

WWW コラボレーションでの非同期データ更新

草場 匡宏 串田 高幸

日本アイ・ビー・エム東京基礎研究所

WWW を新しい情報交換の基盤とする WWW コラボレーションの研究が行われているが、インターネットでは帯域の保障がないために本質的に非同期のデータ更新を行なわなければならない。本論文は、WWW コラボレーションの現状と博物館教育への応用を述べるとともに、非同期のデータ更新に対応したユーザーシナリオとデータモデルを提案する。

1. はじめに

World Wide Web(WWW)の普及により、インターネットの上でマルチメディア情報を交換する基盤ができあがった。しかし WWW はデータの転送に伴うサーバーへの負荷を軽減するために HyperText Transfer Protocol (HTTP)を採用したために、サーバーはクライアントからのリクエストに対する一方向の情報伝達を行い、クライアントの状態を保持しないステートレスなプロトコルとなった。その結果、常にクライアントからのメッセージに対してサーバーの再接続を必要とするため、インタラクティブなアプリケーションを作ることが困難である[1]。また、WWW 上の文書は HyperText Markup Language(HTML)によって記述されるため、テキストやイメージの自由なレイアウトや、音声や動画などのマルチメディアを複合した文書を表示することが出来ず、Computer Supported Collaborative Workspace (CSCW)を実現するプラットフォームとして WWW を利用するには問題がある。そこでクライアントの情報を保持するために、サーバー側のアプリケーションはクライアントのフォームの入力を Common Gateway Interface(CGI)によって処理している。しかしこの方法では、クライアントは指定したフォーマットの情報しか与えないし、クライアントとサーバーが連動して変更しなければならない。したがってインタラクティブなアプリケーションの開発には向いているとは言えない[2]。

Asynchronous Update for WWW Collaboration,
Masahiro Kusaba, Takayuki Kushida,
IBM Research, Tokyo Research Laboratory

WWW コラボレーションとは、このように WWW という制約されたプラットフォームの上で、WWW を単に情報を引き出す基盤だけでなく、ユーザーが相互に情報を交換したり、その上で新しい情報を構築するための基盤として利用したり、コミュニケーションするための協調作業環境を示している。WWW コラボレーションはそのプロトコルの形態によって、同期型と非同期型に大別される。同期型コラボレーションは、WWW 以前より用いられてきた既存のネットワークコミュニケーションである電子メール・ニュース・メーリングリストをより直接的・対話的に行うために研究が行われている。一方非同期型コラボレーションは、それらとの連携や機能強化を意識しながら、会議システムとグループウェアでは、多数のユーザー間のコミュニケーションやワークフローの管理を主として研究や開発が進められてきた。しかし、WWW 上で収集した素材を整理・再構築してそれを他のユーザーと共有する協調作業空間(Shared Work Space)としての応用はまだ十分に行われていない。

我々は、博物館教育における教材の作成や、生徒同士の学習を支援するシステムを目的として、非同期型 WWW コラボレーションの研究を行っている。しかし非同期に協調作業によってコンテンツが作成される場合、同期システムとは異なりデータをどのように更新するか問題が生じる。同期システムの場合、コンテンツの一部があるユーザーによって変更された場合、その状態の変化を他のユーザーに通知し同期を行うことができる。また、もし同時に複数のユーザーが作業を行った場合、そのコンテンツに対する

アクセスを制限することで同期を行うことが可能である。しかし、WWW のような分散環境での非同期なシステムで同一のコンテンツに対する協調作業の場合、システムがリアルタイムに同期しないために、コンテンツの同期をとることが困難である。そこで、本稿は WWW コラボレーションの研究を概況を述べ、非同期コラボレーションにおけるデータ更新方法について検討する。

