

経験ウェブにおける個人コンテンツの協調的活用

牛 尼 剛 聡[†] 渡 邊 豊 英^{††}

現在、我々はユーザの経験に基づいて個人コンテンツを管理するシステムとして経験ウェブ (X-Web) を開発している。X-Web では、利用者の経験をモデル化することにより、様々な種類の個人コンテンツを統合化し、コンテキストに基づいた多様な検索と推薦を実現する。個人コンテンツは、個人の活動を表現したり、個人の活動に取り込まれるものである。同一の活動に参加した利用者間でコンテンツを共有することにより、個人コンテンツを効果的に活用できる。本論文では、X-Web において他人の個人コンテンツを利用した効果的な検索について考察する。また、個人コンテンツはプライバシー情報であり協調的活用のためには高度なアクセス管理が必要とされるため、個人コンテンツのアクセス管理についても考察する。

Collaborative Usage of Personal Contents on an Experience-Web

TAKETOSHI USHIAMA[†] and TOYOHIDE WATANABE^{††}

Recently, we are working on X-Web, which is a system for managing personal contents based on experiences of a user. X-Web provide a facility to integrate various types of personal content in a unified manner, and enable a user to search and to recommend based on context. A Personal content represent an activity of a user and is took in an activity of a user. Personal contents are used more efficiently if they are shared in the users who attended the same activity. In this paper, we discuss the search that using personal contents of other users. Also, personal contents require access right management because personal contents are privacy information. So we discuss access right management of personal contents on X-Web.

1. はじめに

近年、個人情報管理 (PIM: Personal Information Management) が注目され、活発に研究されている^{1)~6)}。PIM では、個人が生活する中で現れる多種多様かつ大量の情報を統一的な枠組みで効率的に管理して活用できるようにすることを目的としている。Jones らは PIM において以下の 3 種類の処理が重要であると述べている⁷⁾。

- (1) Keeping
- (2) Finding
- (3) Maintaing

Keeping は個人情報を獲得し、保存できる形式に変換して蓄積する事である。Finding は蓄積された情報から必要な情報を発見するための機能である。Maintaing は蓄積されている情報から不必要なものを削除したり、重複する情報を整理したり、整理したりするための機能である。

代表的な finding の方法に検索とブラウジングがある。検索では、利用者は必要な情報要求をクエリとして表現する。システムは与えられたクエリに基づいて情報を絞り込み、利用者の検索要求に対する適合度を与える。

PIM で管理する情報としては様々なものが考えられる。たとえば、知人の住所、電話番号、電子メールアドレスは PIM の管理対象である。本研究ではパーソナルコンテンツを対象とする。

現在、一般に普及した PIM 検索のためのツールとしてデスクトップサーチがある。デスクトップサーチは、利用者が与えたキーワードに基づいて個人のパーソナルコンピュータ上に蓄積されたファイルを高速に検索する。現在、様々なデスクトップサーチ用のソフトウェアが公開されている。現在普及しているデスクトップサーチ用のソフトウェアは基本的に、テキストファイルに対する全文検索技術とあらかじめ記述されたメタデータ検索技術を利用している。このため、テキスト以外の形式であらかじめメタデータが付与されていないファイルを検索することが困難である。

パーソナルコンテンツの代表例に写真がある。写真をキーワードで検索する場合いくつかの方法が考えら

[†] 九州大学大学院芸術工学研究院
Faculty of Design, Kyushu University

^{††} 名古屋大学大学院情報科学研究科
Graduate School of Information Science, Nagoya University

れる。写真を画像処理で検索することは難しい。写真にアノテーションを付加する方法もある。しかし、利用者が自分が撮影したすべての写真に対してアノテーションを付加することは現実的ではない。もう一つのやり方は、写真に GPS データを対応付けるものである^{8),9)}。撮影場所の地理的な名称を利用して検索する。地理情報に基づいて検索したい場合には有効である。しかし、GPS により撮影場所がわかったとしても利用者の要求に応えられないことがあるな場合がある。例として、職場や自宅といった場所で撮影した写真を検索することを考える。ユーザが職場や自宅に滞在する時間は長く。撮影場所ではほとんど検索対象を絞り込むことができない。

