

学習者に適した学習教材の推薦と配信

高橋 泰樹[†] 松澤 俊典[†] 山口 未来[†] 土肥 紳一^{††} 和田 雄次^{††}
東京電機大学大学院 情報環境学研究所[†] 〒270-1382 千葉県印西市武西学園台 2-1200
東京電機大学大学院 工学研究所[†] 〒101-8457 東京都千代田区神田錦町 2-2
東京電機大学 情報環境学部^{††} 〒270-1382 千葉県印西市武西学園台 2-1200
E-mail: {DEN03059,DEN03302}@nifty.com[†]

概要

近年、e-ラーニングが企業や高等教育機関等で積極的に取り入れられている。これらの e-ラーニングの普及の要因として、SCORM や WebCT などの LMS の発展があり、コンテンツの利用や運用が簡単に行えるようになった。その一方で、LMS を利用して配信されているコンテンツは、単純な制御で表示されることが多く、学習者が必要と考える最適なコンテンツが必ずしも配信されていない。このような状況から本研究では、コンテンツを利用した学習者の中から過去に同じような学習方法をしている、学習者の閲覧履歴情報を見つけ、もっとも類似性の高い学習者の閲覧履歴情報を活用し、最適と考えられるコンテンツを提示するシステム AIRS の構築を行った。本稿では、著者らが開発した AIRS の概要ならびに本システムについての有効性の評価について述べる。

キーワード e-ラーニング, 協調フィルタリング, AIRS

Recommendation and Delivery of the Optimum Texts for e-Learners

Yasuki Takahashi[†] Syunsuke Matuzawa[†] Mikuru Yamaguchi[†] Shinichi Dohi^{††} Yu'zi Wada^{††}
Graduate School of Information Environment, Tokyo Denki University[†]
Graduate School of Engineering, Tokyo Denki University[†]
School of Information Environment, Tokyo Denki University^{††}

Abstract

Recently, E-learning has been accepted in companies and educational institutions. The fact behind this is the development of LMS such as SCORM and WEBCT. LMS simplifies the utilizations and the operations in E-learning systems. However, the contents delivered method in LMS can only be operated in a simple way. This occurs a problem where the op possible content not being delivered to the e-Learners. From this circumstance, we have established AIRS, a system that finds students who used the same contents from historical logs and recommends optimum contents using Collaborative Filtering.

In this paper, we show the summary of AIRS and the evaluation of the effectiveness.

Key words e-Learning, Collaborative Filtering, AIRS

1. はじめに

インターネットを活用した e-ラーニング教育 [1]が、企業の社員教育や高等教育機関等のサポート的な役割として、積極的に取り入れられている。本学においても、ネットワークを利用した遠隔講義や講義のレポート提出を Web で行うなど、様々な場所で e-ラーニングが活用されている。こうした e-ラーニングの普及の要因として、ADL(Advanced Distributed Learning)が提唱している WBT(Web Based Training)の標準化規格の SCORM(Sharable Content Object Reference Model)の登場や WebCT などの LMS(Learning Management System)の発展 [1]が考えられる。これらの登場により、コンテンツの制作・再利用・流通が行いやすくなり、また、アップロードや公開

といったコンテンツの管理も簡単に行えるようになったため、幅広い分野で利用されるようになった。その一方で、LMS を利用して配信されているコンテンツは、学習者のオンラインテストの成績や講義の進捗状況などでの、単純な制御で配信されることが多く、学習者が必要と考える最適なコンテンツが必ずしも配信されていない。

このような状況から本研究では、学習者のオンラインテストの成績や講義の進捗状況によるコンテンツの制御でなく、コンテンツを利用した学習者の中から、過去に同じような学習方法をしている学習者の閲覧履歴情報を見つけ、もっとも類似性の高い学習者の閲覧履歴情報を活用し、最適と考えられるコンテンツを提示するシステム AIRS (An Individual Reviewing System) の構築ならび教育支援機能の実装をおこなってきた [2][3].

