

WWWによるグラフィカル情報サービスの構築

—WWWのユーザ行動モデル—

新井 克也, 坂本 啓, 中畝 弘, 桑名 栄二

NTT ソフトウェア研究所

広大な情報空間を構成する、今日のコンピュータネットワークによる情報流通サービスでは、いかにしてユーザを特定のサービスに導くか、いかにして利用者が望むサービスを開発するか、といったサービスプロモーション、サービス構成技術が重要である。特にCG作品等の視覚的な情報を扱うサービスでは、サービスに要求される特性が、利用者の目的や嗜好によって多岐にわたり、全利用者の要求を一元的に満足する代わりに、個々の利用者の要求を分析し、これを解決する技術が求められている。本稿では、WWWサービスにマーケティング理論のユーザ行動モデルを適用することを試み、「WWWユーザ行動モデル」を定義した。そして、実際の利用者行動が、このモデルに適合するとの実験結果から、ユーザ行動モデルの適用可能性の見通しを得た。また、この実験結果に基づき、WWWサーバにおけるサービスプロモーションの方策についても考察した。

1. はじめに

World Wide Web (WWW) は、インターネット上に分散化した情報をハイパーテキストによってリンクし、統一的なインタフェースでアクセスすることを可能にする技術である[1]。WWWの利用者(ユーザ)は、WWWブラウザと呼ばれるプログラムを用いて、このリンクをたどることで情報空間を移動し、適切な情報を参照する。WWWで提供される情報の単位は、「ページ」と呼ばれる。このページとページをリンクで接続することにより、ハイパーテキストが構成される。WWWが提供するサービス(WWWサービス)とは、このリンクされたページの集合を提供することであり、インターネットに接続したコンピュータにおいて、WWWサービスプログラムを実行することで実現できる。

今日のインターネットにおいて、WWWサービスが急速に増加した背景には、インターネット接続サービスの商用化とWWWサーバプログラムの流通がある。これらによって、インターネットの利用者が、WWWサーバを運用し、情報を発信するサービス提供者となる機会が増加した。すなわち、WWWは、サービス提供者とその利用者の差が小さい、フラットなサービス環境であり、個人の嗜好や情報を発信する手段としての利点は大きい。

ところで、WWWサービスをユーザの視点で見た場合、WWWサービスの増加によって、広大な情報空間が構成された結果、欲しい情報が見つからない、あるいは、情報空間上の自分のいる位置が分からない、いわゆる「迷子問題(Lost in Hyperspace)」といった問題点が指摘されている。これらの問題に対する研究としては、ハイパーテキスト上のナビゲーション技術[2]や、インターネットにおける情報検索・フィルタリング技術の研究[3,4]がある。

一方、WWWサービスをサービス提供者の視点で見た場合、ユーザをいかにして自分のWWWサービスに導くか、あるいは、いかにして、利用者が望むサービスを提供するか、といったマーケティング、プロモーションの方法論が期待されている[5]。特に、等質の製品を多数の購買者に販売することを目的とした従来の流通とは異なり、WWWサーバによる情報流通では、ユーザそれぞれの特性に応じた、きめ細かいマーケティングの可能性もある[6,7]。たとえば、単一のサービスメニューを提供する代わりに、個々のユーザの特性によって適応的に決定される「お勧めメニュー」を提供するインタラクティブサービス[7]などが、その例である。このようなサービスを実現しようとした場合、1) ユーザのモデル化、2) 特徴パラメータの抽出、3) 特徴パラメー

タをサービス構成へフィードバックする、といった技術が必要となる。しかし、現在のところ、確立された方法論は存在せず、サービス提供者のノウハウによるところが大きい。

本稿では、電子的な情報流通サービスにおけるサービスプロモーション技術、サービス構成技術を確立することを目的として、サービス利用者に対して、マーケティング理論のユーザ行動モデルを適用することを試みる。以下では、電子的な情報流通サービスの1つとして、WWW サービスを取り上げ、WWW サーバのプロモーションについての考察、ユーザ行動モデルの定義を行う。そして、実際のユーザ行動の観測結果から、ユーザ行動モデルの適用可能性について議論する。また、ユーザ行動に関する議論に基づき、サービスプロモーションの方策についても考察する。

