

パーソナライズ情報配信基盤 PIDEM の提案*

5 F-2

小池雄一† 小向孝典‡ 神場知成† 古関義幸†

†NEC C&C メディア研究所 ‡NEC ソフトウェア(株)

E-mail: koike@mmp.cl.nec.co.jp

1 はじめに

インターネットにおける情報提供サービスは、ここ数年で急速な成長を遂げた。今や、World Wide Web (WWW)上の電子新聞サービスからアプリケーションプログラムの自動バージョンアップ[1]に至るまで、実に多くのサービスがインターネット上で提供されている。このように配信される情報の量が増大する中、より重要度や必要性の高い情報を選択したいという要求がユーザから生まれて来た。これに応える手段として、各個人の嗜好や環境に応じて情報を加工・選択して提供する**パーソナライズ技術**が注目を集めている。パーソナライズ技術を適用した情報サービスの例に読売COLINS[2]が挙げられる。この電子新聞サービスは、ユーザが指定したキーワードに関連する記事を優先的に表示する等の機能を備えているため、ユーザは自分の興味のある記事を効率良く閲覧する事が可能となる。

パーソナライズ技術を用いた情報配信サービスは、今後ますます増大していくと考えられる。この時、サービスを構築する際にパーソナライズ機能、配信機能、コンテンツ管理機能などを個別に実装することは非効率である。このため、本論文ではパーソナライズ情報配信基盤 PIDEM (Personalized Information Delivery Middleware: パイデム)を提案する。PIDEM は、コンテンツ管理、パーソナライズ、配信といったパーソナライズ情報配信システムを構成する各要素を再利用可能なソフトウェアコンポーネントとして実現し、それらを統一的に扱うための API を規定したものである。システム構築者は、PIDEM に対応したコンポーネントを組み合わせることにより容易にサービスを構築する事が可能である。

2 PIDEMの構成

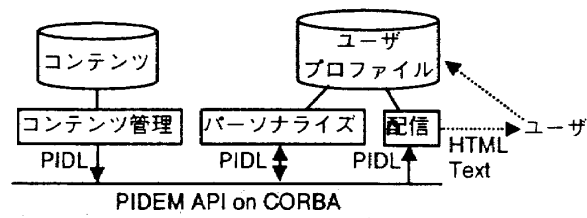


図1 PIDEMの構成

PIDEM の主要な構成要素を図1に示す。PIDEM 文書は、PIDEM を構成する各コンポーネントに共通の文書フォーマットである。筆者らは、パーソナライズされた情報を格納するため、PIDL(Personalized Information

Description Language)という XML に基づいた独自のマークアップ言語を開発した。PIDL の詳細に付いては後述する。

コンテンツ管理コンポーネント群は、様々な形式のコンテンツを PIDEM 文書形式に変換する。コンテンツとしては、データベースに格納された社内情報、WWW ページから収集した情報等、多様な形態があるが、その形態に対応したコンテンツ管理コンポーネントを用意する事で、統一的な I/F でコンテンツを扱う事が可能となる。

パーソナライズコンポーネント群と配信コンポーネント群は、情報のパーソナライズを行う。パーソナライズコンポーネントはいわゆるパーソナライズ、すなわちユーザの嗜好や属性に基づいた情報の加工・選択を行う。また、筆者らはユーザの嗜好等に加え、ユーザが情報を受信する環境、すなわち

1. ユーザが情報を受信するプロトコル (SMTP, HTTP, FAX, Pager 等)
2. ユーザの端末の属性 (端末の表示面積、通信速度、端末の CPU 能力等)

に応じて情報を加工する事も、広義の意味でのパーソナライズと考える。このため、配信コンポーネントは、ユーザの情報受信環境に基づいた情報の加工・選択を行う。同時に、PIDL 形式で記述された PIDEM 文書を、ユーザが受信可能な形式 (HTML, テキスト等) に変換する役割も果たす。ユーザの情報受信環境は非常に多様であるが、受信環境に応じた配信コンポーネントを用いる事で、サービス構築者は情報受信環境をあまり意識せずにサービスを構築する事が可能である。

コンテンツ管理、パーソナライズ、配信の各コンポーネント群を操作する API は CORBA に基づいており、サービスを構成する各コンポーネントを分散させておくことも可能である。

3 PIDEMにおける情報交換フォーマット

3.1 PIDEM文書フォーマットの要件

第2章で述べたように、PIDEM を構成する各コンポーネントは PIDEM 独自の形式の文書を交換する。各コンポーネントが同一の形式を持つ文書を交換するため、PIDEM ではコンポーネントを組み合わせる事が可能である。パーソナライズされた情報を格納する形式として、PIDEM 文書には以下のような特長が要求される。

* PIDEM: Personalized Information Delivery Middleware. † Yuichi Koike, ‡ Takanori Komukai, † Tomonari Kamba, † Yoshiyuki Koseki. † NEC C&C Media Research Laboratories. ‡ NEC Software Ltd.

