

分散型ノウハウ蓄積システム GoldFISH の検討

関 良明 藤木 直人

NTT通信網総合研究所

組織における知識の共有促進を目指して、コンピュータと通信ネットワークを利用したシステムの研究開発を進めている。その一環として、適用領域を単一グループ、単一サーバに限定したノウハウ蓄積システム FISH を構築し、2年間の利用実験を進めてきた。その結果、ネットワーク環境の進展に伴う分散環境への適応について、さらに検討を要することがわかった。本論文では、情報の有効範囲/ジャンル/用途等が同一な情報群、複雑なグループ構成、グループ構成と独立に構築されるネットワーク環境の3層に対応できる分散型ノウハウ蓄積システム GoldFISH の検討結果を述べる。

GoldFISH: Group Oriented Long-term Distribution with FISH

Yoshiaki Seki, Naoto Fujiki

NTT Telecommunication Networks Laboratories

e-mail: {seki, fujiki}@nttmhs.ntt.jp

1-2356-523A, Take, Yokosuka, Kanagawa 238-03, JAPAN

Our research focuses on knowledge propagation in organizations. To cope with this issue, we developed FISH, which stands for Flexible Information Sharing and Handling system. FISH was designed to provide cooperative information sharing in a group work context and to support know-how information propagation. Based on the results from a two-year FISH experiment, adaptation to the following three layers; (1) information groups, (2) organizational structure, and (3) network environment came to be key issues. This paper presents design and method arguments on GoldFISH which consists of distributed servers.

1. はじめに

組織やグループにおける知識の蓄積／共有が注目されている [1],[2]. 個人が所有している知識や、グループにおける過去の経験を、コンピュータを利用して、メンバー間で利活用しようとする試みである。また、非公式で非定型な情報が大規模な組織で失われ易いことに着目し、このような情報を流通させることによって協調行動を支援するシステムと、その運用結果が報告されている [3],[4].

一般に組織は複数のグループから構成され、そのグループは互いに独立に存在するのではなく、相互に関係を保ちながら "マルチグループ" として、組織を構成している。このような環境で用いられる通信ネットワークは、電話や電子メールに代表される個人間の通信から、電子掲示板に代表されるグループ内の情報共有、さらには、マルチグループにおける知識伝播を支援する方向に発展すると考えられる。

筆者らは、組織における知識の共有促進を目指して、個人が所有している動的／非定型／断片的な情報を、通信ネットワークで結ばれたコンピュータに蓄積し、グループ内で共有するノウハウ蓄積システム FISH: Flexible Information Sharing and Handling system を3年前に開発した [5]. また、FISH の適用領域を単一グループ、単一サーバに限定して、研究者の勤務するオフィスに導入し、2年間に渡る利用実験を進めてきた [6]-[9]. その結果、今後の研究課題として、次の3点が重要であることが明らかになった。

- (1) 分散環境への適応.
- (2) ATM 技術を背景とするマルチメディア化.
- (3) あいまい検索や意味理解等の言語処理の導入.

ここでは、複数のグループが複合された環境で FISH の利用実験を行い、マルチグループにおける知識の共有／伝播の支援形態を模索するために、"分散環境への適応" について、まず検討することとした。具体的な実験システムとして、分散型ノウハウ蓄積システム GoldFISH: Group Oriented Long-term Distribution with FISH を設計した。

本論文では、情報の有効範囲／ジャンル／用途等が同一な情報群、複雑なグループ構成、グループ構成と独立に構築されるネットワーク環境の3

層に対応できるマルチサーバ化等の機能拡張を図った GoldFISH の研究展開と方式の選択について述べる。

2. FISH の概要と利用実験

FISH の概要と、2年間の利用実験の結果を概観し、FISH の有効性と限界を述べる。

2. 1 FISH の概要

(1) 管理手法

FISH では、内容が多様で従来のデータベースのような属性と値による定型化が困難な情報を仮想的な複数のカードに細分化し、各カードにキーワードを付与して表現する。さらに、図1に示すように、ハイパーテキストの概念を応用して、カード間をリンクで結び、関連情報の参照を容易にしている。このリンクはシステムが自動生成する。また、質問に回答する形で情報を蓄積する質問／回答カードを設け、インタラクティブなノウハウ蓄積を支援している。