本稿では、著者らが開発した AIRS の概要ならびに、その本システムについての有効性の評価について述べていく。

2. AIRS の開発

AIRS は、学習者に最適な学習教材の配信を目的としたシステム[4]である。また、学習者の学習ならびに復習を支援するための教育支援機能を実装している。以下に本システムの概要図を図 1 に示す。本システムは、Java 言語の Servlet と JSP 技術を利用して開発しており、学習用データベースならびに推薦用データベースは MySQL を利用している。また、Servlet コンテナとして Tomcat5.5、WEB サーバとして Apache2 を利用している。

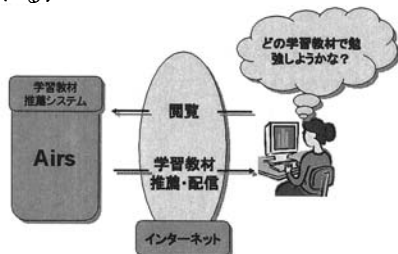


図 1. 概要図

2.1 学習教材&コンテンツ

本システムで配信を行っている学習教材とコンテンツは、本学で開講されている「データベース」の講義を参考に、一連の講義の要点をそれぞれ章と節に分けてまとめたもので、全 6 章構成になったものを利用して配信を行っている[5]。各節に対して 1 つずつのコンテンツを用意するのではなく、各節に対して別々の表現方法を用いたコンテンツを 3 つずつ用意している。この表現方法とは、それぞれ同一の内容のコンテンツを、文章量の多いコンテンツ、中間のコンテンツ、図が多いコンテンツと 3 分割したもので、それぞれを、1, 2, 3 と定めている。

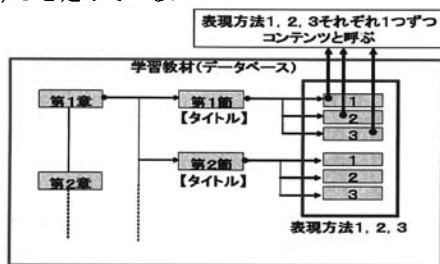


図 2. 学習教材の階層構造

2.2 協調フィルタリング

本システムが利用している技術として、協調フィルタリング[6]がある。この協調フィルタリ

ングとは、学習者の特徴を利用し、似た特徴を持つ他の学習者の情報を元に、新たな情報を推薦する技術である。上記で述べた学習教材を利用し、配信を行うために、協調フィルタリングの技術を利用して、学習者に最適なコンテンツの配信を行っている。

2.3 協調フィルタリングの仕組み

(1). 本システムのデータベースには、各学習者がどのコンテンツをどの表現方法で閲覧したか、という閲覧履歴情報が残されている。仮に A さんの閲覧履歴情報を以下の図 3 とすると

A さんの閲覧履歴

コンテンツ	1	2	3	4	5
履歴	0	1	2	0	3

0 → 未閲覧
1~3 → 表現番号

図 3. 閲覧履歴情報

A さんの閲覧履歴情報とは、コンテンツ 1 を A さんは未閲覧であり、コンテンツ 2 を A さんは表現方法 1 番で閲覧していることを表している。

閲覧履歴情報は、利用者した学習者全員分をデータベースに保存している。

(2). 次に、A さんの閲覧履歴情報を利用し、A さんの閲覧しているコンテンツの特徴を表す特徴ベクトルを生成する。以下に特徴ベクトルの生成方法を図 4 に示す。



図 4. 特徴ベクトルへ変換

生成方法は、表現方法 0 を 000 に置き換え、表現方法 1 を 100 と、表現方法 2 を 010、表現方法 3 を 001 というように各々をベクトルに変換していく。変換して出来上がったものが A さんの特徴ベクトルになる。

(3). 次に、生成された各々の特徴ベクトル同士の類似度を算出する。類似度とは、A さんと他の学習者を比較しどの程度、類似している特徴を持っているかを示す。本システムの中では、どの程度、同じコンテンツの表現方法を閲覧しているのかを算出することになる。類似度の算出方法は、二つの特徴ベクトルのなす角の余弦を利用する事により求める。これにより、二つのベクトルが完全に一致している場合には 1、全く一致していない場合には 0 が求まる、ベクトル空間モデルを利用した方法を使用する[7]。以下に学習者 X と学習者 Y の類似度 a_{XY} を求める式を示す。

$$\alpha_{XY} = \frac{\sum_{i=0}^{\beta} x_i y_i}{\sqrt{\sum_{i=0}^{\beta} x_i^2} \sqrt{\sum_{i=0}^{\beta} y_i^2}} \quad (1)$$

$$\beta = \chi \times \delta \quad (2)$$

a の x_i と y_i は、各学習者の特徴ベクトル i 番目の要素、 β の x は、登録されている学習教材数、 d は表現方法数。学習者同士の類似度 a を計算し、 a がもっとも 1.0 に近い学習者を選び出し、推薦するコンテンツの表現方法の選択をする。例として、B と C の類似度が近く、A が遠いベクトル空間モデルのイメージを図 5 に示す。

