

デジタルライブラリにおける履歴・個人情報の保護及び利用

—新たな電子図書館サービス構築へ向けた個人情報保護モデル—

池田 大輔

九州大学附属図書館
daisuke@lib.kyushu-u.ac.jp

安東 奈穂子

九州大学大学院法学研究院
n-ando@law.kyushu-u.ac.jp

田中 省作*

九州大学情報基盤センター、高等研究機構研究戦略企画室
sho@srp.kyushu-u.ac.jp

概要

電子図書館サービスの多くは、従来からある図書館サービスを電子的に行うことにより、図書館員が行う作業の効率化や資料への効率的なアクセスを実現するものであり、新たなサービスを提供するものは少ない。筆者らの研究グループは新たな電子図書館サービス構築を目指し研究を進めており、その過程で個人情報を含んだ貸出履歴をサービスに使うことの是非が大きな問題点として浮上した。本稿では、図書館の役割や利用者の個人情報について考察し、新たな電子図書館のサービスによる利便性向上と個人情報保護を両立するモデルを提案する。

1 はじめに

従来提案されてきた電子図書館サービスは、省力化・効率化を目指したものと資料を電子化して電子図書館とするものの二つに分けられる。前者としては、古くは事務電算化や図書館システムの導入、最近ではバーコードやRFIDタグ、自動書庫などの導入がある。後者としては、電子ジャーナルや文献データベースに加え、各図書館が保有する貴重資料などの電子化がある。

これらの電子図書館サービスは、基本的に従来からある図書館サービスを形を変えて提供したり、図書館員の作業効率を上げるものと期待されるが、利用者にとって新たなサービスを提案するものではない。また、予算があればすべての図書館で実現可能であり、これらの電子図書館サービスによって各図書館の独自性を出すのは困難である。

例えば、九州大学附属図書館は、2003年2月から筑紫分館に実験的にRFIDシステムを導入し、図書館におけるRFID利用の実証的な研究を行っている[5]。全国の図書館の中でも比較的早い時期の導入であったため、広く注目を集め、現在でも多くの見学者が訪れる。しかし、RFIDシステム自体を九州大学附属図書館が独自に構築したわけではなく、予算があればどの図書館にも同様のシステムは導入可能である。実際、公共図書館を中心に全国的に同様のシステムが数多く導入されつつある。

筆者らは、新たなサービス提供を視野にいれた電子図書館モデルを構築するため、九州大学内の様々な分野の若手研究者によるプロジェクトを平成16年度より開始した。このプロジェクトは、RFIDやデータマイニングといったIT関連技術を用い、新たな電子図書館のサービスを構築し、Webや携帯による情報提供や資料へのアクセスを可能にすることを目的としている。

上記プロジェクトを進める中で、図書館にとって

*平成17年4月より立命館大学文学部所属。
sho@lt.ritsumei.ac.jp

は古くから存在する問題である利用者の貸出履歴の利用に関する問題が、新たな形で浮上してきた。また、「個人情報保護法」および、九州大学に適用される「独立行政法人等の保有する個人情報保護に関する法律」などを含めた個人情報保護関連 5 法（以下、まとめて個人情報保護法と呼ぶ。）が、平成 17 年 4 月 1 日より施行された。この法律によると、個人情報とは

生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）をいう。

と定義され、これをデータベース化したものが個人情報データベースである。

通常図書館が保有する利用者データベースは、氏名や ID 番号、所属や住所、連絡先などから構成されるため、個人情報データベースである。また、図書館においては貸出履歴データベースも重要である。貸出履歴は、ID 番号とその利用者が借りた本の対応表である。通常は ID 番号だけでは個人を特定できないが、利用者データベースと照合し個人の特定が可能であることから、貸出履歴も個人情報と考えられる。

本稿では、電子図書館において考えられるサービスモデルを提示し、これらのサービスと利用者の個人情報の関係について考察する。さらに、九州大学システム LSI 研究センターが提案する PID (Personal ID) システムを用いて利用者の利便性の向上と個人情報を保護とを両立するモデルを提案する。

2 新たな電子図書館サービス

筆者らは“ユビキタス社会における電子図書館のソフト面高度化に関する研究”という研究テーマで九州大学教育研究プログラム・研究拠点形成プロジェクト（略して P&P と呼ばれる。）による助成（D タイプ）を受けて、電子図書館関連のプロジェ