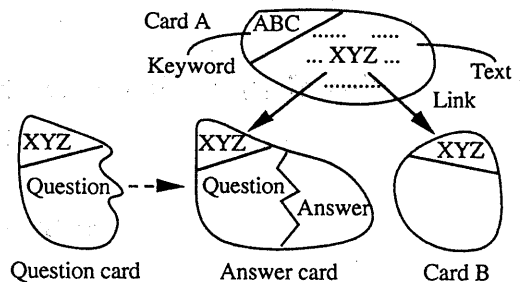


図1 カードとリンク

(2) システム構成

FISH は、図2に示すように、クライアント／サーバ・ワークステーションと、パソコン等の端末、及びこれらを結ぶ通信ネットワークで構成される。複数のユーザが、複数種類の端末から通信ネットワークを介して、カードの登録／参照処理を実行できる。また、システム内のすべてのユーザにカードを公開する Public と、他ユーザには公開しない Private の2つの格納領域を設けている。

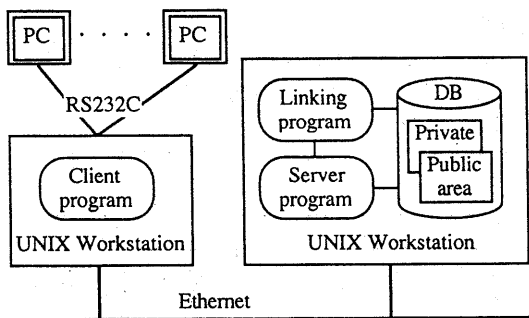


図2 システム構成

(3) ユーザインタフェース

多様な端末に対応するために、FISHは、検索／参照／登録の命令をキーボードを使ったコマンド入力形式で提供している。図3の画面例に示すように、参照コマンドでカードを表示した時、そのカードからリンクが張られている関連カードのリストが同時に表示される。ユーザは、情報の内容による分類や登録した順序に制約されることなく、どのカードからでも参照を開始できる。

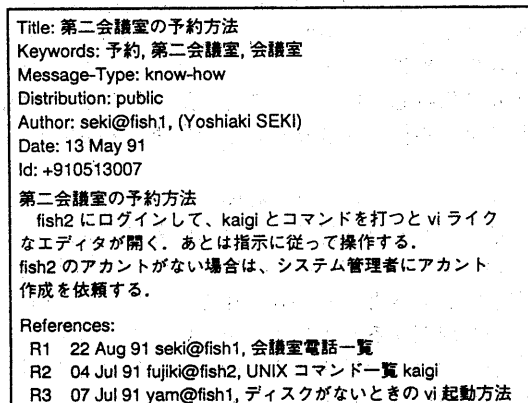


図3 画面例

2.2 利用実験

(1) 実験環境

FISH を以下に示すオフィスに導入した。

- ・ ユーザ数： 16～20 人
(同一グループ内の研究者, 途中一部異動あり)
- ・ 端末台数： 1 ユーザ 1 台
- ・ 実験期間： 24 ケ月

・ 既存サービス： 電子メール, 電子掲示板
電子メールと電子掲示板は、業務連絡用(打ち合わせ開催通知, 業務依頼, 出張通知, 情報交換等)にグループ内で利用が義務付けられている。

(2) 実験方法

FISH の導入に際して、利用方法をユーザに周知し、オンライン・マニュアルを用意した。また、FISH に関するユーザからの質問には、主に電子掲示板で回答した。さらに、有効性や利用動向を観察するために、FISH の利用をユーザに義務付けたり、特定業務への適用は行わなかった。