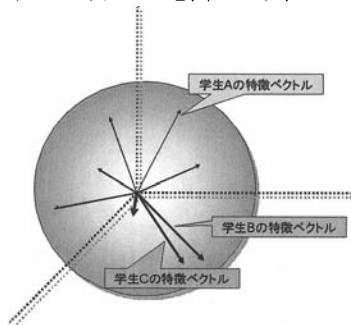


図 5. ベクトル空間モデル

3. 教育支援機能

本システムにおいて、よりの確に学習や復習を行うために、bookmark 機能、学習履歴表示機能、比較・分析機能の実装を行った。これらの機能は、学習者がどのように復習をしたのか、振り返ることにより学習や復習の効果を高めるものである。

3.1 bookmark 機能

学習者が、重要と考える単語やコンテンツに対して、キーワード、コメント、重要度を設定し登録するための機能である。学習者が重要と考える程度により、赤（最重要）、青（重要）、緑（確認）と 3 段階での色による重要度の設定が出来る。また、この機能を利用して、キーワード、コメントと重要度を設定し登録することで、自動的にコンテンツへのリンク情報も付加することが出来る。これにより、登録されたキーワードをクリックすることで、登録先のコンテンツにリンクすることが出来る。登録されたキーワードなどの情報は、bookmark エリアと呼ばれる中央の表示部分に表示され、表示されたキーワードにマウスをロールオーバーさせると、キーワードに登録されたコメントを表示することが出来る。

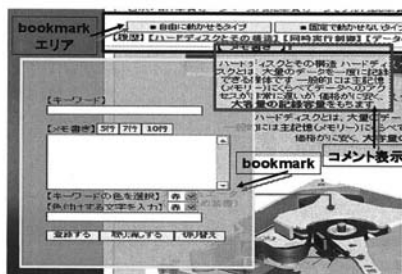


図 6. bookmark 機能

3.2 比較・分析機能

bookmark 機能を利用して、登録したキーワードの情報を元に、学習者間での同キーワードに対しての 3 段階による色の重要度の設定を比較する機能である。同キーワードに対し、どのような色による重要度が設定されていることが多いのかを分析し、表示・比較が可能である。仮に、比較したキーワードの色による重要度が、緑と設定されており、なおかつ他の学習者の多くが同キーワードへの色による重要度を赤と設定している場合など、学習者自身は、同キーワードを重要とは考えていないが、実際には、講義の中で非常に重要なキーワードの可能性がある。そこで、学習者に対しグラフや文章を利用し、再度同キーワードが含まれているコンテンツや関連コンテンツを閲覧するように推薦を促し、うろ覚えや忘れていた部分に対し指摘する。

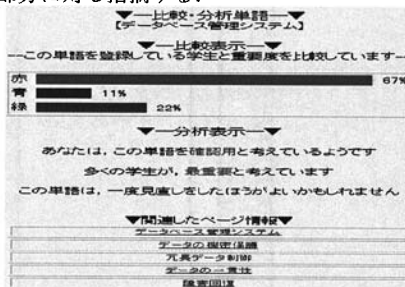


図 7. 比較・分析機能

3.3 学習履歴表示機能

学習者が、今までどのような過程で学習や復習を行ってきたのかを、表やグラフを利用し、時系列で表示する機能である。学習者がどこに着目し学習していたのかを指摘するために、bookmark 機能の色による重要度を活用し、各色の設定によって、結果の表示の方法を変更する。また、学習者が閲覧したコンテンツに関連したコンテンツ情報を一緒に表示することで、学習者が復習する際に、勉強し忘れた部分や関連しているとは考えなかった部分を意識させる。

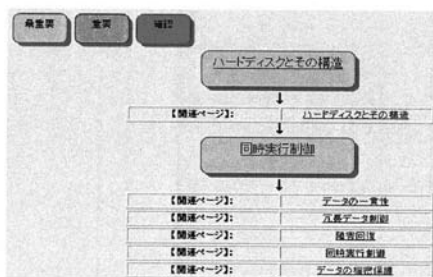


図 8. 学習履歴表示機能

4. 評価

4.1 協調フィルタリングの有効性の評価

本システムで利用している協調フィルタリングの機能が学習者に対し、どの程度適切なコンテンツを表示できているかどうかについて、その有効性の評価実験とし、コンテンツの適合率に関する以下のような実験を行った。なお、本稿で書かれている適合率とは、本システムが推薦するコンテンツと被験者からのアンケート結果の中で、合致した部分の事を指すこととする。