クトを平成 16 年度から始めた。D タイプのプロジェクトは、若手研究者の交流も視野に入れたプロジェクトであるため、構成メンバーは若手中心かつ研究テーマが文理融合であることが条件となっている。筆者らのプロジェクトには附属図書館、法学研究院、情報基盤センター、経済学研究院、システム情報学研究院、比較社会文化研究院の研究者が参加している。

このプロジェクトは、RFID やデータマイニングといった IT 関連技術を用い、新たな電子図書館のサービスを構築し、Web や携帯による情報提供や資料へのアクセスを可能にすることを目的としている。具体的には、図書館が保有する資料とこれに対する潜在的な利用者を IT 関連技術を用いて結びつけることが目的であり、個々のアプリケーションとしては図書推薦システム、仮想書架、読書録などを考えている。

図書推薦システム

図書推薦システムは、ある本と様々な意味で関係のある本を同時に提示するシステムである。考えられる関係としては、同じ著者の本や同じ分野に分類される本といった意味的に近いと思われる本だけでなく、同じサイズであったり同じ日に購入されたり、といった意味的に遠いと思われるものも定義可能である。さらに、単に機械的に関係を定義するだけでなく、図書館職員による手動の関係づけも行うことが可能である。

仮想書架

仮想書架は、自動書庫の欠点を克服する対応策として考えたものである。自動書庫においては、資料は数十冊単位でコンテナに収められ、どの資料がどのコンテナに入っているかが機械的に管理される^{*1}。資料 ID などを用いて出庫命令を自動書庫に対して発行すると、クレーンのようなものが目的のコンテナを自動的に出口まで移動させる。

自動書庫の利点は、収納効率の良さにある。一方、完全な閉架であるため、書架の間を散策しつ

^{*1} ただし、コンテナ内における資料の順番までは分からない。



図1 仮想書架のサンプルイメージ

つ、様々な本を眺めたり、思いがけない本との出会いが不可能になる。そこで、図1のように、背表紙の画像を用いて実際の書架の配列を仮想書架としてコンピュータ上で再現し、本との出会いを仮想的に演出する。このようなインターフェイスは既に提案されている [4] が、実際の背表紙データを用いるところに特徴がある。

単に実際の書架と同じ配列だけでなく、ポータルサービスと組みあわせれば、利用者が見たい書架だけ表示したり、物理的な配置は遠くになっているが仮想書架では近くに配置したりでき、まさに My Library を実現できるものと期待できる。

さらに、図書推薦システムや通常の検索結果のインターフェイスとして用いることも可能である。

読書録

このサービスは、個人認証を含んだポータルサービス^{*2}の一部として機能させることを想定している。ポータルサービスの典型的なサービスとして、現在借りている資料の一覧を表示させる機能がある。これにより、利用者はいつまでにどの本を返せばよいか把握できる。

このサービスを一歩進めて読書録サービスとすることができれば、より便利に利用してもらえると期待できる。読書録サービスとは、本にコメントなどや点数(レーティング)などをつけた上で、書誌情報とこれらの付加情報を保存しておき、後で検索

や点数により自分が借りた本を表示するサービスである。

研究者マッチング

このサービスは、図書館というより大学としての政策の一つと考えるべきである。従来大学における研究者は、個々に独立して研究を行ってきた。一方、COE などの大型プロジェクトやその他の外部資金などを獲得していくには、大学が戦略的に研究を推進する体制を整備する必要がある^{*3}。研究者マッチングは、ある研究テーマに関連する研究者をピックアップする作業であり、上述した観点では非常に重要になる。

九州大学では、教員の活動をデータベース化しておき、このデータベースを有効に利用することで研究者マッチングを行おうとしている [6, 10]。一方、図書館における本や論文の貸出履歴や利用履歴から研究者マッチングも行える可能性がある。

3 電子図書館と個人情報

本節では、2 節で説明した各種サービスを、図書館が保有する個人情報や貸出履歴を用いて実現する方法とその問題点について述べる。これらのサービスは、これらの情報を用いずとも実現可能だと考えられるが、ここでは考察のため用いることを想定する。