実験過程の記録は、登録に関しては、「いつ／誰が／そのカードを／登録した」がわかるようカードに属性を付与した。また、参照に関しては、「いつ／誰が／どのカードを／参照した」がわかるよう履歴データを蓄積した。

2.3 実験からの考察

利用実験によって明らかにされた FISH の機能性／有効性／性能と、ユーザの利用動向及び蓄積された情報を分析した結果を以下に概観する。

(1) FISH の機能性

FISH には、約 1000 件の情報がカードとして蓄積され、約 12,000 件の参照が行われた。また、全カードの 78% にリンクが張られ、システムがリンクを自動生成する機能も予想以上に機能した。しかし、以下のようなユーザ要求もあげられた。

- ・ 情報のジャンルによるサーバの分割。
- ・ カード間の不要なリンクの切断。
- ・ 登録者以外のユーザによるカードの書き換え。

(2) FISH の有効性

利用が義務付けられていないにもかかわらず、電子メールや電子掲示板とともに FISH が 2 年間使われ続けたことは、このようなシステムのニーズを示している。また、FISH を経由した知識伝達や、情報の共有、関連カードによる思いがけない情報との遭遇などが観察された [7],[8]。

(3) FISH の性能

検索／表示コマンドともにリストアップされるカードが20件以下ならば、1秒以下の応答時間を確保している[5]。しかし、蓄積件数がさらに増加すれば、高速化が必須と考えられる。

(4) 利用動向

FISH にカードを積極的に登録するユーザは、電子掲示板へも積極的に投稿するユーザであり、電子掲示板がよく利用されているオフィスでは、FISH の親和性も高いと予想される[6]。

(5) 蓄積情報

カード数の増加に伴いユーザが自由に付与するキーワード数も1次関数的に増加している。1つのカードは20行程度であり、電子掲示板に比較すると小さい単位にまとめられる。内容は、装置／ソフトウェア／仕事の処理手順(how)と各種連絡先(what)が多く、広範囲に渡って蓄積された[9]。

2. 4 FISH の課題

単一グループ対象とした利用実験では、継続した利用が観察されたが、マルチグループにおいてFISH の利用実験を展開するには、以下の技術的課題がある。

(1) マルチグループへの対応

単一グループを対象とする場合は、グループ、そこで蓄積される情報群、サーバは、それぞれ1対1に対応していた。しかし、マルチグループでは、図4の比較モデルに示すように、以下の状況が発生する。

【状況1：情報群とグループ】

グループと情報群が1対1に対応しない。グループ内における情報群の分割や、複数グループに共通な情報群が存在する。ここでは、有効範囲／ジャンル／用途等が同一な情報の集団を情報群と呼ぶことにする。

【状況2：グループと個人】

複数グループに属する者が存在する。組織／地域／職種等、グループ化のための分類基準が複数存在する。

【状況3：グループとサーバ】

グループとサーバが1対1に対応しない。また、クライアントとサーバの配置は、ネットワーク環境に依存する。

このような3つの状況に対応する必要がある。

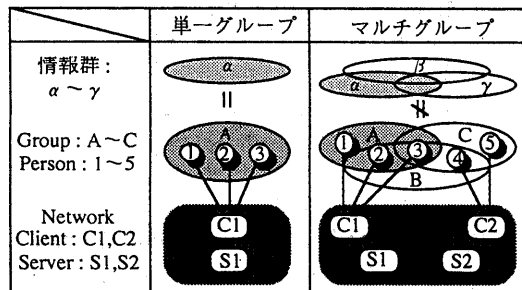


図4 単一／マルチグループの比較モデル

(2) 高速／大容量化

マルチグループにおける利用では、情報群の複数化に伴う情報量の増大と、ユーザ数の増加が予想される。また、ネットワーク環境に依存して生じる処理の集中や、分散化に伴う処理時間の増加等も考えられる。FISH は拡張性とメンテナンス容易性を重視し、インデックス等をテキストファイルで保持していたため性能の限界が低く、今後、システム規模に見合った処理の高速化を要する。

