

XML を用いた知識管理方式の提案と考察

梅田恭子、安田孝美*、横井茂樹

名古屋大学大学院人間情報学研究科、名古屋大学情報文化学部*

近年様々な分野で、ネットワークを介した知識の形成、蓄積、共有、提供・利用を支援する知識情報共有環境の実現が求められてきている。本稿では、特に学術情報の共有に焦点をあて、XML を用いた研究室における知識管理方式を提案する。本方式では、論文等の形式化された知的資産に対して、アイデア等の従来定型化されていない情報を、メモという形で形式化し蓄積する。ユーザが入力時にメモをつける対象を明示することにより、独立して保存されていた文書がメモにより明確につながるようになる。このメモを通して知識情報の共有を行うことにより、文書に補助的情報を与えるだけでなく、文書の集合化が図られ、それにより新たな知識を作り出す効果が期待される。

A proposition of a management system based on XML for knowledge-based information.

Kyoko Umeda, Takami Yasuda* and Shigeki Yokoi

Graduate School of Human Informatics Nagoya University,

*School of Informatics and Sciences Nagoya University

In recent years there is an increasing need for realization of knowledge-based information share environment, which supports knowledge production, accumulation, sharing, and use over network in various fields. In this paper, we focus on a share of study information and propose a management system basing on XML for laboratory knowledge-based information. It formalizes and accumulates practical information such as an advice as memorandums attaching to documents such as paper. Since the user indicates the object of attachment when he makes the memorandum, the independently document can be exactly connected by it. By using this system and analyzing its working results, we can reorganize the existing documents as a new set, and create new knowledge.

1. はじめに

知識共有や知識管理システムの重要性が社会で高まるに連れそれらに関する研究が盛んに行われ、情報の共有の方法や枠組みの提供だけでなく、コンテンツの充実が大切であることが明らかになってきた。¹⁾ また有意義な知識情報を発見する方法等の研究も行われ²⁾、WWW 文書の為の記述方法の標準化も進んでいる。更に知識環境を構築する際には、いかに質の高い知識を蓄積し、どのように利用し、フィードバックするかというサイクルを考慮する必要性³⁾も論じられている。本稿では、特に学術情報の共有に焦点をあて、研究室における知識管理方式を提案する。

研究室における活動は、創造的な部分が多く試行錯誤を繰り返しながら様々な知識が生み出され

る。それらの知識情報がデータ化されたものが論文や発表資料等の形式化された知的資産であり、何らかの方法で蓄積されている。しかし、その他の定型化されていない知識をメンバーで共有して活用することも重要である。研究のヒントを得る機会は、様々な人とのディスカッションやアドバイスからも生まれるが、多くの場合それらの情報はデータ化されていない。本稿ではそのような非定型情報に注目し、既に形式化された知的資産に補助的情報を与えることで、知識情報の集合化を図り、質の高い知識の提供・利用が行われることを目的とした研究室における知識管理方式の提案する。更にはそれをフィードバックして新しい知識が生み出されることを目指している。また、提案した知識管理方式のシステムの試作を行った。

2. 知識管理方式の提案とシステム構成

2.1 情報共有の現状と XML の活用

現在、研究室では論文等の形式化された知的資産は、管理用のメタデータと資料本体が切り離されてライブラリ DB として管理されている。また、CG 等の動画データも蓄積・活用することが試みられている。さらに、研究室における DB だけでなく、他の研究室や機関の DB との連携も視野に入れ、ある特定の DB に特化するものではなく汎用性を持って使えるものにしたいという点と、WWW を情報の入出力の窓口にしたという点から、DB 間やシステム間のデータ交換方式として XML を用いる。また、今後メモやデータ項目が変更される可能性がある為、柔軟なデータ項目の設定やデータの階層化表現が可能であり、項目の任意数の繰り返しを木構造で表現可能であるという利点に注目し、知識処理システムにおける情報の保存にも XML を採用することにする。

今回のシステムの試作は、論文等が保存されているライブラリ DB についてのみ行った。しかし、上述したように XML を交換方式としていること、またメタデータのみを知識処理の対象としていることから、例えば画像 DB が連携されても本方式の基本的な方針は変わらないと考えられる。ただし、説明をわかりやすくするために、以後は定型化された知的資産のことを、論文やレポート等の文書と表現する。