2. WWW サーバにおけるプロモーション

2.1. サービスの特徴

WWW サービスで提供される情報の単位はページである。ページは、他のページとリンクによって関係づけられている。ユーザは、このリンクをたどって、ページからページへ移動（ブラウジング）し、適切な情報を利用する。1 ページの大きさには、特に制限はない。しかし、一般には、1 ページが1 画面程度の、比較的小さな情報単位に分割されている。【特徴1】 ページ単位のブラウジングは、1 ページごとにユーザの判断（次のリンクの選択）と行動（ブラウジング）が必要である。すなわち、比較的小さな情報単位ごとに、ユーザ行動の観測機会が得られるという特徴がある。

【特徴2】 1つのページからリンクするページ数の制限や、ページ間の関係（連結性や閉路等）に関する制限はない。したがって、情報の構造化に関する自由度が大きい。たとえば、1つのページから他の全てのページをリンクすることも、あるいは、1ページずつ順次リンクすることもできる。

【特徴3】 ユーザからのアクセスに応答して、WWW サーバ上でプログラムを実行し、その結果を

新たな提供情報とすることができる。この機能によって、事前に内容が用意された静的なページだけでなく、ページの内容や他のページへのリンク等をユーザのアクセス時にリアルタイムで生成する、いわゆる動的なページが構成できる。

以上3つの特徴から、ユーザの特性を検出し（特徴1）、これに適したページ・リンク（特徴2）やページ内容の変更（特徴3）をアクセス時に適応的に実現する（特徴3）といった、きめ細かいプロモーションの可能性がある。

2.2. サービス構成の方法論

上で述べた、きめ細かいプロモーションを具体化するためには、1) ユーザのアクセス行動の観測から、そのユーザの将来の行動を予測する技術と、2) この予測に基づき、このユーザに適したサービス構成を動的に生成する技術が必要である。これらの技術のうち、ここでは前者の技術について議論する。

ユーザの行動を予測する単純なユーザモデルとしては、例えば、あらかじめ全てのページをいくつかのカテゴリに分類しておき、ユーザがよくアクセスしたページのカテゴリをそのユーザの特性を示す「好み」とする方法が考えられる。

しかし、この「好み」は、ユーザがWWW サービスを利用する時のコンテキストに大きく依存する。例えば、CG 作家の情報を集めたライブラリを利用する場合、CG 作品を楽しむ場面では、自分の好みの作家のCG 作品だけを見たいと考えるだろう。一方、作家の出身地とその作品の特徴の関係を研究する場面では、全ての作品と作家の出身地情報にアクセスするだろう。また、「好み」は、ユーザの経験や学習によっても変化する。例えば、好みの筋書の小説もたくさん読めば、やがて飽きてしまうだろう。

このように、WWW サービスのプロモーション実現のためには、「好み」によるモデル化では不十分であり、利用時コンテキストやユーザの学習的変化をアクセス時に実時間でシュミレートする様な、より詳細なユーザのモデル化が必要となる。

サーバ理解は知覚構造の刺激感受性に影響を与え、今後の（当該サーバに対する）刺激－反応の特性に影響を与える。

3.3. ユーザのアクセス行動ループ

モデルにおいて、ユーザのアクセス行動の経路は、つぎの2つに大別できる。

即時的行動ループ：

ユーザは例えば新着サーバ情報を得て（刺激）、そのサーバへアクセスする事を決定し（意図）、アクセス結果から、自分の好みに合っているか、意図した通りの情報であったか等の主観的判断を行い（評価）、そのWWWサーバの特徴を理解する（サーバ理解）。そして、この知識から、さらにリンクをたどって別のページへアクセスするかどうかを決定する（意図）。

このように、意図－アクセス－評価－サーバ理解の関係を巡るループは、あるWWWサーバの1ページへアクセスするユーザ行動を表している。ユーザは何度かこのループを巡り、「アクセスすることをやめる」という決定をした（意図）時点で、この一連のループが終了する。この連続したループの回数は、ユーザが一時に利用したページ数となる。

以下では、このループを即時的行動ループ（Immediate Decision Loop: IDL）と呼び、連続したIDLをIDL系列と呼ぶ。

後天的行動ループ：

ユーザは、ページへのアクセスとその評価から、そのサーバの特徴を理解し（サーバ理解）、例えば、「このサーバのサービスは自分好みであり、提供予定の新サービスにも期待できる」といった、そのサーバが持つ潜在的性質に関する確信を持つ（サーバ態度）。このユーザは、後日、新サービス開始の情報（刺激）を得たとき、先の確信（サーバ態度）の影響によって（刺激感受性）、再び当該サーバへアクセスする（意図、アクセス）。