^{*2} 九州大学附属図書館では、“MyLibrary” という名称でこのようなサービスを提供している。

^{*3} 筆者の一人である田中の所属していた九州大学高等研究機構研究戦略企画室は、このような大学全体における研究やその方向性の企画立案を行う部署である。

まず、図書館が保有する個人情報としては利用者データベースがある。通常、氏名や ID 番号、所属や住所、連絡先などから構成される。ID 番号は、単なる連番であったり、学校図書館であれば学籍番号などが利用される。

図書館が保有する個人情報は、基本的に利用者データベースのみであるが、重要な機密データとして貸出履歴に関するデータベースが存在する。従来から図書館は、利用者の利用傾向について非常に注意深く取り扱ってきた。そのため、このデータベースも個人情報と同様に慎重に扱うべきである。また、個人情報保護法では、図書館に利用者 ID と個人情報の対応表の双方があり、貸出履歴の利用者 ID が誰なのか特定が容易なため、貸出履歴も個人情報と考えることができる。

新たなサービスとして、読書録、研究者マッチング、図書推薦システムの 3 つを考える。以下に、各サービスと個人情報や貸出履歴との関係について述べる。

読書録：人と本の対応づけ

読書録は、誰が何を read したかの記録に、本ごとのコメントや感想をつけるものである。つまり、人と本の対応づけを行うサービスである。これを実現するには、どの利用者が何を借りたかという貸出履歴を残す必要がある。これ自体が潜在的な問題点である。

しかし、利用者が用いる貸出履歴は利用者本人の情報であり、利用者と貸出履歴が完全に対応している。そのため、サービス開始時に、利用者に対しサービスの概要と潜在的な問題点を十分説明した上で、貸出履歴を残すことに同意した利用者にかぎって貸出履歴を残してサービスを提供することが自然である。

その意味では、貸出履歴の使われ方が利用者から見て明白であり、なおかつ履歴を残したくない利用者はその選択肢が残されている。貸出履歴を活用するのはその情報を生成した利用者本人であるため、このような選択肢が可能になる。

研究者マッチング：人と人の対応づけ

このサービスは、特定の研究テーマに対し、主に学内の研究者を集め、研究グループを組織するものである。マッチさせるべき研究者は、例えば、論文を引用していたり、共著の関係にあったり、同じ専攻に所属していたりすることが自然であるが、ここではマッチさせるべき研究者を、同じ本を多く借りた研究者と定義する。サービスの性質上、サービスの提供結果（この場合は組織された研究グループ）に個人情報が陽に含まれている。

このサービスでは、貸出履歴の利用者は、図書館の利用者ではなく大学全体の研究戦略を担当する部署である。これは読書録における貸出履歴の利用と違う点である。また、少数の利用者の貸出履歴では十分な結果が得られない。つまり、貸出履歴は多ければ多いほど、質の高いマッチングが可能になる。

図書推薦システム：本と本の対応づけ

本と本の関係を、ある一人の利用者が一定期間に借りたという行為により定義する。一定期間を例えば短く 1 日にするとある利用者がその日に同時に借りた本だけに関係がつけられる。そのため、期間は長いほうが多くの関係づけが可能になる。また、研究者マッチングと同様、多数の利用者の貸出履歴があるほうが、精度の高い推薦が可能になる。

このサービスにおいても、研究者マッチングと同様、貸出履歴を直接の利用するのは図書館であり、図書館の利用者ではないという点が問題である。しかし、図書推薦システムと研究者マッチングには違いも存在する。研究者マッチングは、サービスの性質上、貸出履歴とその利用者の両方の情報が必要であり、情報の最終的な解析結果には利用者（この場合は研究者）が陽に現われる。そのため、図書館利用者の事前の同意が必要になる。一方、図書推薦システムは、陽に図書館利用者の利用傾向を必要としない。つまり、どこの誰が借りたかという個人の特定につながる情報は不要で、誰でも構わないから同じ本を借りたかどうかの情報が必要になる。

そのため、貸出履歴を変換して得られる、利用者 ID を含まない本と本のみの対応表から図書推薦シ

システムは構築可能である。そのため、図書推薦システムの出力から利用者やその利用者の読書傾向が漏れることもない。その意味では、3つのサービスの中でもっとも安全なサービスであると言える。しかし、通常は変換前の貸出履歴も、利用者IDと個人情報の対応表も図書館が同時に管理するため、誰が借りたのかが追跡可能である^{*4}。