2.2 知識管理方式の概要

本方式では、アイデア等の従来定型化されていない情報を、論文やレポート等の文書に情報を付加するメモという形でデータ化し、蓄積する。二つ以上の文書に関連する情報であれば、一つのメモを全ての文書に関連付ける。このことにより、独立して保存されていた文書がメモにより明確につながることになる。このつながりは、文書の検索や、現状の研究分析に利用される。更には、検索や分析の結果、新しい関連を作っていくことで新しいアイデア等を形成する、というサイクルを繰り返すことを目指している。形式化された知識と非定型な知識を連携させることに関して、自然言語解析やキーワードで間接的に指定する方法もあるが、初めからユーザがメモをつけることで直接その対象を明示していることに本方式の特長がある。

この方式により、以下のような効果があると考えられる。

- ・ 論文等の文書にメモをつけることにより、文書に補助的な情報を与え、検索する際の助けとなる。
- ・ メモを通して独立していた文書に関連ができ、従来は無関係であった文書を集合として見ることが出来る。
- ・ それらの集合を分析し、再びメモを作ることにより新しい知識を作り出す支援となる。

2.3 知識管理方式のシステム構成

本項では、まず知識メモについての提案を行い、次にシステムについて説明する。本システムは、情報の入出力は WWW 上で行き、入力された知識メモを蓄積する知識メモ群、知識メモと既存の DB の知識情報を結び付ける知識処理システム、WWW 上における知識情報の入出力のための知識情報利用インターフェースからなる。(図 1)

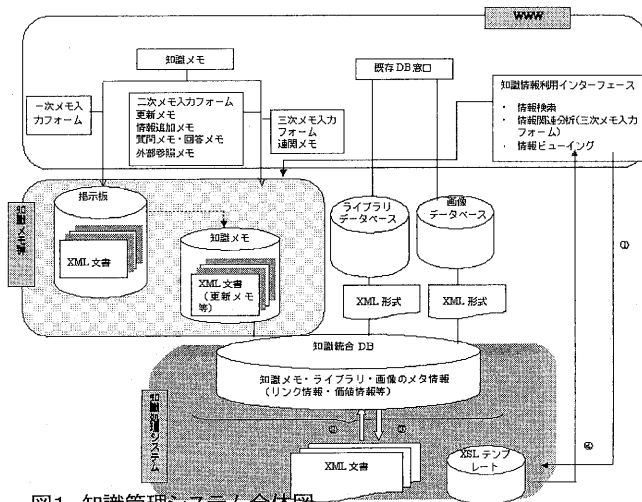


図1 知識管理システム全体図

2.3.1 知識メモ概念の提案

非定型な情報を収集し活用する際、処理がこないやすいように、本稿では論文等の知的資産に付加する情報を知識メモと呼び、関連の付け方により次の3段階に分類した。

一次メモ：メモを付ける対象を明確に持たない情報。特に本稿の場合においては、ライブラリ DB に納められた論文等の文書に直接関連しない情報。
 二次メモ：メモを付加したい対象が明確に分かっている情報。例えば「この論文はライブラリ DB の学会フォルダにある B 論文を更新したものです。」等。
 三次メモ：二次メモを人が分析して別の情報と関連付けや再構成を行ったもの。三次メモは二次メモと考えることもできる。

想定される知識メモの推移をまとめると、図2のようになる。一次メモは自動的、あるいは人により二次・三次メモ、文書となり、二次メモは検索や分析により新たな情報が付加されるなどして三次メモとなる。さらにそれらが再構成され、利用されて新しい文書となり得る。このように、知識メモの利用を通して、非定型的な知識を形成、蓄積し、再利用するということが本方式の一つの目標である。

2.3.2 知識メモ群：知識メモの収集・蓄積

上述した知識メモを次のように扱う。

まず、従来のラウンジ型掲示板でやりとりされている情報を一次メモと考える。図2のように、ラウンジ型の掲示板は一つのトピックに対し、それに関して返信するという形で使われている為、あるトピックに関するまともは関連がある確率が高い。この掲示板でやり取りされる情報を一次メモとして XML 形式で保存する。従来の掲示板方式をとることにより、利用者側に特別なことを行っているという印象を与えずに情報を収集が可能で

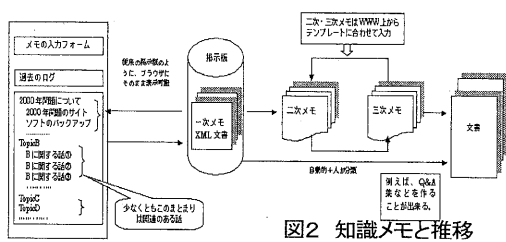


図2 知識メモと推移

あり、また死蔵されていた情報を再利用できる。特に XML で内容が記述されているため、自動的に整理できる部分もある。一次メモは整理されて知識メモ群やライブラリに統合される。