このように、IDL－サーバ態度－刺激感受性の関係を巡るループは、アクセスした経験を有するWWWサーバへ、再びアクセスするユーザ行動を表しており、このループを回る回数が、WWWサーバへの再来回数となる。

以下では、このループを後天的行動ループ（Learned Decision Loop: LDL）と呼ぶ。

3.4. ユーザ行動モデルの適用方策

以上で述べたモデルは、個々のユーザの内的状態変化をシミュレートすることができる。つまり、観測可能な入力（プロモーション情報等）と出力（ページへのアクセス）から、ユーザの内部状態を推定する機能を持つ。この機能に加えて、ユーザの内部状態に適したサービスを自動的に生成する機構を用意できれば、個々のユーザに最適な個別サービスを提供する、個別サービスプロモーションが実現できる。

また、複数のユーザの内部状態を総合し、このユーザ集合が持つ特性や傾向を判断することも可能である。この手法を利用して、より多くのユーザに適した汎用的サービスを提供する、包括的サービスプロモーションが実現できる。

4. oneZero サーバのユーザ行動

以下では、ここまで議論してきたWWWユーザ行動モデルを実際のWWWサービスに適用する試みについて述べる。

4.1. oneZero サーバ

oneZero サーバ（以下、oneZero）[11] は、CG作家の任意団体「デジタル・イメージ」の協力を得て、インターネット上に公開しているWWWサーバであり、CG作品や展覧会の情報など、CGに関する情報サービスを行っている。

oneZero で提供するサービスの特徴は、グラフィカルな情報を主体としている点にある。特に、CG作品だけでなく、画像と文字の配置等のページデザイン自体も、提供する情報の一部であると考えている。したがって、ユーザが利用するカラーパレット、画面サイズ、文字フォント等、個々のユーザの環境構成に影響を受けやすいサービスであると言える。また、オフィスと家庭のいずれからでもWWWサービスを利用するユーザが、どちらでも同一の環境を利用しているとは限らない。このためoneZeroでは、動的に変化する、ユーザの利用時環境それぞれに適した情報の表現方法を採用する必要がある。

また、oneZeroのサービスは、CG作品を鑑賞して楽しむ、いわゆるイメージを伝えるサービスとしての側面と、CG作家の作品特性や経歴等を検索する、いわゆる情報を伝えるサービスとしての側面を持っている。したがって、2.2でも述べたとおり、ユーザの利用時コンテキストによって、それに適した特性を持つサービスを採用する必要もある。

以上のような特徴から、oneZeroでは、ユーザの利用時の状態を把握し、行動を予測する、WWWユーザ行動モデルの適用と、これに基づくサービスプロモーションの実現が期待されている。

4.2. サービス構造

oneZeroが提供するページ間の関係（サービス構造）を図2に示す。図において、矢印で結ばれたページは、矢印が示す方向へリンクされている。並置されたページ集合とは、双方向にリンクされているページの集合である。

Rootは、oneZeroへのアクセスポイントである。oneZeroでは、Rootの下に、3つの下層構造（以下、サービスオプション）、OP1、OP3、OP4を用意している。これらは、いずれもサービス構造が同一である。OP1とOP3は、いずれも、CG画像やボタンの配置、背景や文字の色や形等をデザイナーが総合的にデザインしたページである。また、これらのページを利用するためには、このデザインを表示する機能を持つ、特定のWWWブラウザを利用する必要がある。一方、OP4は、ページデザインに特別な配慮が施されておらず、どのWWWブラウザでも利用できる汎用的サービスである。また、CG画像等も小さいサイズに抑えられ、転送されるデータ量がOP1やOP3に比べて小さいという特徴を持つ。

3つのサービスオプションは、いずれも4つのサービス（以下、サービス要素）へのリンクを持つ。そのうちの1つである作家目次ページ（Digital Artists Library）は、約30人のCG作家の名前と作品の一部が一览できる。そして、個々の作家項目からは、各作家の個人ページ（Artists）がリンクされている。この他に、「デジタル・イメージ」に関する情報ページやoneZeroの紹介ページ等のサービス要素がある。