4.2 実験準備

9月から本学部で行われている後期科目の一つ、「データベース」の講義の受講生ならびに、すでに一度履修したことがある学習者、合計20名に対し実験を行った。対象とするコンテンツの数は、31個である。コンテンツには番号が振られており、No1～No31までである。さらに、この31個のコンテンツには、それぞれ3つの表現方法があり、1は、文章が多いコンテンツ、2は中間のコンテンツ、3は図の多いコンテンツと分けられており、合計では93個のコンテンツとなる。

各被験者は、31個のコンテンツの中から被験者自身が考える最適なコンテンツを一つずつ選択する。

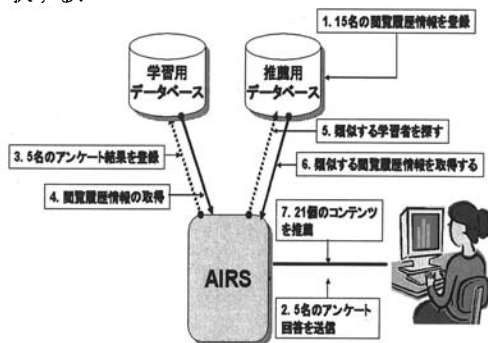


図 9. 実験概要図

4.3 協調フィルタリングの実験方法

(1) 推薦用データベースに履歴データを格納。20名の被験者のうち、15名の被験者が31個のすべてのコンテンツの閲覧を行い、各被験者が、もっと自身に最適と感じられるコンテンツをアンケート形式で選択する。選択したアンケートの結果を集計し、各被験者の閲覧履歴情報とし、推薦用データベースに格納する。

(2) 学習用データベースに履歴データを格納。(1)で回答を行わなかった残りの被験者5名も同様にコンテンツの閲覧を行い、アンケート形式で選択する。ただし、ここで選択するアンケートの数は、No1～No10までの10個のコンテンツだけとする。選択したアンケートの結果を集計し、閲覧履歴情報とし、学習用データベースに格納する。

(3) 本システムからコンテンツを推薦。続いて、(2)の被験者5名が、残りの21個のコンテンツの閲覧を行う。このとき閲覧するコンテンツは、協調フィルタリングを行った結果、本システムが推薦し表示されたコンテンツを閲覧する。

(4) アンケート。実験終了後、(3)の被験者5名が、アンケート形式でNo11～No31までの21個のコンテンツをすべて閲覧し、もっとも自身に最適と感じられるコンテンツを選択する。

(5) 適合率の算出。(4)のアンケート結果と本システムが推薦したコンテンツの結果のログを比較し、選択したコンテンツのうちどの程度適合していたのか調べ、その結果をコンテンツの適合率として報告する。

4.4 協調フィルタリングの評価結果

表 1. 協調フィルタリングの結果

コンテンツ番号	コンテンツタイトル	本システム推薦コンテンツの表現方法(1,2,3)	アンケートで選択したコンテンツの表現方法(1,2,3)
No.11	論理データと独立性	1	2
No.12	ハードディスクとその構造	1	1
No.13	ファイルシステム	2	2
No.29	B*木からのkeyの削除	1	2
No.30	B*木構造	1	1
No.31	木構造のまとめ	3	2

1, 2, 3は、それぞれ選択したコンテンツの表現方法を表しており、薄く色が掛かっている部分が、本システムが自動的に推薦したコンテンツと実験終了後、アンケート形式で選択した残りのNo11～No31までのコンテンツの中で適合していた部分を示している。

4.5 考察

(1) 「コンテンツの適合率について」。協調フィルタリングの計算に利用したNo1～No10までを除いたNo11～No31までのコンテンツの適合率は、約57%。全体としてNo1～No31までのコン

テンツの適合率は約 65%だった。協調フィルタリングの計算に、利用した部分を除いた結果が悪くなる点については、No1～No10 までのコンテンツの適合率が、表 1 の場合 80%もあるのに対し、それ以降のNo11～No31 までの適合率が 57% になっているためだ。この原因として下記の検証にも記述しているが、4.3 の実験の場合 No1～No10 までで、もっとも適合率が高い被験者を協調フィルタリングで求める。ゆえに、No10 までの時点でもっとも類似度が高い被験者の閲覧履歴情報が選択されることになる。しかし、閲覧履歴情報の中には、No10 まで適合率が高くそれ以降の低くなってしまうケースがあり、4.3 の実験では、No11 以降が低くなってしまった結果、適合率の低下に繋がってしまったと考えている。