つまり、貸出履歴を用いた図書推薦システムの実現は、他の2つのサービスと異なり、利用者の事前の同意なしにサービスに必要なデータベースの利用が可能であることを意味しているようにも思える。しかし、個人を特定することができなくても、図書推薦システムに貸出履歴を用いることは、図書館が貸出履歴を収集している本来の目的から外れるかもしれない。そこで、次に図書館の目的と貸出履歴を収集する必要性について考察する。

4 (電子)図書館の目的

まず、電子図書館ではない通常の図書館の目的から考えてみる。図書館は、本を体系的に集めて、分類し、提供している。これにより、目的の本が明確になっていなくとも、分類をもとに利用者が目的の本にアクセスしやすくなる。また、目的の本が明確な場合でも、分類・関係づけられているため、その本の近辺に参考になるような本が他にも見つかる可能性がある。言い方を変えると、図書館は、分類や配架の工夫によりアクセスの容易さを実現していると言える。

一方、筆者らの提案する電子図書館は、IT技術を用いて関連づけ・分類・付加価値を大規模にかつ多様につけて提供する新しいタイプの図書館である。その意味では、従来の資料の単なる電子化や、RFIDや自動書庫による効率化を目指した電子図書館は、その付加価値が狭く限定されている。

つまり、電子図書館であれ、通常の図書館であれ、本と本に関係づけることにより、アクセスのしやすさを実現しており、電子図書館の場合は、さら

に、IT技術により大規模にできる。ここまでの議論では、特に個人情報が必要なく、本と本の間に関連のみが必要であった。

さて、図書館は資料の貸出を行っている。これは、一時的に資料を利用者に預ける行為であり、図書館は資料を確実に返却してもらわなければならない。そのため、図書館は利用者の個人情報が必要になり、結果的に利用者の利用傾向などが図書館による収集可能になってしまうという現状がある。

利用者の利用傾向の扱いについては、図書館は古くから慎重であった[7, 8]。例えば、日本図書館協会総会により採択された「図書館の自由に関する宣言」には「第3 図書館は利用者の秘密を守る」として、利用者の利用傾向はプライバシーに属する問題であり、その事実を外部に漏らさないと明記している[2]。さらに、より具体的に電子的に蓄積されていく利用者の利用傾向の利用を禁止するものとして、例えば、日本図書館協会総会により採択された基準がある[3]。また、この方針を具体的な運営上の準則として導入した例もある[8, p. 180]。つまり、図書館は本と人を対応づけてはならない。

その意味では、読書録は人と本の関連づけ、研究者マッチングは人と本の関係を用いた人と人の関連づけであり、本来の図書館の目的の範囲はずれるサービスである。そのため、図書館利用者に対する事前の同意は不可欠である。

一方、図書推薦システムは、本と本の関連をつけるものであり、図書館本来の目的に合致する。この本と本の間関係を用いれば、図書推薦システムだけでなく、配架の工夫にも援用できる。従来は物理的な配架を頻繁に変えることはコストが高く、このような配架の工夫の必要性はあまりないと思われてきた。しかし、電子図書館では、上述した仮想書架のように自由に配架を変えることが可能である。そのため、本と本の間関係を積極的に利用する必然性は十分にある。

本と本の間関係を作るために従来よく行われている方法は、ある本に引用されている本を、元の本に関連づけるものがある。その他に、図書館職員や教員による本の推薦などが考えられる。また、本稿の

^{*4} もちろん、図書館が追跡すると言っているわけではない。しかし、利用者側からすれば可能性が残っていることは不安になるであろう。

考察の対象である貸出履歴を使った場合は、貸出履歴と個人情報を図書館が同時に保有していることから、上述したように、一時的にせよ、個人の利用傾向が追跡可能な状態になってしまうことが問題となる。

5 PID を用いた図書推薦システムの実現

本節では貸出履歴は用いるが、誰がどの本を借りたかという意味では個人情報を含まない形で図書推薦システムを実現するためのモデルを提案する。まず、比較のために貸出履歴を残さないモデルと残すモデルを説明する。