我々は、経験を中心とした個人情報管理 (PIM) を実現する枠組みとして X-Web(Experience-Web) を開発している¹⁰⁾。

現在利用されている多くの個人情報管理システムは基本的に自分のファイルを検索することを前提としている。情報の共有を考えているシステムもあるが、明示的にアクセス管理をするものが多い。他のユーザが管理する情報を安全かつ効率的にアクセスするために必要な十分な機能を有していない。

本論文では、X-Web を利用して他者のコンテンツを利用できるようにする枠組みについて考察する。

2. X-Web

X-Web は個人の経験に基づいて個人コンテンツを管理する枠組みである。本節では X-Web のデータモデルの概要を述べる。

2.1 コンテンツの種類と経験の関係

個人コンテンツは個人が作成したコンテンツと、無償または有償で個人が取得したコンテンツに大別できる。個人が作成したコンテンツをプライベートコンテンツと呼ぶ。プライベートコンテンツの代表例には、デジタルカメラで撮影した写真、送受信した電子メールなどがある。個人が有償または無償で取得したコンテンツをパブリックコンテンツと呼ぶ。パブリックコンテンツの例として、聴取した楽曲、視聴した TV 番組、映画などがある。

プライベートコンテンツは、個人の活動の一部を切り取ったものである。いわば、プライベートコンテンツは経験を外化したものである。一方、パブリックコンテンツは個人が聴取したり視聴したりすることによって、内容が個人の経験となる。いわば、パブリックコンテンツは経験として内化されるものである。

個人コンテンツは、個人の経験に対して、外化また

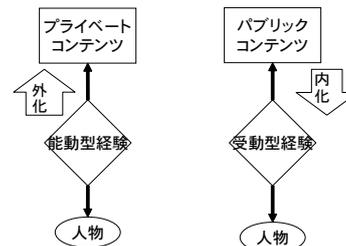


図 1 能動型経験と受動型経験の概念図

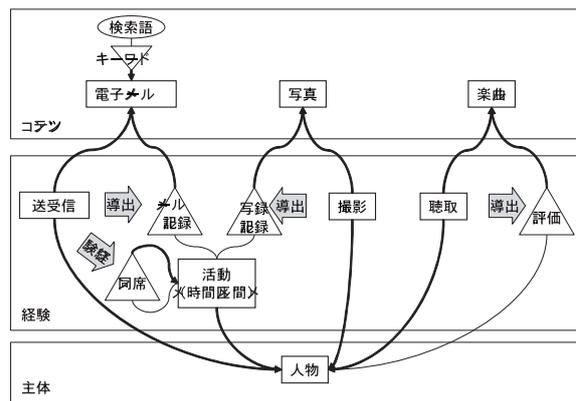


図 2 X-Web のスキーマの例

は内化する対象として関係づけることができる。図 1 に能動型経験と受動型経験の概念図を示す。

X-Web はこのような能動型経験と受動型経験をモデル化し蓄積することにより、様々な種類の個人コンテンツを意味づけ、統一的に組織化し高次の活用を実現する。

2.2 X-Web のデータ構造

2.2.1 ユニットとコンテナ

X-Web において管理対象とするのは、個人コンテンツ、経験、主体である。これらをユニットと呼ぶ。個人コンテンツユニットは、本システムで管理対象となる、写真や楽曲などを表す。主体はユーザを表す。経験ユニットは個人コンテンツと主体の関連として表現される。すべてのユニットは識別子を持ち一意に区別される。