2. WWW コラボレーション

WWW コラボレーションに関するホームページは、非常に膨大であり、また日々更新されているのですべてを網羅するのは不可能であるが、ホームページに大半のものは紹介されている [URL1]。WWW コラボレーションは、同期的に情報の交換・共有を行うシステムと行わない非同期的なシステムに大別される。厳密に言えば、HTTP プロトコル上では、クライアントからのリクエストに対して即座にサーバーから応答が来ることは保証されないので、同期的なシステムはありえない。そこで以下の議論では、“同期的 (Synchronous)” を物理的な実時間で処理を行う実時間システムとそうではない非同期的 (Asynchronous) に処理を行うシステムとして定義する。また時間を時計で計測できる厳密な時間 (Clock Time or Hard Real Time) での同期を“同期的”、実時間 (Soft Real Time) での同期を“非同期的”と定義することもできる [3]。ここで同期と非同期によってどのような WWW コラボレーションシステムが実在するか示すと次の表ようになる。

同期・非同期による WWW コラボレーションの分類

	同期型	非同期型
テキスト	チャット	会議システム 電子掲示板 投票システム 情報検索システム
画像 & テキスト	共有黒板	グループウェア
音声	音声対話	
ビデオ	TV 会議	

2.1 同期型 WWW コラボレーション技術

同期型 WWW コラボレーションは、1 対 1、1 対多、多対多のユーザが WWW 上でリアルタイムのコミュニケーションを行う事を目的とする。このようなシステムでは、ネットワークの負荷による影響が少ない環境に限られ、またサーバーを経由してクライアント同士が直接相手クライアントとの接続状況をモニターするエージェントプログラムを WWW と並行して動かさなくてはならない。また、何らかの原因で通信中に接続が切断しても、再接続時にもとの状況を再現できるようなメカニズムが必要である。テキストやイメージに比べ、音声やビデオなどの連続的なメディアでは、ネットワークの負荷を考慮した Quality of Service (QoS) のメカニズムも必要であるが、現在市販されているインターネット用の音声・TV 会議システムでは十分な対策が取られていないため、実用的なシステムはまだ少ない。同期型の WWW コラボレーションでは、ネットワーク負荷の影響を受けやすいため、また参加できるユーザー数に制限があるために、実用的に利用されているシステムは少ない。今後、通信規格の制定が進み、また十分な回線が整備されるまでは、試験的な研究が必要である。

(1) チャットシステム

複数のユーザーがテキストメールを共有したり個別に交換するチャットシステムをベースにしている。このシステムでは、発言を管理するチャンネルを用意して、どのユーザーにたいして発言を許可するのか、また誰が発言したかを認識するユーザーのアクセス管理を行う。

代表的なシステムとしては、Internet Relay Chat (IRC) [URL2]、MS Internet Explorer ver 3.0 [URL3] などがある。

(2) 共有黒板

画像 (大抵はテキストも含んで) を扱うシステムは、1 対 1 あるいは 1 対多のユーザーがスクリーンにお互いが書き込める共有黒板をもち、ポ

インタを使って情報を書き込んだり、またスキャンした画像データなどを貼り込んで画像情報を伝達する。画像をリアルタイムに交換するネットワークフロー制御と複数のユーザーからのポインタの共有制御と表示を行っている。この分野では、CSCW の応用として TCP/IP 上で古くから研究が進められてきたが、帯域の確保が保証できない WWW 上では、パフォーマンスの観点から実装しているものなかった。しかし、T.120(126)によるデータ会議用の通信プロトコルが標準化されたため、実現が容易になった。代表的なシステムは、MS Internet Explorer ver 3.0[URL3]がある。

(3) ネットワーク音声対話システム

HTTP によって、データの送信・受信をコントロールすることにより、マイクから音声キャプチャカードによって、デジタル化した音声をファイルとして転送し、受け手のクライアントのスピーカーで再生する。それによって、音声情報をリアルタイムにクライアントに送出したり、あたかもインターネットで電話で行うように音声での対話を実現することが可能になる。(1)のチャットシステムと比べて、データを管理する専用のサーバーと複雑なネットワーク制御が必要だが、よりインタラクティブなユーザーインタフェースを提供することができる。(2)と同様に T.120 の勧告によって、実現が容易になった。代表的なシステムは、WebPhone [URL4]、CoolTalk on Netscape Navigator [URL5]などがある。