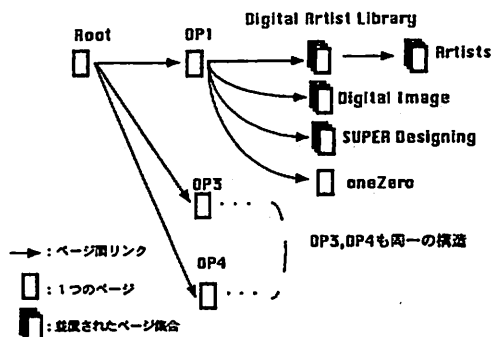


図2：oneZeroサーバのサービス構造

4.3. ユーザ行動の検出

ユーザ行動の検出には、WWWサーバプログラムのアクセス履歴情報（以下、ログ）を利用した。ログでは、あるホストから1ページアクセスされる毎に、

ホスト名：アクセス時刻：ページ名

のデータの組（以下、ログエントリ）がアクセス時刻順に記録されている。以下では、このログの解析にともなう条件仮定と、ホストの同一性、連続アクセス、再来アクセスを定義する。

仮定1：1つのホストからアクセスするユーザは、同一人物、ただ一人である。

仮定2：あるユーザのIDL系列の最中に、別のユーザからのアクセスはない。

仮定1は、例えば代理サーバ等からのアクセスを考えると、明らかに不合理な条件である。しかし、ログから個々のユーザを識別することは、今のところ困難であるので、この条件を仮定することはやむをえない。また、仮定2は、一般には成り立たない条件であるが、今回扱うログに関しては、あらかじめ、結果に影響が無いことを確認した。必須な条件ではないが、以降の議論を簡単にするための仮定である。

【ホストの同一性】ホストが同一であるとは、ログエントリのホスト名フィールドの値が同一であることを言う。

【連続アクセス】連続アクセスとは、同一ホストか

らの連続したアクセスを指す。ログ上では、同一ホストが記録された、連続した、1つ以上のログエントリの集合を指す。ユーザ行動モデル上では、1つのIDL系列にあたる。

【再来アクセス】連続アクセスのうち、同一ホストの2つ以上の連続アクセスを指す。ユーザ行動モデル上では、1度の再来アクセスが1つのIDLにあたる。

以上の仮定、定義にもとづき、ログから個々のユーザのアクセス状況を抽出し、これをユーザ行動の観察結果とした。

5. ユーザ行動の観測結果と考察

5.1. 観測結果

観測を行った oneZero サービスの諸元を表1に示す。サービス開始時期は1995年6月であり、提供している総ページ数は、約600ページである。本稿で扱うログの対象範囲は、1995年6月から1995年9月の間の約100日間であった。また、ログを解析した結果、この期間にわたって、同一でないホスト数（ユニークホスト数）は約2000ホストであり、これらによって、延べ18000ページがアクセスされた。

まず、サービスオプションそれぞれに対して、サービス要素毎のアクセス総数を測定した。その結果を図3に示す。図において、各サービスオプションの積み上げグラフは、サービスオプションに属するアクセスの総数を1とした時の、各サービス要素へのアクセスの分布の割合を示している。

次に、サービスオプションごとに、連続アクセスの総数を計測した。そして、アクセス総数をこの連続アクセスの総数で割ることにより、1連続アクセスあたりの平均アクセス数を計算し、これを「平均滞在時間（単位：ページ数）」とした（表2）。

最後に、各サービスオプションのHomeへのアクセスのうち、再来アクセスの数とユニークホスト数を計測し、前者を後者で割ることにより、1ホストあたりの平均再来アクセス回数を計算し、これを「平均再来アクセス数（単位：回）」とした（表2）。

5.2. 結果の考察

4.2で述べたとおり、サービスオプションのうち、

表1：oneZeroサービスの諸元

サービス開始時期	1995年6月
サービス総ページ数	約600ページ
使用ログ対象期間	1995年6月から1995年9月 約100日間
ユニークホスト数	約2,000ホスト
総アクセスページ数	延べ約18,000ページ

OP1とOP3は、いずれもデザイナーによってデザインされたページを提供する。一方、OP4は、デザインに関して、なんら配慮されていないページを提供する。すなわち、OP1、OP3とOP4の相違点は、ページの品質に関する違いにあると言える。ところで、画像の品質を評価する方法としては、画素数、色の数やその配色等に基づき定量的データで示す方法がある。しかし、本稿では情報提供者側の意図と利用者側の主観的評価の関係を扱うものである。そこで、本稿では、デザイナー（情報提供者）が意図してデザインを施したページを「高品質」とし、デザイナーの主観による定性的特徴をパラメータとした。