コンテナはユニットの集合を表現するモデル化要素である。コンテナは名前を持ち、名前によって参照される。

経験ウェブではデータ構造のスキーマはグラフとして表現できる。図 2 はスキーマの例を示す。図中の矩形はコンテナを表す。すべてのユニットは「コンテンツ」「経験」「主体」のいずれかのコンテナに所属する。

2.2.2 ランキング関数とクエリパス

ランキング関数はユニット間の重み付きの関連であ

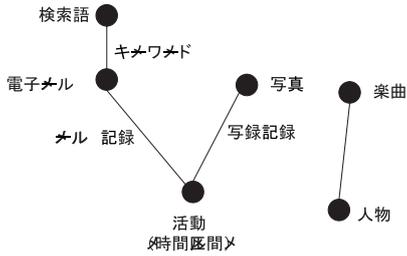


図3 ランキンググラフの例

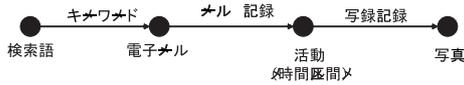


図4 クエリパスの例

る。2 個のコンテナに属するユニット間の関連の強さを表現する。ランキング関数は、蓄積された経験やコンテンツの特徴に基づいて導出される。図2における三角形はランキング関数を表す。

ランキング関数は、コンテナ内のユニットを意味的に結びつける働きを持つ。コンテナをノードとし、ランキング関数をエッジとするようなグラフを考える。たとえば図2に示したスキーマに対しては、図3のようなグラフを考えることができる。このようなグラフをランキンググラフと呼ぶ。

X-Web のクエリはランキンググラフ上のパスを利用して表現する。図4は図3上のクエリパスの例を表している。クエリパスの先頭ノードは、利用者がクエリの対象とするコンテナである。コンテナはユニットのグループを表す概念である。末端ノードは利用者の検索要求が含まれているコンテナである。

クエリパスを用いて検索を行うために、利用者は先頭コンテナに対して初期の重みを与える。与えられた重みを伝搬させて末端コンテナ U_n に属するコンテンツそれぞれに対して重みを決定する。最終的な u_j^n の重みを w_j^n と表記する。 w_j^n の計算には以下の式を用いる。

$$w_i^n = \sum_j \sum_k \cdots \sum_l w_i^0 w_{ij}^1 \cdots w_{kl}^{n-1} \quad (1)$$

最終的な重みが適合度を表しており、重みが高いほど利用者の要求に対する適合度が高いと考える。図5に上記のモデルの概念図を示す。円がユニットを表し、波線の角が丸い四角形がコンテナを表す。