(4) TV 会議システム

音声だけでなくビデオや TV の動画像も制御するシステムである。クライアントには音声取り込み・再生カードのほかにビデオ取り込み・再生カードが必要になるが、基本構成は(3)と同等である。音声データは実時間で送信を重視するのに対し、動画像はフレームレートをコントロールして、ネットワークの負荷に応じて情報

量をコントロールする必要がある。(2)(3)と同様に T.120 によるデータ会議用の通信プロトコルが標準化され、WWW での開発基盤が整ってきた。しかし、T.120 では QoS 制御はまだ実現されていない。Internet Engineering Task Force(IETF)では RSVP や QoS 予約システムが検討されている。代表的なシステムは、neT.120 [URL6]である。なお同期型 WWW コラボレーションのシステムについての研究の概要については、[URL16]に概要が述べられる。

2.2 非同期型 WWW コラボレーション技術

一方、非同期型 WWW コラボレーションは、リアルタイム性を必要としないが、その代わりユーザー数は同期型よりはるかに多い。このようなシステムに情報を蓄積したり検索するために、非同期型 WWW コラボレーションは、データベースと密接な関係にある。したがって、非同期型 WWW コラボレーションは、複数のユーザーが自分の必要とする情報を知りたいときに引き出し、また必要に応じて入力したものをグループや個人で利用することを目的とするシステムである。

(1) 会議システム

複数のユーザーが、それぞれ固有の ID を持ち議題について意見を書き込んでゆくことで会議室を構成するシステムである。電子会議室は、すでにインターネットニュース(NNTP)やパソコン通信の中で実現されているが、これらは、テキストベースである、フォーマットが固定されている、ナビゲーション機能が弱い、コメントチェーンをサーバーが管理していない、などの点で劣っている。一方 WWW 会議システムでは、WWW による GUI とデータベースを活用して、関心のある議題をコメントツリーとして表示したり、会議室での位置づけを示すことにより、優れたナビゲーション機能を実現している。

代表的な製品については、Byte Article[URL7]に詳しいレビューが掲載されている。研究では、草分け的存在として、W3 Interactive Talk

(WIT) Project [URL8] が有名である。その他のシステムについては、Conferencing on the World Wide Web[URL9]を参照する。

(2) 電子掲示板

既存のニュースや電子会議室を WWW 上で実現したもので、(1)の電子会議室の機能を簡素化して、不特定多数のユーザーに対して自由に書き込みや読み出しができる掲示板あるいはメッセージボードの機能を提供する。製品としてではなく、フリーウェアとして配布されているものが多い。代表的なシステムは、HyperNews[URL10]があげられる。

(3) 投票システム

このシステムでは、さまざまな議題について、ユーザからの意見を投票によって集め、集計・分析した情報を再びユーザーに提供することを目的としている。ユーザーとサーバー間の情報の流れと、扱うデータが定型化しているため既存の CGI だけで対応できる場合が多く、研究対象になるものは少ない。代表的なシステムは、CNET's Online poll [URL11]にある。

(4) 情報検索システム

このシステムは、多様な WWW 上の情報から必要な情報を抽出・集積し、ユーザーからのリクエストに応じて情報を検索し提供する。サーバには、通常分散検索エンジンとデータベースを備え、コンテンツから必要な Index を抽出して登録することで、検索したい URL を高速に検索する。個人の WWW の利用が増えるに従って、ホームページの数も加速的に増え、情報検索サービスの必要性がますます高くなった。しかし、キーワードの全文検索からインデックスを作成するため、検索結果が期待されるものと異なる場合が多い。代表的なシステムは、Lycos[URL12]、Yahoo![URL13]、AltaVista Search[URL14]などがある。