要件一：複数コンポーネントの加工結果を格納可能
 通常のパーソナライズ情報配信サービスでは、複数のパーソナライズコンポーネントによって加工された結果をユーザに配信する。このため、PIDEM 文書は複数のパーソナライズコンポーネントによる加工結果を効率良く格納できなければならない

要件二：複数ユーザに対する加工結果を管理可能
 パーソナライズコンポーネントはユーザ毎に異なる加工を行う。このため、文書はユーザ毎に異なる加工結果を効率よく蓄積、参照できなければならない

要件三：ネットワーク透過
 各コンポーネントが分散している場合に対応するため、PIDEM 文書はネットワークを通してやり取りされることが可能である必要がある

3.2 PIDLの構造

第一の要件を満たすため、PIDL は元々のコンテンツと、パーソナライズコンポーネントによる加工結果とを別に保持している。このため、図2に示すように複数のコンポーネントが同一文書を加工した後でも、いずれか一方のコンポーネントのみによる加工結果を参照する事が可能である。

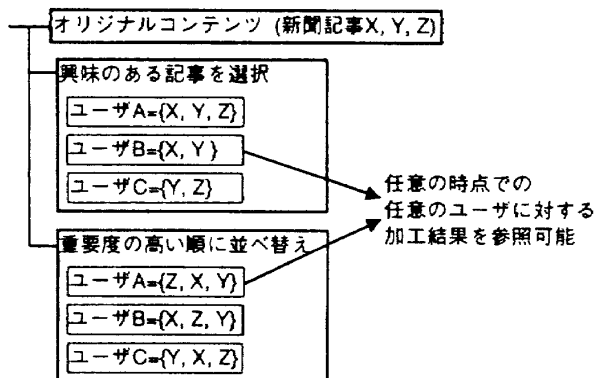


図2 PIDLの構造

更に、PIDL では加工後の文書そのまま格納するのではなく、加工方法とそれに必要な付帯情報のみを格納する。例えば、加工が「複数の新聞記事からユーザが興味を持つものだけを選択する」というものである場合、PIDL では各記事を選択するかどうかのフラグだけを加工結果として保存する。そして、外部からの加工結果へのアクセスに応じて、加工後の文書を作成する。このため、複数のコンポーネントによる加工結果を一つの文書に格納する場合でも、少ない容量で効率よく加工結果を保存する事が可能である。この格納方法を用いる事で、複数ユーザに対する加工結果も効率良く格納出来るため、第二の要件も満たされる。

3.3 PIDLの操作

第三の要件を満たすため、PIDEM 文書を CORBA オブジェクトとして実現した。PIDEM のコンポーネントが PIDEM 文書をやり取りする際には、文書へのポイントのみを扱う。各コンポーネントは、必要に応じて PIDEM 文書の内容にアクセスするため、ネットワークを通して PIDEM 文書を効率良く扱うことが可能

である。また、PIDEM 文書は

1. 特定のユーザのみに関係する加工結果
2. 特定のコンポーネントに関連する加工結果

を木目細かく指定して取り出す事が可能なインタフェースを備えている。このため、図2の例で言えば「ユーザ A が興味を持つ記事のみを選択」、「ユーザ B が興味を持つ記事を重要度の高い順に並べる」等の加工を行った後の文書を自由に取り出す事が可能である。

4 PIDEMの適用例

PIDEM の有効性を確認するため、PIDEM を用いたパーソナライズ情報サービスを二例構築した。一つはイントラネット向け技術情報配信システムであり、ユーザの職位や興味分野に応じて情報をパーソナライズし、WWW、Pager 等多様な方法で配信する。もう一つは、新製品情報を表示する Web ページにパーソナライズ機能を付加したもので、ユーザの設定したキーワードに関連する新製品情報を強調表示する等のパーソナライズ機能を持つ(図3)。二例は、互いに異なるサービスであるが、キーワードマッチ等の汎用パーソナライズ部品を用いる事で、構築の工数を削減する事が出来た。

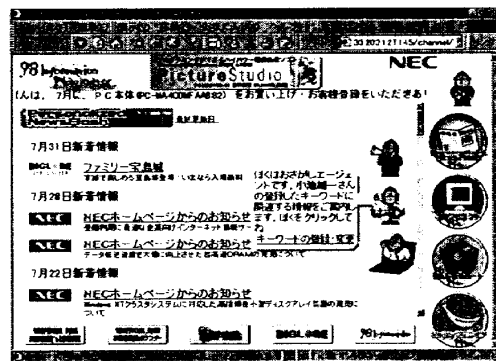


図3 PIDEM適用例

5 まとめ

本稿では、パーソナライズ情報配信基盤 PIDEM を提案した。PIDEM ではパーソナライズ情報配信システムを、1) コンテンツ管理、2) パーソナライズ、3) 配信、の三要素にモデル化し、それぞれの部品を同一の操作で扱うための API を規定した。このため、様々な部品を組み合わせる事でパーソナライズ情報配信システムを構築する事が可能である。

参考文献

- [1] Marimba: Application Distribution and Management (<http://www.marimba.com/>)
- [2] 坂上秀和、神場知成、古関義幸. 「パーソナル電子新聞ANATAGONOMYの開発と評価」 インタラクティブシステムとソフトウェア IV (WISS'96) pp 21-30. 近代科学社, 1996.