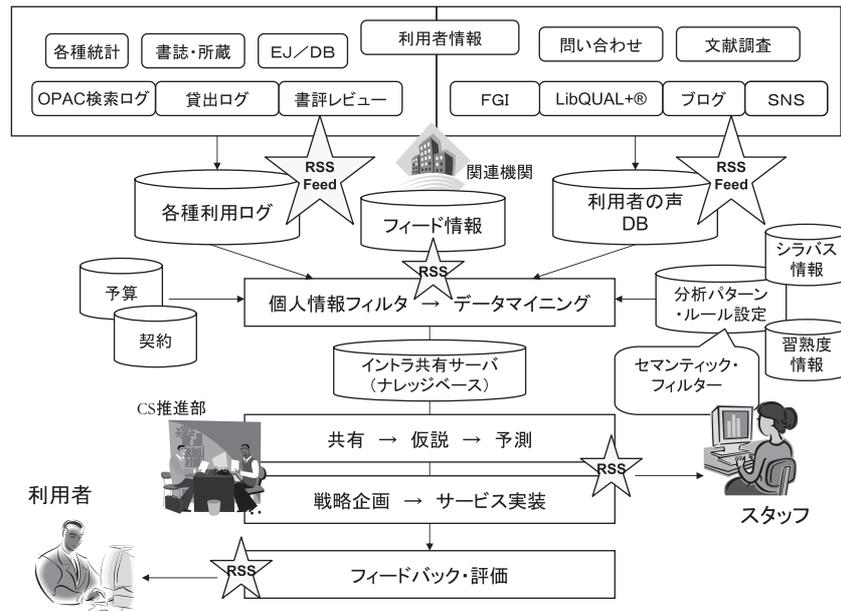


図2. 図書館におけるデータマイニングのしくみ

がっているという理論であり、ランダムなネットワークに、人間の特徴である同類志向（つまり好きなもの同士が結びつく傾向）を追加し、つなぎ直し（リワイヤリング）をすることによって、世界を小さくすることが可能になると説いている。

西口敏宏氏（一橋大学教授）の『遠距離交際と近所づきあい 成功する組織ネットワーク戦略<sup>4)</sup>』によると、「遠距離交際」とは「偶然に起こるリワイヤリング（情報伝達経路のつなぎ直し）によって、普段は地理的な距離が離れているために入りにくい情報がダイレクトに入ってくること」、また、「近所づきあい」とは、「自分を取り巻く近隣の人たちとの日々の規則正しい密接な関係のこと」をさしている。このバランスを取ることが、組織の成功の鍵となる。

## 7 図書館を支える3つのマネジメント リソース

図書館におけるデータマイニングの1モデルを図示（図2）し、「利用者の声マネジメント」の実現に必要なリソースを3つに分けて概説する。

### (1) データ

図書館が日々管理している情報には、書誌情報（MARC）・所蔵情報・予算契約情報・利用者情報・貸出履歴・OPAC検索ログ等がある。これだけを取り上げても、十分マイニングができる。「この本

を借りた人はこんな本も借りています」、「この雑誌は高いお金を出して購入した割に利用されていない」、「粘着質な利用者が発するコメントには共通性が多い」など面白い分析結果が得られると予測される。これらのデータと連携する形で、問い合わせ、要望・クレーム、サービス評価、書評レビュー、口コミなど利用者の生の声をデータベースに蓄積する。また、書店、代理店、出版社、データベースベンダーなど図書館関連機関の情報やインターネット上の関連情報をRSSリーダー経由で日々収集しておく。

### (2) システム

上記それぞれ蓄積されたデータ群をコンピュータ処理に掛けてキーワードや関心テーマ、関連文献等を抽出する訳だが、より効果的に分析するためには、前述した「セマンティック・フィルター」を設定する。ここで図書館員の経験と勘が必要となる。何を見出したいか考え、分析パターンやルールを決める。授業のカリキュラム（シラバス情報）と受講者情報を登録しておき、情報リテラシーテストなどを通じて受講者の習熟度をフィルター情報に加味すれば、レベルに応じた情報の提供が可能となる。また、授業後に実施するフィードバックアンケートの結果と連動して「受講者がわかりにくい所ランキング」等を教員側に提供することで、次の授業や教材