(5) グループウェア

電子メール、電子掲示板、ワークフロー管理、グループスケジュール管理、電子会議、文書管理を一つのソフトウェアに統合したもので、共通の目的をもつグループの作業を支援するシステムである。WWW とは全く別の独自のプロトコルによって、メール管理、文書管理、セキュリティ管理、エージェント機能、アプリケーション開発機能を提供してきた。またロータスノートに代表されるように WWW と相互に情報をアクセス出来るシステムも登場している。しかし既存の WWW ブラウザーから情報を共有する仕組みはまだ十分ではない。代表的なシステムは Lotus Domino [URL15]がある。

3. WWW コラボレーションでのデータの更新

次にこれらの WWW コラボレーションにおける問題点について考える。通常同期システムの場合では、再接続までに要する時間がユーザーによる更新よりも小さいと考えられるため、最後に更新をかけたユーザーに対するアクセス権を優先して、他のアクセスに対しては更新に制限を加えながら、更新の履歴を管理する方法がとられる。すべての接続が復旧した時点で、履歴に基づき更新を加えて接続がきれる直前の状況(バージョン)まで再現する。

通信量の制限が行われていないインターネット上では、一般的に同期通信を保証するメカニズムがない。そのため通信量の増大による通信速度の低下や、サーバーの障害や、サーバー間でのネットワークの障害によって通信不可能な状態、あるいは非常に速度が低下してユーザーに対する応答時間が非常に大きくなって同期が困難な状態になる可能性がある。したがって次に接続が復活するまでの間各サイトで独立して更新を管理することが必要になる。WWW 上でのコラボレーションを行うには、システムは本質的に非同期の協調作業空間を前提にする必要がある。

WWW 上に非同期な分散協調作業空間を構築

するにはデータ更新時の問題を明らかにする必要がある。非同期型コラボレーションでも、サーバーが接続すれば、同期型コラボレーションと同等の同期のメカニズムを考慮しなければならない。以下の議論では、コンテンツをWWWで表現される情報の集合とし、またオブジェクトはコンテンツを構成する最小単位の構成物とする。まず、ユーザーのアクセス権によって更新の優先度を決定するメカニズムが必要である。これは、UNIXのファイルシステムのようにアクセス権の高い順にユーザーによる更新を制御する方法である。このメカニズムは、教師と生徒、グループとグループリーダーという上下関係のあるコラボレーションを実現するのに適している。アクセス権の低いユーザーに対しては更新ではなく読込のみ許可を与えるなどアクセス方法の制御を行う方法も有効となる。次に、それぞれのユーザーに対して更新の履歴を管理するメカニズムが必要である。同一の書き込みが可能なコンテンツに対して、複数のユーザーがアクセスして更新処理を行った場合、複数のバージョンができるためその同期処理を行う前に、どのようなバージョンが存在するのか調べて管理しなければならない。その後、アクセス権や更新時間などの基準にしたがってバージョン間の整理を行い、最新バージョンを生成する。以上のメカニズムは、同期の協調分散システムでも、同期の時間が存在する限りにおいて解決しなければならない問題点である[3]。しかし、非同期な協調分散システムにおいて、同一のアクセス権を持ったユーザーによる更新が行われた場合には、更新履歴の情報が保証されないために、どのような基準で更新するか基準を定める必要がある。この基準はコンテンツを管理する者あるいは、製作者がグループ全体によって決定される。すなわち、対象となるコンテンツの管理者あるいは、そのグループ内での同期に対する方針に基づき更新の管理を行うメカニズムを用意しなければならない。