(3) カスタマイズ性の確保

単一グループを対象とした利用実験では、ユーザの要望に応じて、FISH にユーティリティ機能を追加してきた。しかし、複数のグループで利用実験を展開するためには、予めいくつか情報の管理形式や検索手段を備え、それぞれの利用形態に応じて機能を選択可能とするカスタマイズ性が必要である。

3. GoldFISH の研究展開

多種多様な情報群の中から、必要とする情報を検索する手段の研究が行われてきたが、ユーザ要求の同定等に困難な問題を残している[10]。FISH は情報検索システムでもあるが、同時にボトムアップ的に情報を蓄積するシステムでもあり、

情報登録時にその有効範囲／ジャンル／用途等に
 応じた情報群を予め形成することが可能である。
 そこで、GoldFISH は複数の情報群を設定し、マ
 ルチグループに対応するとともに、既存のネット
 ワーク環境に依存しないシステムを目指す。検討
 ポイントを以下に示す。

(1) マルチグループへの対応

情報の有効範囲／ジャンル／用途等が同一な情
 報群、複雑なグループ構成、グループ構成と独立
 に構築されるネットワーク環境の3層に対応でき
 るサーバを実現する。

(2) 高速／大容量化

情報の集中／分散に柔軟に対応できるサーバの
 高速／大容量化を図り、1つの情報群につき、1
 万件の情報を蓄積した状態で、1秒以下の応答時
 間を目標とする [11]。

(3) カスタマイズ性の確保

複数の利用形態に対応できるように、カスタマイ
 ズ性を確保する。想定する利用形態としては、ソ
 フトウェア開発現場における仕様書／連絡票管理、
 営業窓口における顧客情報管理、ソフトウェア/
 ハードウェア製品の障害状況管理等がある。

4. GoldFISH で採用した方式

研究展開で示した3点について、GoldFISH で
 採用した方式と検討の過程を以下に述べる。

4. 1 マルチグループへの対応

(1) サーバの構成

FISH の課題で述べた3つの状況（情報群とグ
 ループ、グループと個人、グループとサーバ）に
 対応し、グループ構成、及び情報の有効範囲／ジ
 ャナル／用途等に応じた情報のグルーピングを可
 能とするためには、以下の2つの前提条件のもとに、
 3つの機能を満たす必要がある。

【前提1】

機能の効率的分担、システムの拡張性を考慮し、

クライアント／サーバ形式とする。

【前提2】

情報管理の簡略化のため、異なる情報群に蓄積
 されたカード間では、リンクを自動生成しない。

【機能1】

任意の情報群をシステム内に複数設定可能。

【機能2】

複数の情報群にまたがる検索が可能。

【機能3】

グループ単位の読出し／書込み制御が可能。

この3つの機能を満たすサーバの構成方式とし
 て、図5に示す3つが考えられる。

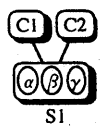
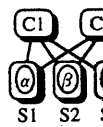
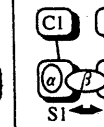
方式	A: 単一サーバ	B: サーバ独立	C: サーバ連携
Client: C1, C2 Server: S1 ~ S3 情報群: α ~ γ			
特徴	S内を情報群 別に分割	Cは情報を持つ Sにアクセス	S間で互いに 情報交換
機能1	S内に設定	S毎に設定	S間にも設定可
機能2	Sで実現	Cで実現	S間で実現
機能3	Sで実現	Sで実現	Sで実現
長所	情報再構築容易	Sの増設が容易	Cの負担が小
短所	規模に限界	Cの負担大	S間の管理複雑