二次メモ・三次メモは、WWW上でいくつかのテンプレートを用意し、それらにメモ付けしたい対象情報を含めて入力し、知識メモとしてシステムに XML 形式で蓄積される。本システムでは主に二次メモ用として文書の更新・情報追加・質問・回答・外部参照メモ、主に三次メモ作成用として連関メモを用意する。このメモの用途は以下のように想定している。

更新メモ：レポート等の更新情報や更新理由を記述。更新前と更新後の文書を関連付ける。

情報追加メモ：レポート等にアドバイスや参考情報を追加。

質問メモ：文書等への質問を記述。メモに記入すると同時に質問する相手にメールが届く。

回答メモ：質問メモに対する回答を記述。同様に質問者へメールが届く。

外部参照メモ：研究室外への文書やインターネット上のサイトと関連付ける。

連関メモ：文書検索や知識メモの検索から得られた情報を分析して、関連づけたい情報を記述。少

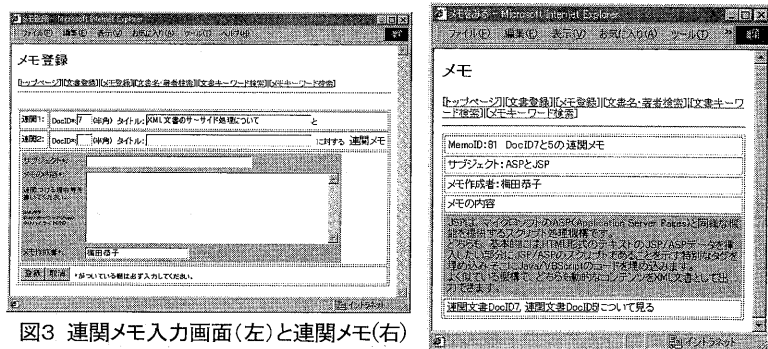


図3 連関メモ入力画面(左)と連関メモ(右)

なくとも二つ以上の文書、もしくはメモが関連付けられる。(図3)

基本的にはこれらのメモは、ライブラリ DB に文書を登録する際や、文書を検索した結果に対して入力することとなる。

このように、可能なものは入力時から明確な文書への関連付けを行い、文書のグループ化を図ることに本方式のねらいがある。

2.3.3 知識処理システム：文書と知識メモの統合

知識処理システムに、ライブラリ DB から必要なメタ情報を取り出し、XML 形式で記述する。この記述は Dublin Core⁴⁾ を参考にして行った。同様に XML で記述された知識メモと統合させ、リンク情報や価値情報を含んだメタ情報 DB を作る。メタ情報には、リンク情報等を追加・更新し、XLink⁵⁾⁶⁾ の仕組みを用いて文書と非定型情報を結び付ける鍵とする。

例えば、図4のように文書Aから文書Bへの更新メモをこの二つの文書に作成することを考える。まず、ユーザが更新メモに文書AとBの文書IDを記入して登録すると、システム内で更新メモのリンク要素に文書AとBのURLを書き込む。次にシステム内で文書AとBのリンク要素に更新メ

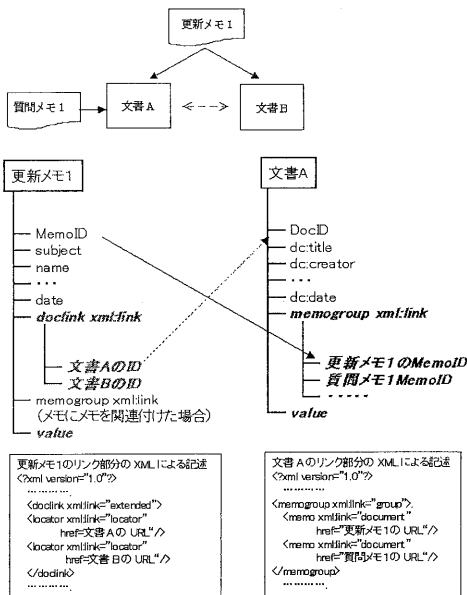


図4 知識処理システムにおけるリンク処理

モの URL を反映させる。このことにより、XLink の拡張リンクグループを用いて、文書検索で文書 A が結果として表示されると、更新メモや文書 B を呼び出すことが可能になる。また文書 B にもメモがあれば、それも同時に呼び出すことができる。XLink はまだ W3C でのドラフトの段階であり、今後も規格は変わる可能性があるが、基本的に XLink の機能を用いれば同様のことができると考えられる。