また、OP1とOP3の相違点は、それぞれに必要なWWWブラウザが異なるという、要求されるユーザ環境の違いにある。

ユーザ行動モデルの適用可能性：

まず、WWWサービスに対するユーザ行動モデルの適用可能性について議論する。

表2において、平均滞在時間が大きいOP1、OP3は、OP4と比べて、平均再来回数においても大きい値を示した。WWWユーザ行動モデルによれば、平均滞在時間はIDL系列の長さを表している。また、平均再来回数は、IDLの数である。したがって、観測されたユーザ行動は、IDL系列におけるユーザの評価がIDLへ影響を与え、その結果としてIDLにおけるユーザ行動が変化するという、WWWユーザ行動モデルに合致する結果であった。

以上の考察から、WWWサービスへのユーザ行動モデルの適用可能性についての見通しが得られた

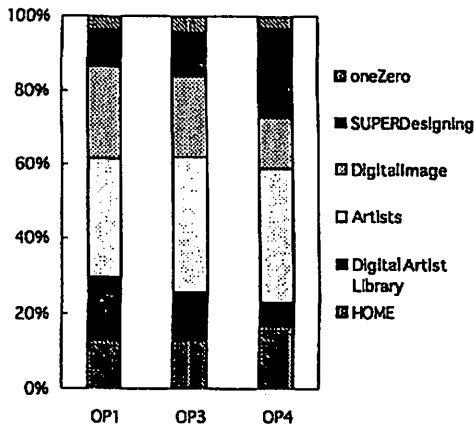


図3：サービスオプション毎のアクセス分布

と結論する。

ユーザ行動と影響パラメータ：

次に、図3と表2の観測結果をページ品質の観点から考察してみる。

アクセス分布に関しては、図3ではOP1、OP3とOP4の間に大きな差はみられず、サービス要素を構成するページ数の割合と同様の分布を示している。一方、表2からは、平均滞在時間と平均再来回数いずれも、OP4に比べてOP1、OP3の方が高い値を示していることが分かる。

また、同様にして、ユーザ環境の観点から考察すると、図3と表2の結果からは、OP1とOP3の大きな差は認められない。

以上から、今回比較実験を行った行動影響パラメータ（ページデザインの品質、WWWブラウザの種類）とユーザ行動の関係について、以下に示す考察結果を得る。

【考察1】高品質なページデザインは、多くのユーザから肯定的な評価を得る。また、サーバの潜在的特性としても肯定的な影響をあたえらるものである。

【考察2】高品質なページデザインは、サービス要素へのアクセス分布には、影響が少ない。

【考察3】使用するWWWブラウザの違いは、アクセス分布、平均滞在時間、平均再来回数のいずれに

表2：観測結果

サービス オプション	平均滞在時間 (ページ)	平均再来 アクセス数 (回)
OP1	3.9	4
OP3	4.0	5
OP4	1.0	2

も影響が少ない。

サービスプロモーション方策：

最後に、以上の考察に基づき、実験に利用したoneZeroのサービスに対して示唆されるプロモーション方策を考察する。

高品質なページデザインは、多くのユーザを連続アクセスさせるための手段として有効であった（考察1）。したがって、ユーザ・ナビゲーションの視点からは、サービス構造のうち、「見せたいページ」へ至るパスのページデザインを他よりも高品質なものとする事で、ユーザに選択の余地を残しながら、比較的ゆるやかにユーザを導くことができる。また、「より多くのアクセスを獲得する」という包括的プロモーションの立場からは、全てのページデザインを高品質化する方策が有効である。

ページの品質が再来アクセスを増加させる影響を持つという結果からは、多くのユーザのアクセスが集中するWWWサービスのエントリポイント付近には、高品質なページを配置し、より多くのユーザを再来アクセスに誘うというプロモーションも考えられる。

5.3. 今後の課題

上述の考察では、ページデザインの品質という定性的データによる比較を行った。しかし、このパラメータは、oneZeroのページデザインを担当したデザイナーによる主観的評価であり、上記の考察結果をただちに一般のWWWサービスへ適用することはできない。また、トレードオフの関係にあるパラメータなど、パラメータ間の関係も明らかになっていないという問題も残されている。今後は、客