図5 サーバ構成の方式比較

A: 単一サーバ方式

1つのサーバですべての情報群を管理する。処
 理の簡略化／機能追加が容易であるが、一点に
 処理が集中し、システムとしての高速／大容量
 化に限界が生じる。

B: サーバ独立方式

複数のサーバが独自の情報群を管理する。サー
 バにまたがる検索は、クライアントの判断で行
 う必要がある。

C: サーバ連携方式

Aのサーバを物理的に分割した方式。性能と容
 量的な限界は解決できるが、サーバ間の情報交
 換が必要となる。情報交換方式として、X.500
 で規定している分散ディレクトリモデル（連鎖
 /多方向展開／紹介） [12] などが適用できる。

システムを構築する上で、Aは記憶容量、検索

速度等の点で大規模化に限界があり、C は複数サーバ間の問い合わせ処理と、データ更新時の一貫性保証が複雑化し、オーバーヘッドが大きい。一方、B は、サーバの選択をクライアントに委ねているが、システム内に存在するサーバの一覧性を高めることで問題を解決できる。GoldFISH は、以下の条件を満たす B: サーバ独立方式を採用する。

- ・ 1つのサーバが1つの情報群を管理する。
- ・ 1つのワークステーション内に複数のサーバと複数のクライアントが共存できる。
- ・ サーバ毎にアクセス可能なグループを設定できる。

(2) サーバのアドレス情報管理

サーバ独立方式では、アクセスできるサーバの名前や通信アドレス等のアドレス情報を予めクライアントに教えておく必要がある。これらの情報を記述した Server Table をシステム内で管理する方式として、図6に示す3つが考えられる。

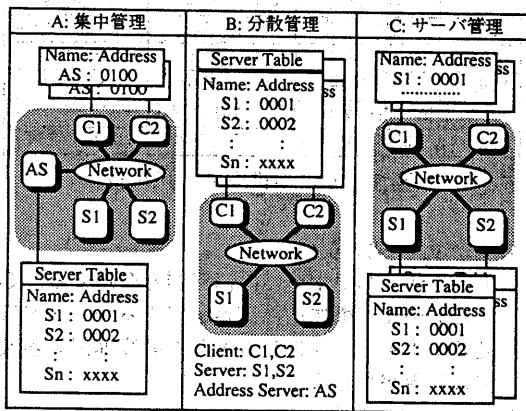


図6. アドレス情報管理の方式比較

A: 集中管理方式

Server Table を集中管理するアドレスサーバを設け、サーバ接続時にクライアントがアドレスサーバに問い合わせる。

B: 分散管理方式

クライアント毎に Server Table を管理する。管理方法には、Server Table の無矛盾性を確保する手順をもつ方法と、別ルートを使ったオフライン管理の方法が考えられる。

C: サーバ管理方式

サーバ毎に Server Table を管理する。クライアントは、Server Table のうち必要最低限な一部を保持する。

A はアドレス情報をアドレスサーバに問い合わせる必要がある。そのための時間と、通信トラフィックの増大、アドレスサーバへの負荷集中による応答性能の悪化が危惧される。B はアドレス情報に更新が生じた場合、すべてのクライアントの Server Table を更新する必要がある。そのため、すべてのクライアントのアドレス情報を管理する手段と、Server Table の変更を通知する手段が必要となる。これらは、処理を複雑化し、クライアントによっては不要なアドレス情報の管理に CPU 時間やメモリを浪費したり、更新ミスやシステムダウン等による Server Table の矛盾を引き起こす危険性がある。適用領域を考えると、アドレス情報が頻繁に更新され、その一貫性がシステム全体として重視される場合は A、アドレス情報の更新がほとんどなく、クライアントからサーバへのアクセスが局所的な場合には B が適する。

GoldFISH では、アドレス情報の更新頻度と、アクセスの局所性が A と B の中間に位置する環境での利用が主であると考えられ、以下に述べる C: サーバ管理方式を採用する。