このように、文書のメタ情報に、知識メモのリンク情報を反映させることにより、文書間の関連付けを容易に持たせることが可能になる。

このメタ情報 DB から欲しい情報を取り出し、WWW の知識情報利用インターフェースに表示する。また、取り出された情報を変換する XSL テンプレートも置く。

2.3.4 知識情報利用インターフェース

ユーザは、WWW上で知識情報の利用、再構築を行う。情報検索、情報ビューイング、情報関連分析3つのプロセスを通して知識情報が循環していく支援となることを目指している。

情報検索には従来の文書検索結果に加えて関連づけられている知識メモ情報も表示する文書検索と、逆に知識メモをキーワードで検索し、関連づけられている文書を表示する知識メモ検索の2種類を用意した。前者はよく使われる検索方法であり、文書をタイトルやキーワードで検索すると、更新メモがついていれば、それにより文書に対して訂正があったことがわかり、検索結果の文書よりも新しい文書を得ることが出来、情報検索の助

メモ ID	メモのサブジェクト	メモの作成者	登録年月	カテゴリ	メモが付けられている文書ID
22	インテリジェントエージェントとJava	原崇	00/02/15	更新メモ	5, 6
25	エージェントとは？補足	上捨石弘	00/02/16	情報追加メモ	1
26	エージェント=プログラム？	原崇	00/02/16	関連メモ	1, 6
28	人工知能=エージェント？	加藤智也	00/02/16	関連メモ	4, 1
29	知識エージェント=AI+インターネット	ハウザー・クリストフ	00/02/16	関連メモ	1, 2

図5 知識メモのキーワードによる検索

けとなる。また更新理由を見て、自分が原稿を書く際の参考になる可能性もある。後者は漠然としたアイデアに対して情報が欲しいときに役立つと思われ、検索結果が関連付けられた文書やメモの集合になっているため、従来より効率より検索が可能となる。(図5)

情報ビューイングでは、上述した検索結果をたどっていくと文書とメモのつながりが分かるようになっていく。例えばメモを作成した時点では無関係と思われた文書が、リンクをたどっていくとつながっていることが分かり、新しいことを発見する可能性がある。

情報連関分析においては、情報ビューイングなどで分析した結果である新しい知識を、連関メモや情報追加メモなどを利用して三次メモとして作成する。このように新しい情報が登録されることにより、上述した流れが繰り返される。このことにより、文書の有効利用が促進され、断片的であったメモ情報自体が価値を持つものになる効果も期待される。(図6)

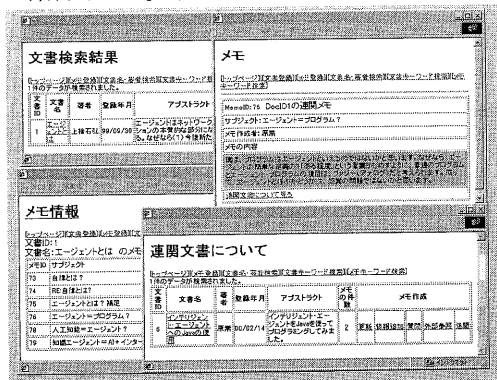


図6 知識メモを用いた連関分析例

3. 本手法の使用例と考察

本手法を考察する為、研究室内のエージェント技術をトピックとした研究グループにおいて行われた議論や作成されたレポートに基づいた、本システムの使用例を示す。

- (1) まず、A君が「エージェントとは?」というレポートを作成し、文書をライブラリDBに登録する。
- (2) ディスカッションに欠席したBさんがレ

ポートを読んで、「自律とはどういうことですか?」という質問メモを作成する。これは、レポートに関連付けられた質問メモとして登録されると同時に、A君にメールが届く。

- (3) メールが届いたA君は回答メモで回答する。これも、レポートに関連付けられた回答メモとして登録されると同時に、Bさんにもメールとして回答が届く。

- (4) また、Javaを用いてエージェントをプログラムしようとしたC君が、「エージェント」というキーワード検索を行い、エージェントレポートを活用して自分のレポートを完成させたとする。例えばエージェントレポートのどの部分を引用し、どの様に活用したかを連関メモとして登録する。

以上のように文書登録やメモ登録、分析を重ねることにより図7のような情報の連関分析が可能となる。このとき以下のような効果がある。

- ・ 文書を検索するとき、文書のアブストラクトだけでなく、メモの情報も参考にできる。
- ・ 「ネットワーク通信」「人工生命」といった一見無関係であるような文書同士に関連ができる。さらに、いくつかの文書が結び付き文書の集合ができる。
- ・ その集合から、「エージェントとは?」という最初のレポートに対して新しい定義やより深い理解が得られたことになる。