4. WWW コラボレーションと博物館教育

我々は、博物館教育における教材の作成や、生徒同士の学習を支援するシステムを目的として、非同期型WWWコラボレーションの研究を行っている。博物館に蓄積される情報は、非常に多様なデータ構造を持ち、ホームページのようなファイル単位のハイパーテキストであらわすのは困難である。なぜなら、画像やテキストなどのデータは、単体として意味を持たず、必要時の状況に応じて関連するデータを組み合わせる一つの意味のある情報を形成できる。したがって博物館教育では、取り出したさまざまな情報を組み合わせ複数のユーザーが相互に共有する協調型ハイパーメディアが必要である[5]。博物館教育では、

- (1) 研究者による素材の収集
- (2) 学芸員・教師による教材の作成
- (3) 生徒への教材の配布
- (4) 生徒による教材の実習
- (5) 生徒による博物館の探索
- (6) 生徒同士の意見交換
- (7) 生徒による課題の解答
- (8) 生徒の課題の提出
- (9) 教師からの解答へのコメント

などの作業が考えられるが、非同期型コラボレーションとして考えた場合、(6)での複数のユーザーによる非同期なコンテンツのアクセスとアップデートを矛盾なく行うには、コンテンツ(ホームページ)単位のデータ管理ではなく、その中に含まれる多様なオブジェクトに対する細かいバージョン管理が必要である。

5. ユーザーシナリオ

博物館教育におけるコンテンツの更新について、代表的なシナリオを考えると次のようになる。(a)教材の配布(b)個人学習(c)グループ学習(d)クラスルーム
(a)のシナリオは、学校教育で先生が教材をつくって、配布する場合がこれに相当する(図1)。アクセス権の高いユーザー(先生)に対して更新

を許可し、通常のユーザー(生徒)に対しては、読み込みのみ可能な制限をもうける。更新可能なアクセス権は、コンテンツのオーナーやグループのリーダーなどの限られたユーザーに限定しなければならない。また、更新されることが明らかな場合、通常のユーザーに対して読み込みの操作を禁止することによって更新によるコンフリクトを避けることができる。

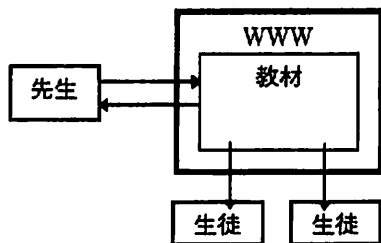


図1 教材の配布

次に(b)のシナリオは、生徒が自己学習のために教材や博物館の素材を取込み・加工して新しいコンテンツを作る場合に相当する(図2)。一度取込んだコンテンツは、アクセス権の低いユーザー(生徒)がコンテンツのアクセス権を意識せずに更新をかけることができる。ユーザー(生徒)がコンテンツにアクセスするとシステムは、ユーザーの状態に関係なくコンテンツの複製を作り、もとのコンテンツに対しリンクをはって関係づける。リンクをはられたコンテンツは新しくバージョンがつけられ、更新がかけられる。

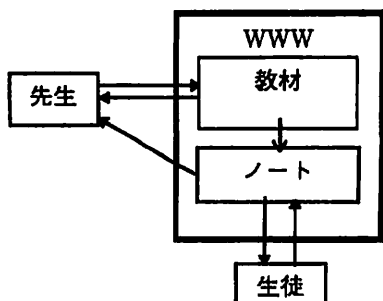


図2 個人学習

(c)のシナリオは、グループで議論をしながらコ

ンテンツを作ってゆく場合に相当する(図3)。共有されたコンテンツに対して一人のユーザーだけが更新することができ、更新が完了するまでは他のユーザーは読み込みだけのアクセスしか許可されない。このシナリオでは、同期・非同期に関わらず、サーバーとすべてのユーザー間のネットワークが十分高速でない限り成り立たない。もしそうでなければ、ユーザーはひたすら更新の許可が下りるのを待つことになる。