【クライアントからサーバへのアクセス手順】

1. サーバ毎に、システム内に存在するすべてのサーバの名前（システム内で重複しない）とアドレス等が記された Server Table を管理する。
2. クライアントは、最低1つ以上のサーバのアドレス情報を保持する。
3. サーバは、要求に応じてクライアントに Server Table を通知する。
4. クライアントは、サーバの名前を指定してサーバにアクセスする。

【サーバを生成/変更/削除する際のアドレス情報の更新手順】

Server Table の変更の際には、サーバ間で対等な関係をとる方式 [13] もあるが、ここでは、図7に示すような原本を設定し、プロトコルの簡易な以下の手順をとる。

1. システム内のすべての Server Table の原本である Master Server Table を 1 つ設定する。
2. サーバに変更が生じた場合、当該サーバは、Master Server Table を管理するサーバに変更を申請する。
3. Master Server Table の変更結果は、その時点で起動している各サーバに通知される。
4. サーバは、起動時に Master Server Table の内容を自分の Server Table に書き写す。

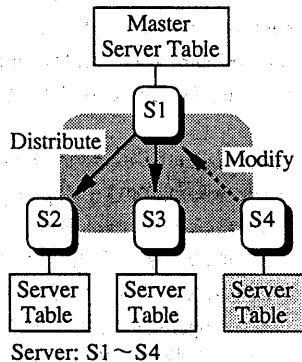


図7 Server Table の変更手順

(3) アクセス制御

グループ単位の読出し/書き込み制御を可能とするために、サーバ単位に以下の2つのリストを管理し、サーバへの不法なアクセスを制限する。

- ・アクセス可能なホスト名とユーザ名
(ワイルドカード使用可)
- ・ユーザ名、User ID とグループ名等の別名

User ID は、次に述べるユーザ名のコード化に利用し、グループ名等の別名は、ユーザ名へ再帰的に展開される。これらのアクセス制御は、グループの多様化と、サーバの増設容易性を考慮して、情報を実際に格納しているサーバ単位に独立して設定する。

4. 2 高速/容量化

(1) 高速化

ユーザ要求として、最も高速化が必要とされる処理は次の2つである。

- A: 検索処理実行時のリスト表示。
- B: 読出し処理実行時の内容表示。

A の検索処理と結果出力を高速化するため、インデックスデータと、検索結果として出力されるリストデータ (ユーザ名を含む) の2つをバイナリ/コード化し、メインメモリで管理する。バイナリ/コード化は、検索対象をコンパクトにし、アルゴリズム上の高速化が期待できる。また、一般にメインメモリのアクセス速度は、2次記憶のそれよりも3桁以上速いと言われており、メインメモリで管理することにより、IO上の高速化が期待できる。さらに、属性ファイルと情報ファイルを分離し、インデックスとして保持している以外の属性を用いた検索処理を高速化する。

B の読出し処理を高速化するため、サーバからクライアントへのファイル転送に ftp を利用し、信頼性の高い大容量ファイル転送を簡易に実現する。さらに、関連カードのリスト検索をバックグラウンドで処理することにより、見かけ上の表示開始時間を短縮する。もちろんこの検索処理には、先に述べた高速化も貢献する。

(2) 大容量化

多数の情報や、容量の大きな情報を扱うために、属性ファイルを情報ファイルと独立に管理する方式を採用し、大規模化/マルチメディア化に対応する。蓄積件数の増大と、マルチメディア化によるファイル容量増大への対応が必要であるが、前者に関しては、高速化で対応する。また、後者に関しては、テキストファイルと、その他の拡張ファイル (イメージ、音声等) を分離格納格納することにより、出力プログラムと記憶装置を最適化する。さらに、ハードリンクを多用することにより、複写等による記憶容量の増大を抑制する。

4. 3 カスタマイズ性の確保

マルチグループにおける用途の多様化 (例えば仕様書や連絡票管理等への応用) に対応するため、以下の属性パラメータを追加し、機能拡張を図るとともに、ユーザ定義可能な属性パラメータを設定可能とする。