3. 個人コンテンツの協調的活用

ある人物が管理する個人コンテンツが、他者にとっても意味があり、アクセスしたいという要求がある場

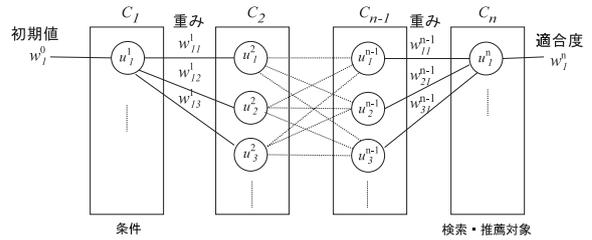


図5 重みの伝搬

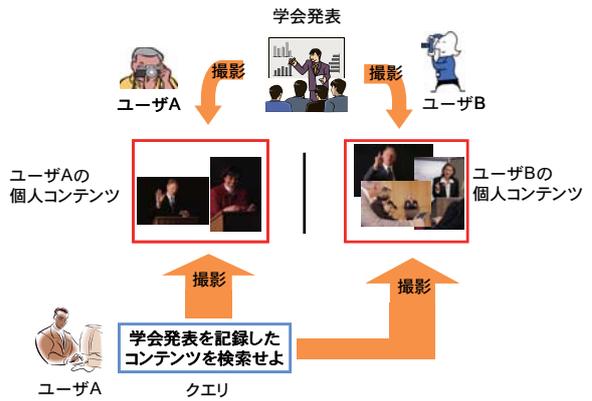


図6 経験に基づくコンテンツ共有の例

合がある。

たとえば、図6に示すように、ユーザAとユーザBが同一の学会に参加し、記録の写真を撮影した場合を考える。ユーザAは報告のために学会発表を撮影した写真が必要となったとき、個人コンテンツ管理システムに対して「学会発表を記録したコンテンツを検索せよ」というクエリを実行するとする。このとき、検索対象にユーザAが撮影した写真だけではなく、ユーザBが撮影した写真が含まれていれば、利用者Aが撮影できなかった場面を記録した写真を使用できるかもしれない。

本論文では、他社の個人コンテンツを利用することを協調的活用と呼ぶ。本節では、X-Webにおいて個人コンテンツの協調的活用を実現するための基本的な考え方について考える。

3.1 個人コンテンツの共有と活動のすりあわせ

活動のすりあわせとは、同一の活動に参加した複数の人物の個人コンテンツを必要に応じて相互に参照可能とすることである。たとえば、同一の授業に参加した学生は同一の活動に参加していると考えられる。学生の授業ノートは、学生の経験を記録したものであると考えられる。学生は試験勉強中に自身のノートを見るが、解らないところが出てくる。同じ授業に参加した友人のノートを見せてもらうことにより、

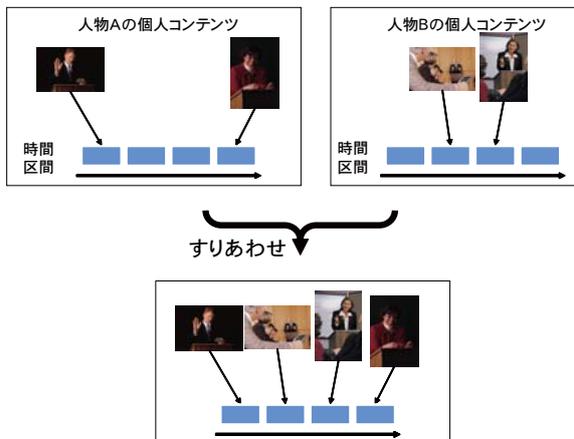


図7 すりあわせの例

自分のノートの不十分な点を補うことができる。活動のすりあわせによって個人コンテンツが、グループコンテンツとして有効的に活用できるようになる。

活動のすりあわせによる写真の共有の例を図7に示す。図の上部は、人物Aと人物Bそれぞれの個人の活動を表す時間区間に対して、そこで本人が撮影した写真が対応づけられていることを示している。人物Aと人物Bが同一の活動に参加していたことがわかり、2者の活動をすりあわせた結果が図の下部に示されている。

3.2 活動の索引付けとすりあわせの種類

X-Webでは利用者のクエリはクエリパスとして表現する。先頭コンテナから末端コンテナへの経路は末端コンテナの要素に対する索引としてとらえることができる。このように考えたとき、X-Webは利用者の要求に応じて様々な間接的な索引を利用して個人コンテンツのアクセスを行うシステムと言うことができる。

これまで、コンテナ間のランキング関数は同一ユーザ内のコンテナを対応づけるもの限定して考えてきた。しかし、異なる利用者が管理するコンテナ間にも対応関係を考えることができる。他者の個人コンテンツやランキング関数を利用したクエリパスを想定することによって活動のすりあわせを実現できる。