作成に役立つこともできる。スモールスタートで構わないので、目に見える形で現場や経営トップへプロトタイプを示すことが重要だ。

### (3) 組織

慶應 MC では、利用者調査 WG というキャンパス横断の組織体を設けているが、利用者調査専門の担当ではなく、メンバーは全員兼任(いわば片手間)である。可能であれば「CS推進部」や「戦略企画室」といった専門部署を設け、「情報流通のハブ」としての役割を担い、カウンタースタッフや IT 部門、経営トップそれぞれに対し、有用な情報をリアルタイムに提供することで活性化を計るといった体制を確保することが理想であろう。また、分析者としてのセンスを養うために人材の育成が急務である。そのためにも、知識とスキルの蓄積・継承が期待される。理想的な「セマンティック・フィルター」の実現を目指すためにも、全社の意識を「利用者の声」に傾ける努力が必要となる。

## 8 おわりに

図書館などの情報サービス機関においては「利用者の声をマネジメントする」ことがこれからの使命の1つであり、勝ち残り戦略になるといっても過言ではない。利用者の「今」を知り、「今」にリーチするためには、利用者の多様性を受け入れるしくみや体制が必要となる。図書館と利用者の叡智を結集し、効果的にデータマイニングを行い、有用なサービスをマッシュアップしていくことがCSに直結する。それは、利用者に迎合することでは決してない。あくまで、利用者の声を積極的に活用して、見通しのよいサービスを提供し、利用者と良い関係を築くことが目的である。利用者自身が、膨大な情報から「玉」を得て、オリジナリティを創造し発信できる環境を整えることができれば社会貢献にもつながる。

これらのしくみや体制づくりは、図書館だけの力だけでは実現し得ない。できるだけ情報をオープンにして周囲を巻き込み、外部から斬新な発想や優秀なリソースを取り入れて、内部リソースとうまくリワイヤリングさせ、サービスの活性化を図る必要がある。「夢は宣言すると叶う」というが、図書館が

「脱！図書館宣言」をするとき、夢に一步近づくのではないだろうか。

### 謝辞

原稿執筆にあたっては、株式会社関心空間・代表取締役社長の前田邦宏氏、株式会社さざしカンパニー・代表取締役専務テクノロジー統括の稲垣陽一氏、グランドデザイン&カンパニー株式会社・代表取締役社長の小川和也氏に御協力を頂いた。この場を借りて感謝を申し上げます。

### 参考文献・注

- 1) CGM/UGM (Consumer Generated Media/User Generated Media)とは、インターネットなどを活用して消費者やユーザーが内容を生成していくメディア。個人の情報発信をデータベース化、メディア化した Web サイトで、Web 2.0 的なもののひとつとされる。クチコミサイト、Q&A コミュニティ、ソーシャルネットワーキングサービス (SNS)、ブログ、COI (Community Of Interest) サイトなどがこれにあたる。
  - 2) 湯川鶴章の IT 潮流. “経済リワイヤリング—ボツにした未完成原稿 vol.0”. (オンライン), 入手先<[http://it.blog-jiji.com/0001/2008/04/vol0\\_86e3.html#more](http://it.blog-jiji.com/0001/2008/04/vol0_86e3.html#more)>, (参照 2008-07-30).
  - 3) ダンカンワッツ. スモールワールド・ネットワーク—世界を知るための新科学的思考法. 東京, 阪急コミュニケーションズ, 2004. 389p.
  - 4) 西口敏宏. 遠距離交際と近所づきあい 成功する組織ネットワーク戦略. 東京, NTT 出版, 2007. 486p.
- ・三室克哉, 鈴木賢治, 神田晴彦. 顧客の声マネジメント: テキストマイニングで本音を「見る」. 東京, オーム社, 2007, 232p.
  - ・岡嶋裕史. 数式を使わないデータマイニング入門: 隠れた法則を発見する. 東京, 光文社, 2006, 211p.
  - ・浅尾千夏子, 藤本優子. “利用者調査 WG 活動報告”. 第5回メディアセンター研究発表. (オンライン), 入手先<<http://www.lib.keio.ac.jp/kenshu/pdf/2007/3/riyoshawg.pdf>>, (参照 2008-05-09).