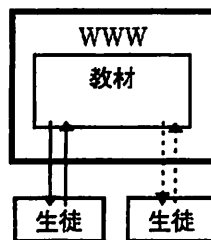


図3 グループ学習

最後に(d)のシナリオは、クラスルームで与えられた課題に対して、生徒が自由に議論を行う場合に相当する(図4)。このシナリオでは、非同期システムで一般的な解決の難しい問題を含んでいる。なぜなら、ユーザーのアクセス権が同じため、ユーザーの更新要求が完全に衝突する可能性があるからである。したがって、どのような場合にどのように更新するのか基準を設ける必要がある。

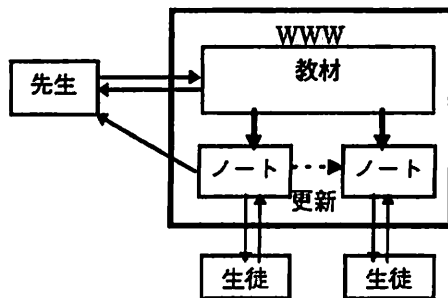


図4 クラスルーム

まず、更新と通知の時間をならべて一番古い順

にバージョンをつける。一定の時間間隔で、更新履歴の古い順に更新基準を参照し更新を行っていく。我々は、博物館教育における更新基準として次の5つの更新基準を採用している。

(1)通知時間

ユーザーがロードしたオブジェクトの日時とクライアントからサーバーに更新の命令が通知された日時を比較して最新のものを採用する。もっとも直接的な更新基準であるが、非同期コラボレーションでは更新の内容の整合性は保証されない場合がある。そこで通常、更新時間と組み合わせ使用。例えば、宿題やレポートの提出など期限がある作業を行う場合に有効である。

(2)更新時間

ユーザーがクライアント上にあるオブジェクト・コンテンツに更新を行った日時とサーバー上のコンテンツの日時を比較して最新のものを採用する。(1)に比較してネットワークに依存しないので、時間的な制限の少ない協調作業を行うのに適している。

(3)ユーザー数

共有しているユーザーの数によって更新の可否を判断する。議論をまとめクラス全体の意見を統一する場合には、数が基準値を越えた場合に更新を許可する。

(4)ユーザー許可

コンテンツの管理者や作成者は常に更新を許可されるが、他のユーザーは参照しかできない。そこで特定のユーザーやグループに対して更新を許可する。これによって、クラスの中に小人数のグループを作って議論をすることが出来る。

(5)複製許可

オリジナルに対して複製を許可し、かならず複製に対して更新を行う。複製はオリジナルのオ

ブジェクトに対して関係付けられ、ユーザー毎に履歴が管理される。この基準はもっとも緩い同期に相当し、更新時の衝突は起こり得ないが、バージョン管理が困難になる。生徒個人の意見やレポートを求める場合に採用する。

現実のプロセスは非常に複雑であるが、これらの更新基準に対して順位をつけたり、また組み合わせることによって、目的とする協調作業にあった更新を行う必要がある。

また、更新の対象は、オブジェクトの属性によってさまざまな場合がありうる。例えば、オブジェクトがクライアントのスクリーン上にイメージやアイコンで表示されると仮定する。画像オブジェクトに対して、複数のユーザーが更新を行った場合でも、画像事体を内容を更新する場合と、その位置や大きさだけを更新する場合がある。博物館教育で教材で生徒同士で地図上である素材と関連する国家を結び付ける作業を考えると、地図や素材のアイコンや画像は場所を移動することはあっても内容を変更することはできない。したがって、オブジェクトに対する共有・更新の制御だけでなく表示や表現に対する制御が必要である。

6. データモデル

このようなユーザーシナリオを実現するために、図5のようなデータ構造を採用した。各オブジェクトは、リソース属性(file、URL、名前、作成情報などオブジェクトの内容を示すもの)・ステータス(オブジェクトの共有状態)・リンク属性(他のオブジェクト、表示オブジェクトへの関連情報)の他に履歴属性を持つ。履歴属性は、リソース属性と別に更新のためのユーザー情報と日時を管理する。更新基準とリンク属性以外は基本的にシステムによって管理され、オーナー以外のユーザーは変更することが出来ないようになっている。