索引と検索対象コンテンツの組み合わせパターンとして図8に示す4種類を考えることができる。4種類の特徴は以下の通りである。

【自己索引-自己コンテンツ】 自己の索引を利用して自己のコンテンツにアクセスする。これは、同一人物のシステム内で閉じており、他者との協調的な活用は行っていない。

【自己索引-他者コンテンツ】 自己の索引を利用して他者のコンテンツにアクセスする。他者と活動を共

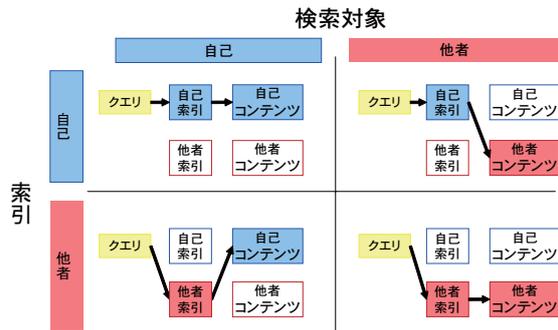


図8 索引と検索対象コンテンツの組み合わせのパターン

有し、自己と同じ目的でコンテンツを記録した場合に有効なパターンである。

【他者索引-自己コンテンツ】 他者の索引を利用して自己のコンテンツをアクセスする。自己の活動の特徴を十分に獲得できない場合に有効である。

【他者索引-他者コンテンツ】 他者の索引を利用して他者のコンテンツをアクセスする。単に、他者の検索システムを利用するだけであり、協調的な活用ではない。

上記の4種類のすりあわせパターンのうち他者の個人コンテンツを利用することにより、検索対象が増大する。自信が保有していないコンテンツを利用したいという要求に対応する。

他者の個人コンテンツを利用する場合と他者の索引を利用する場合に分けられる。他者の個人コンテンツを利用する場合の利点は、検索対象が増大することにある。これにより、ユーザ自信が保有していないコンテンツを利用したいという要求に対応できる。一方、他者の索引を利用することにより、索引の精度の向上が期待できる。X-Webでは個人の経験を様々な形式のデータで特徴づけることができる。この特徴付けには不確実性が含まれる。基準となるデータが増大すれば、特徴付けの精度の向上が期待できるため、他人の索引を利用することで検索の精度向上が期待できる。

3.3 協調的活用のためのクエリ表現

協調的活用のクエリを行う場合、自分と活動を共有していない人物の索引や個人コンテンツを利用しては、希望する結果は得られない。他人の索引や個人コンテンツを利用して検索する際は、利用対象となる人物をクエリ内で明示的に指定する。

協調的活用のクエリは具体的には、キーワードとすりあわせ条件から構成される。すりあわせ条件は、誰の索引を利用するか、および誰の個人コンテンツを検索対象とするかということ指定する。

協調的な活用を行うためのクエリを指定する際に、

すりあわせの対象となる利用者を明示的に指定することは利用者にとって大きな負担となる。個人コンテンツの中のプライベートコンテンツは利用者の経験を記録している。何らかの形で同一の活動を行ったと判断できることがコンテンツのすりあわせが成功する必要条件である。もし、利用者が蓄積しているプライベートコンテンツの情報を利用して個人の活動の同一性を導出することができれば、利用者が与えたクエリに基づいて自動的に、関係する他者の活動とのすりあわせ条件の指定を支援することが期待できる。

図4に示したクエリパスでは、利用者が送受信した電子メールに含まれる時間概念に基づいて写真を検索する。これは、電子メールに含まれる時間概念に基づいて、個人の活動を表す時間区間をキーワードで索引付けするものである。ここでは、メールの本文のみを利用したが、電子メールの送信者と受信者に関する情報は利用者の人間関係を表す重要な情報となる。電子メールに含まれる時間概念に基づいて時間区間をキーワードで重みづけたやり方と同じ考え方を利用すれば、電子メールの送信者と受信者の情報にもとづいて、人物をキーワードで重みづけることができる。この重み付けは、キーワードに対する他者の関連性を示しているために、すりあわせが可能となる可能性の程度として利用できる可能性がある。

4. アクセス制御

本システムの利用者はコンテンツと索引の提供者および被提供者という2種類の役割を持つ。この役割は固定的なものではない。状況に応じて動的に決定する。個人コンテンツの協調的な活用を行う場合には、利用者が望まない他者からのアクセスを排除したいという要求があるため、アクセス制御が非常に重要な問題である。