観的な行動影響パラメータを複数抽出し、これらを用いた比較実験を行い、考察結果の検証を進めてゆく。

また、今回の実験では、ログから個々のユーザを識別することが困難であったために、ユーザ集合の行動を中心に考察した。つまり、3.4 で述べた包括的サービスプロモーションを主に検討した。今後は、モデルによるユーザ行動の予測技術と、そのユーザ行動に応じて、サービス構造や情報表現を適応的に変更する技術を統合し、個別サービスプロモーションを実現する技術に発展させて行きたい。

6. まとめ

本稿では、電子的な情報流通サービスに対して、マーケティング理論のユーザ行動モデルを適用することを試みた。その結果、WWW サービスの1つである oneZero サーバにおいて、「WWW ユーザ行動モデル」における即時的行動ループと後天的行動ループの因果関係が、実際のユーザ行動と合致するとの結果を得た。この結果から、電子的な情報流通サービスに対する、ユーザ行動モデルの適用可能性についての見通しを得たと結論する。また、以上の結果をふまえ、oneZero のサービスを改善するための、包括的サービスプロモーションの方策も示した。

電子的な情報流通に対してマーケティングや経済的モデルを適用する他の研究としては、情報流通の特性を経済的モデルで説明する試み[12]や、デジタル情報のための流通システムの研究[13]などがあるが、本稿と同様なアプローチをとる研究の絶対数は多くない。しかし、マーケティング理論、経済学、社会学等の研究成果の適用を図るこれらの研究は、今後の電子的な情報流通サービスにおいて不可欠なものであると考える。

今後は、本稿で述べたWWW ユーザ行動モデルを詳細化し、個々のユーザに適応的に対応する、個別サービスプロモーションの方法を明らかにして行くとともに、WWW サービス以外の電子的な情報流通サービスに対して、ユーザ行動モデルを適用することを検討してゆく。

謝辞

oneZero サーバの構築にあたり、ページのデザイ

ン、作品提供等のご協力をいただいた、長田智行代表をはじめとするデジタル・イメージ所属の作家、デザイナーの皆様に感謝します。また、クリエイタの立場から、ユーザモデルに関して議論していただいた、宮崎正弘氏、山田英治氏に感謝します。

参考文献

- [1] T.Berners-Lee, R.Cailliau, A.Luotone, H.F.Nielsen and A.Secret "The World-Wide Web", Comm. ACM, Vol.37, No.8, pp.76-82 (1994).
- [2] D.M.Edwards and L.Hardman, "Lost in Hyperspace: Cognitive Mapping and Navigation in a Hypertext Environment", R. McAleese編, Hypertext: theory into practice, Ablex, New Jersey (1989).
- [3] N.T.Bilkin and W.B.Croft, "Information Filtering and Information Retrieval: Two sides of the same coin?", Comm. ACM, Vol.35, No.12, pp.29-38 (1992).
- [4] P.Maes, "Agents that Reduce Work and Information Overload", Comm. ACM, Vol.37, No.7, pp.31-40 (1995).
- [5] R. Dedrick, "A Consumption Model for Targeted Electronic Advertising", IEEE MultiMedia, Vol.2, No.2, pp.41-49 (1995).
- [6] 中村雄三, "マルチメディア情報流通技術の動向", NTT R&D, Vol.44, No.10 (1995).
- [7] 本橋健, 荒野高志, 桑名栄二, "インタラクティブマルチメディアシステムの構成モデル", 情処第51回全国大会論文集, 1Q-7 (Sep,1995).
- [8] 田中由多加編, 新・マーケティング総論, 創成社, pp.62-83(1990).
- [9] J.A.Howard and J.N.Sheth, "The Theory of Buyer Behavior", John Wiley&Sons, (1996).
- [10] What's New, <http://www.ntt.jp/WHATSNEW/>.
- [11] oneZero, <http://onezero.sl.cae.ntt.jp/>.
- [12] 後藤滋樹, 野島久雄, "人間社会の情報流通における三段構造の分析", 人工知能学会誌, Vol.8, No.3, pp.348-356(1993).
- [13] 森亮一, 河原正治, "歴史的必然としての超流通", 超編集, 超流通, 超管理のアーキテクチャシンポジウム論文集, 情処学会, Vol.94, No.1, pp.67-76(1994).