- ・カードの輸出元サーバ名
サーバ間のカード移動，サーバの分割／統合を支援する。
- ・カードのバージョン
カードの修正履歴をバージョン管理し，オリジナルな情報を保持することにより，登録者以外のユーザによるカードの書き換えを可能とする。
- ・参照先／元カードの ID
議論や情報交換の記録を残すため，カード間の reply 関係を管理する。
- ・拡張ファイル名と再生手順
マルチメディア情報管理へ進展させるため，拡張ファイルを管理する。

5. おわりに

ノウハウ蓄積システム FISH を用いた 2 年間の利用実験の経験から，情報の有効範囲／ジャンル／用途等が同一な情報群，複雑な組織構成，組織構成と独立に構築されるネットワーク環境の 3 層に対応できる分散型ノウハウ蓄積システム GoldFISH の検討を行った。本論文では，マルチグループへの対応と，高速／大容量化，カスタマイズ性の確保に焦点を絞って，研究の展開方向と方式の選択について述べた。

サーバの構成方式検討においては，複数のサーバが独自の情報群を管理し，サーバにまたがる検索は，クライアントの判断で行う方式を採用した。サーバを情報群別に設け，情報管理の簡略化，高速／大容量化を図ったわけだが，今回，クライアントの判断に委ねたサーバの選択手段が，グループ対応に情報を再構築するポイントでもある。今後，システム規模や，適用業務に対応したサーバ選択の自動化を検討する必要がある。

謝辞

いつもご指導，激励頂いている NTT 通信網総合研究所 中田 寿 高機能ネットワーク研究グループリーダ，杉田 恵三 主幹員に厚く感謝いたします。また，ご討論頂いた清水 明宏 主任員，山上 俊彦 主任員をはじめ，研究グループの皆様

に深謝いたします。

参考文献

- [1] E. Jeffrey Conklin: Capturing Organizational Memory, Groupware'92, pp.133-137 (1992)
- [2] Lucy M. Berlin, Robin Jeffries, Vicki L. O'Day, Andreas Paepcke, and Cathleen Wharton: Where Did You Put It? Issues in the Design and Use of a Group Memory, INTERCHI '93, pp.23-30 (1993)
- [3] K.C. Burgess Yakemovic and E. Jeffrey Conklin: Report on a Development Project Use of an Issue-Based Information System, CSCW'90, pp.105-118 (1990)
- [4] Mark S. Ackerman and Thomas W. Malone: Answer Garden: A Tool for Growing Organizational Memory, COIS'90, pp.31-39 (1990)
- [5] 関良明, 山上俊彦, 清水明宏: ノウハウ蓄積システム FISH の実現とその評価, 信学論, J76 D-II, No.6, pp.1223-1231 (1993)
- [6] 関良明: ノウハウ蓄積システム FISH の利用動向分析, 信学技報, OFS92-7 (1992)
- [7] 山上俊彦, 関良明: 協調行動過程に着目したノウハウ支援の拡張の検討, 情処学マルチメディア通信と分散処理研報, 92-DPS-53-4 (1992)
- [8] 山上俊彦, 関良明: Knowledge-awareness 指向のノウハウ伝播支援環境: CATFISH, 情処学マルチメディア通信と分散処理研報, 93-DPS-59-8 (1993)
- [9] 関良明: ノウハウ蓄積システム FISH の蓄積情報分析, 信学技報, OFS92-2 (1993)
- [10] Peter W. Foltz, and Susan T. Dumais: Personalized Information Delivery: An Analysis of Information Filtering Methods, Communications of the ACM, Vol.35, No.12, pp. 51-60 (1992)
- [11] Ben Shneiderman: ユーザインタフェースの設計, 日経マグローヒル社 (1987)
- [12] JIS X 5731-1990 ディレクトリ
- [13] K.Ravindran, and Samuel T. Chanson: Application-specific Group Communications in Distributed Servers, IEEE INFOCOM '91, pp.1104-1113 (1991)