前節で個人の活動のうち、複数の同一の活動に参加したという事実は個人コンテンツのすりあわせに重要であると述べた。同じ活動に参加しているのであれば、その活動を記録した対象を隠蔽する必要は無い事がある。同一の活動に参加しているという事実はコンテンツの共有の基準として重要である。

アクセス権の指定方法として、もっとも基本的なものは、対象となる個人コンテンツ自体に直接アクセス権を指定する方法である。この方法は利用者に対して多くの手間を要求する。

個々のマルチメディアコンテンツに個別にアクセス権を設定するのは利用者にとって多大な負担を強いる。これは、個々のマルチメディアコンテンツにアノテーション

ンを付加するのが現実的でないことと同様である。

経験ウェブでは経験に基づいてコンテンツを組織化する。経験は時間区間に対応づけることが可能である。この性質を利用して、ここのコンテンツではなく、時間区間に対してアクセス権を指定することが考えられる。時間区間は階層性を持ち、利用者の活動の抽象度や詳細度をふまえて柔軟にアクセス権を指定可能である。また、前節で述べた、電子メールなどの個人コンテンツに基づいて導出された活動の類似性に基づいて、基礎的なアクセス権を指定することも考えられる。

5. おわりに

本論文では、個人コンテンツを対象としたPIM環境であるX-Webにおける個人コンテンツの協調的活用に関して考察した。今後、本論文で示した考察に基づいて実データに基づく予備実験を行い、X-Webにおける協調的活用のための枠組みを明確化していく。

謝 辞

本研究は科学研究費補助金若手研究(B)の助成を受けている。日頃からお世話になっている九州大学大学院芸術工学研究院藤村直美教授及び富松潔教授に感謝します。また、本研究に関して熱心に議論していただいた、名古屋大学渡邊研究室の皆様にも感謝します。

参 考 文 献

- 1) J. Teevan, W. Jones and B. B. Bederson: "Personal information management", *Communications of the ACM*, **49**, 1, pp. 40–43 (2006).
- 2) S. Dumais, E. Cutrell, J. Cadiz, G. Janck, R. Strin and D. C. Robbins: "Stuff I've Seen: A System for Personal Information Retrieval and Re-Use", *Proc. of SIGIR'03*, pp. 72–79 (2003).
- 3) J. Gemmell, G. Bell, R. Lueder, S. Drucker and C. Wong: "Mylifebits: fulfilling the memex vision", *Proc. of ACM Multimedia'02*, pp. 235–238 (2002).
- 4) Y. Cai, X. L. Dong, A. Y. Halevy, J. M. Liu and J. Madhavan: "Personal information management with semex.", *SIGMOD Conference* (Ed. by F. Ozcan), *ACM*, pp. 921–923 (2005).
- 5) E. Adar, D. R. Karger and L. A. Stein: "Haystack: Per-user information environments.", *CIKM*, *ACM*, pp. 413–422 (1999).
- 6) M. Ringel, E. Cutrell, S. T. Dumais and E. Horvitz: "Milestones in time: The value of landmarks in retrieving information from personal stores.", *INTERACT* (Eds. by M. Rauterberg, M. Menozzi and J. Wesson), *IOS Press* (2003).
- 7) W. Jones and J. Gwizdka: "Personal information

- management in theory and in practice”, CHI 2006 Course Notes (2006).
- 8) 上田, 天笠, 吉川, 植村: “位置情報と地理情報を用いた映像データの索引付け手法”, 第 12 回データ工学ワークショップ (DEWS2001) (2001).
 - 9) T. Hori and K. Aizawa: “Context-based Video Retrieval System for the Life-Log Applications”, Proc. of MIR'03, pp. 31–38 (2003).
 - 10) 牛尾, 渡邊: “X-web: 経験に基づく個人コンテンツ管理の枠組み”, 第 17 回データ工学ワークショップ (DEWS2006) (2006).