貸出履歴を残すモデル

このモデルでは、適当な期限を切って貸出履歴をデータベースに残しておく。その後、貸出履歴をデータベースから削除する際に、同じ利用者が借りた本の情報から、本と本の対応表に変換したデータベースを作成する。新たに作成したデータベースは、個人情報を含まないが、これを作成する際には、下のモデルより長い期間貸出履歴を保持する必要がある。

貸出履歴には、利用者 ID が含まれており、その意味で個人の利用傾向が分かる形になっている。つまり、前節の最後で指摘した問題点は残る。

貸出履歴を残さないモデル

このモデルでは、本が返却されたら貸したという事実をデータベースから削除する。そのため、貸出履歴を最低限しか持たない。しかし、利用者の個人情報と履歴が一時的にせよ、同時に図書館にあるという意味では、こちらのモデルも前節の最後で指摘した問題点は残る。

PID モデル

上述した 2 つのモデルは、どちらも一時的にせよ図書館に貸出履歴と個人情報が同時に存在する。そのため、個人の利用傾向を追跡可能な状態にある。ここで、個人情報を図書館が持たず、図書館が個人の利用傾向を追跡できないモデルを提案する。

PID とは Personal ID の略であり、九州大学システム LSI 研究センターが提案する新たな個人認証の仕組みである [1, 9]。PID 設計の大きな方針とし

て、個人のプライバシーや権利の保護が第一義的に挙げられている。

PID を用いた個人認証では、発行元、ユーザ、サービス提供者の 3 種類の参加者が定義される。大学図書館の場合は、発行元は大学学長に、ユーザは図書館の利用者になり、図書館はサービス提供者と位置づけられる。

まず、発行元がユーザの身元と信用を確認する。その上で、各ユーザに個人識別子 (PID) として十分に長いビット列を与える。このビット列が IC カードなどに記憶されて各ユーザに渡される。この枠組みでは、発行元のみが個人情報と PID の対応表を持ち、その管理に責任を負うが、サービス提供者は個人情報を持つ必要がない。

次に、図書館などのサービス提供者は、発行元にサービスを行う旨申請する。申請が認められれば発行元は各ユーザの PID の一部 (サブ PID と呼ぶ) をサービス提供者に渡す。この時、ユーザにはどの部分をどのサービス提供者に渡したかが通知される。サービス提供者は、このサブ PID を用いて認証を行う。

PID の利点として、まず、仮に利用者のあるサブ PID が漏洩しても、別のサービスを利用するには別のサブ PID を用いるので、被害が限定される。例えば、ある大学において図書館、情報基盤センター、成績証明を発行する部局がサービス提供者になったとしよう。この時、図書館のサブ PID が漏洩したとしても、他のサービスの利用には影響を与えない。

また、PID の仕組みでは、図書館はそもそも個人情報を持たずにすむ。PID を用いることで、図書推薦システムを実現する一つ的手段として、個人情報を含まない形式で貸出履歴を用いることが可能になる。実際、サブ PID は、学籍番号などとは異なり、所属や入学年度の情報を含まないため、このビット列からだけでは、図書館が個人を特定することは不可能である。

しかし、PID を用いた図書館における個人認証でも、督促の問題が存在する。なかなか本を返却しない利用者に督促を行う時には、連絡先などの個人情報

報が必要である。PID の枠組みでは、一度個人情報
を図書館が入手してしまえば、その後の利用傾向に
についても図書館が把握できてしまうことになるが、
サブ PID の再発行（発行元に別のサブ PID を割り
あて直してもらう）により、図書館が知る貸出履
歴と個人の対応関係を限定的にすることは可能で
ある。

限定できるとはいえ、一定期間貸出履歴と個人情
報が同時に図書館に存在することになることに
変わりはない。つまり、期限内に本を返却する利用
者の個人情報は図書館に知られないが、期限を
超過してもなかなか返却しない利用者は自らの
個人情報を知られるリスクを負うことになる。
PID を用いたモデルを実現する場合は、この
点をよく周知する必要があるだろう。

6 まとめ

本稿では、筆者らが進めている P&P プロ
ジェクトを紹介し、このプロジェクトにおいて
遭遇した問題点、つまり、電子図書館にお
ける新たなサービスと個人情報の保護の
関係について考察した。さらに、PID
を用いた個人情報の保護をできるだけ
可能にする電子図書館サービスの
運用モデルを提案した。