以上のことから、本手法は知識情報の形成、共有、利用、更にはそれらをフィードバックして新しい知識を生み出すという一連の流れを支援することが可能である。

また、非定型な知識を収集する方法⁷⁾については、主に二次メモ用として作成したメモテンプレートは、無理な作業を課すことなく作成できる場合が多い。例えば、研究室のゼミで発表したレポートを登録したときに、教官から受けたアドバイスを情報追加メモとして登録する、学会で発表した資料を発展させて得られた新しい研究成果を登録するときに、更新メモを作成する等が考えられる。三次メモに関しては二次メモよりも労力が必

要とされるが、誰かの知的資産を利用したときに、その代わりとして自分の分析結果を登録するというようなルールを作っておけば、知識の収集・蓄積は、無理なことにはならないと思われる。

問題点としては、現在はライブラリ DB から知識処理 DB に変換されるのに時間がかかることから、登録したばかりの文書を分析することができないこと、IE や Netscape といった良く使われるブラウザが XLink をサポートしていないため、その効用を十分には使用できていないこと、また本手法ではある程度知識メモが蓄積されるまで上述した効果が得られないことが挙げられる。前者の二つについては、ソフトやアプリケーションの開発に従って解決されることと思われる。⁸⁾ 後者については利用履歴やキーワード等による関連付け手法との複合が考えられる。

4. むすび

本方式では、アイデア等の従来定型化されていない情報を、論文やレポート等の文書に情報を付加するメモという形でデータ化し、蓄積した。独立していた文書と、断片的な知識情報が関連付けられることにより、知識の共有や提供、利用が行える。例えば、今まで個別で無関係であった文書がメモの内容により集合としてみる事が可能

になり、検索の助けや新たな知識集合の発見ができる。また、断片的な知識メモが既知の文書を中心に集合することにより、アイデアを生み出し易い環境を提供している。

技術的側面から考えると、本方式では、二つ以上の文書に関連する情報であれば、一つのメモを全ての文書に関連付けているが、これは XLink の仕組みを用いることによって容易に実現できると考えられる。また、メタ情報のみを管理しているので、メモをつける対象は文書に限らず、画像でも動画でも同様に扱える。

今後の課題として、知識情報への価値基準の付加が考えられる。価値を付与することにより情報の削除などの整理や分類が可能になり、知識情報の再利用がさらに促進されると思われる。これに関して様々な研究が行われていることから、それらを参考に検討したい。

参考文献

- 1) 中山,真鍋,竹林: 知識情報共有システム (Advice/Help on Demand) の開発と実践: 知識ベースとノウハウベースの構築 情報処理学会論文誌 Vol.39 No.5 1186-1194
- 2) 大森,斉藤: 利用履歴に基づくデジタルコンテンツへの価値情報付与方式の提案 情報処理学会 DD 研究会 No.21 pp1-8
- 3) 池田,山本,高田,中小路: コミュニティ知識ベース環境の構築へ向けての知識の形成と利用に関する調査と分析 情報処理学会論文誌 Vol.40 No.11 pp3887-3894
- 4) Dublin Core Metadata Initiative: <http://purl.org/dc/>
- 5) XLink Working Draft
3-March-1998:<http://www.w3.org/TR/1998/WD-xlink-19980303>
- 6) XML/SGML サロン: XML 完全解説 技術評論社 1998
- 7) 森田,高梨: 入門ナレッジマネジメント かんき出版 1999
- 8) BtoB の標準言語 XML の“期待値” ネットワークコンピューティング 2000.1月号 pp26-37

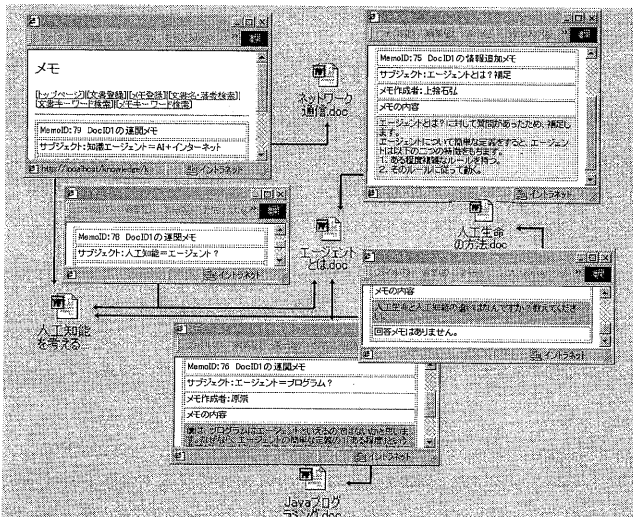


図7 知識メモによる連関分析により結果得られた知識情報の集合 (矢印は分かりやすいように書いたもの)