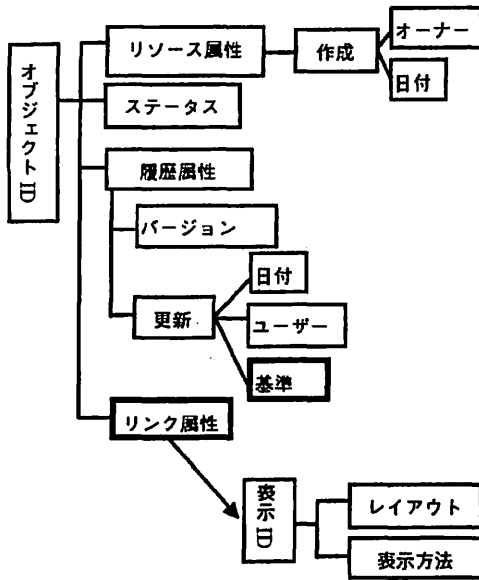


図5 非同期コラボレーションのデータモデル

7. 今後の課題

複数のユーザーによる WWW での非同期なコンテンツのアクセスとアップデートを矛盾なく行うには、コンテンツ(ホームページ)単位のデータ管理ではなく、その中に含まれる多様なオブジェクトに対する細かいバージョン管理が必要である。本稿では、WWW コラボレーションの位置づけと問題点を示した上で、博物館教育を目的とした非同期なデータ更新時のプロセスとデータ構造について考察したが、プロトタイプを通じてその実効性を確認する必要がある。

また Java によって、WWW 上でプラットフォームに依存しないプログラムがブラウザ上から実行できるようになり、インタラクティブで自由度の高いデータ表現が可能となり優れたユーザーインターフェースを提供できるようになった。今後はプロトタイプの開発・試験を通してユーザーインターフェースでの問題点もあきらかにしてゆきたい。

リファレンス

[1] Sankar Virdhagriswaran, Shared Information Space: An Interactive, Collaborative System

Ennoblement Perspective, Workshop on WWW and Collaboration, Sep. 11-12, 1995

[2] Martin Gleeson and Tina Westaway, Beyond Hypertext: Using the WWW for Interactive Applications, 1995, AusWeb95 Conf.

[3] Flaviu Christian, Synchronous and Asynchronous Group Communication, Communications of the ACM, 1996 Apr. Volume 39 No.4 p.88-97

[4] John L.Schnase, et. al., The StudySpace Project: Collaborative Hypermedia in Nomadic Computing Environments, Communications of the ACM, 1995 Aug. Volume 38 No.8 p.72-73

[5] Kaj Gronbeak, et. al., Systems: A Dexter-Based Architecture, Communications of the ACM, 1994 Aug. Volume 37 No.2 p.65-74

[URL1] <http://union.ncsa.uiuc.edu/HyperNews/get/www/collaboration.html>

[URL2] <http://www.kei.com/irc.html>

[URL3] <http://www.microsoft.com>

[URL4] <http://www.jabra.com/html/webphone/webphone.html>

[URL5] <http://live.netscape.com>

[URL6] <http://www.databeam.com/Products/net.120/prodinfo.html>

[URL7] <http://www.byte.com/art/9606/sec10/art1.htm>

[URL8] <http://www.w3.org/pub/WWW/WIT/>

[URL9] <http://freenet.msp.mn.us/people/drwool/webconf.html>

[URL10] <http://union.ncsa.uiuc.edu/HyperNews/get/hypernews.html>

[URL11] <http://www.cnet.com/Forum/Polls/index.html>

[URL12] <http://www.lycos.com/>

[URL13] <http://www.yahoo.com/>

[URL14] <http://www.altavista.digital.com/>

[URL15] <http://domino.lotus.com/>

[URL16] <http://union.ncsa.uiuc.edu/HyperNews/get/www/collab/synchronous.html>