読書録を除けば、本稿で考察した
サービスを実現する方法として、
貸出履歴は必ずしも必要ない。
しかし、従来から図書館は貸出
履歴を慎重に扱ってきたことと、
個人情報保護法が施行されたこと
を考えると、新たな電子図書館
モデルを構築するためには、
貸出履歴と電子図書館について
考察する必要があると考え、
本稿で取り上げた。

読書録と研究者マッチングに
貸出履歴を用いる場合、必ず
個人情報が介在し、サービス
提供後も個人情報が残ってい
る。そのため、これらのサー
ビスに貸出履歴を用いること
については、事前に利用者
からの同意が必要であること
は自明であろう。特に、研
究者マッチングについては、
図書館だけでなく、大学全
体における同意が必要であ
らう。

一方、図書推薦システムにつ
いては、個人情報は

サービス提供後には残らない。
さらに、このシステムに必要
な情報は本と本との関係であ
り、システムの出力も本と本
の関連づけであるため、この
システムの利用は図書館本来
の目的の範囲内と考えること
も可能かもしれない。さら
に、PID という個人認証の
枠組みを用いれば、この関
連づけに個人情報を
用いる必要もなくなる。その
ため、仮にこのようなサー
ビスを行う場合には、貸出
履歴^{*5}を用いること
についての利用者からの
事前の同意は不要に感じ
られるかもしれない。し
かし、現状ではこのよう
なサービスが図書館本来
の目的であるという
広い合意が得られて
いるわけではないこと
や、仮に個人が
特定されなくても
内面的精神活動
（例えば思想
など）の表れ
である履歴情
報を提供し
それが利用
される利用
者の心情に
配慮する
ならば、
貸出履歴
を用いる
目的を
限定し、
これを
利用者
に通知
し、同意
を得る
ことが
重要
である。

謝辞

本研究の一部は、九州大学教育研究プログラム・
研究拠点形成プロジェクト (P&P) の助成を受けて
行ったものです。

参考文献

- [1] 安浦. 九州大学全学 IC カード導入プロジェクト. http://www.slrc.kyushu-u.ac.jp/japanese/presentation/p_yasuura04_1.pdf, 9 2004.
- [2] 社団法人日本図書館協会. 図書館の自由に関する宣言. <http://www.jla.or.jp/ziyuu.htm>, 1954. (1979 年改訂).
- [3] 社団法人日本図書館協会. 貸出業務へのコンピュータ導入に伴う個人情報の保護に関する基準. <http://www.jla.or.jp/privacy/kasidasi.html>, 1984.
- [4] 神谷, 呂, 原, 宮井. 3 次元ウォークスルーと CG 司書を用いた電子図書館インターフェイ

^{*5} ここでいう貸出履歴は、サブ PID と書誌情報が含まれており、個人との対応づけは図書館はできないものである。

- スの開発. 第 19 回情報処理学会情報メディア研究会, pp. 27-34, 1 1995.
- [5] 南. 九州大学図書館における RFID 実証実験の試み. 第 1 回 SLRC Workshop-ユビキタス・コンピューティング社会に向けての RFID の可能性-, 7 2004. <http://www.slrc.kyushu-u.ac.jp/japanese/information/workshop/workshop07.pdf>.
- [6] 田中, 関, 石野, 金丸, 杉本, 竹田, 廣川. 大学経営における大学評価システムの活用. 情報処理学会 第 67 回全国大会, 3 2005.
- [7] 図書館の自由に関する調査委員会(編). 図書館は利用者の秘密を守る. 図書館と自由第 9 集. 社団法人日本図書館協会, 1988.
- [8] 図書館の自由に関する調査委員会(編). 図書館の自由に関する事例 33 選. 図書館と自由第 14 集. 社団法人日本図書館協会, 1997.
- [9] 浜崎, 安浦. PID を用いた安全な社会システムの構想. 九州大学大学院システム情報科学紀要, 7(2):139-148, 9 2002.
- [10] 関, 廣川. 教員連想検索システム. 第 48 回情報処理学会デジタル・ドキュメント研究会, 1 2005.