(2) 「コンテンツの推薦について」。従来からあるようなオンラインテストを繰り返し行うことで、学習者の学力レベルを判断し、表示するコンテンツを決定する方法の場合、一度オンラインテストで学力レベルを決定してしまうと、次回に表示されるコンテンツは、その学力レベルと同様のレベルのコンテンツが単一的に推薦されてしまうことがあるが、本システムが推薦するコンテンツは、表 1 の結果から例えば、No12 では 1 を推薦、No13 では 2、No30 では 1 を推薦しており、単一的な学力レベルのコンテンツの推薦ではなく、より学習者の状況に合わせた木目細かい推薦が出来ている。

(3) 「全体として」。約 6 割の適合率を示すことが出来た。現在は、まだ被験者数があまり多くなく本システムからのコンテンツの推薦の精度には、ばらつきが大きい部分もあるが、従来のオンラインテストの結果によるコンテンツ推薦よりも、木目の細かい推薦ができるため、コンテンツの推薦には有効であると考えている。また、SCORM2004 から規格化された学習シーケンシングのように、シーケンシングルールを作成することなく、閲覧履歴情報から特徴ベクトルを算出し動的にコンテンツの推薦を行うことが出来るため、管理者やコンテンツ作成者にシーケンシングルールの作成の負担をかけることなく学習者に適切なコンテンツを推薦することが可能である。

5. 推薦用データベースの人数変更の検証

4.3 の実験では、推薦用データベースの人数を 15 名と固定し、協調フィルタリングの有効性の評価実験を行ったが、推薦用データベースの人数を変更することで、本システムが推薦するコンテンツと被験者が選択したコンテンツの適合率に

変化があることが予想される。そこで、推薦用データベースの人数を変更し、適合率にどの程度影響があるのか検証を行った。

5.1 推薦用データベース人数変更による検証結果

推薦用データベースの人数を 5 名、10 名、15 名と変化させた結果を下に示す。

表 2. 人数変更の検証

推薦用DBの人数	No11～No31の適合率(%)	No1～No31の適合率(%)
5名	42.8 (9/21)	45.1 (14/31)
10名	33.3 (7/21)	41.9 (13/31)
15名	57.1 (12/21)	64.5 (20/31)

推薦用データベースの人数を 5 名から 15 名に変更することで、適合率が高くなっていることが分かる。これは 5 名の時に比べ 15 名のほうが、適合率が高い被験者がいたことによる結果だと考えている。その一方で、5 名から 10 名に推薦用データベースの人数を変更した時に、一時的に適合率が低くなってしまっていることが分かる。これについては、4.5(1)の考察で書いたことが原因ではないかと考えている。またこの要因として、4.3(2)で行ったアンケートの回答数に問題があったのではないかと考えている。そこで、次の検証を行った。

5.2 アンケート回答数の変更

4.3(2)の評価実験では、アンケートの回答数を No1～N10 までの 10 個のコンテンツとし、協調フィルタリングの有効性の評価実験を行ったがアンケートの回答数によって適合率に影響がある可能性が 5.1 の検証で分かった。そこでアンケートの回答数を 10 個、15 個、20 個と変更することで適合率にどの程度影響があるのか検証を行った。

5.3 アンケート回答数の変更による検証結果

表 3. アンケート回答数の変更

アンケートの回答数	No11～No31の適合率(%)	No1～No31の適合率(%)
10個	57.1 (12/21)	64.5 (20/31)
	No16～N31の適合率(%)	No1～No31の適合率(%)
15個	68.7 (11/16)	70.9 (22/31)
	No21～N31の適合率(%)	No1～No31の適合率(%)
20個	68.7 (11/16)	70.9 (22/31)

アンケートの回答数を変更することで、全体の適合率を高くすることが出来ることが分かった。

今回の検証では、15 個から 20 個へ回答数を増やした場合に、差は出なかったが推薦用データベースに登録している人数を増やすことにより、さらに適合率が向上する可能性があると考えている。しかし、アンケートの回答数を増やすことは、利用者に回答してもらうアンケートの数を増やすことになり、増やすすぎると逆に大変な手間になってしまう可能性がある。適切な回答数ならびに適切なアンケート項目の検討がこれから必要

だと考えている。

6. 教育支援機能の有効性評価

本システムに、実装されている教育支援機能が、学習や復習をする際に、どの程度利用されているかアンケート形式で評価を行った。

6.1 実験準備

本学部で行われている科目、「データベース」と「情報処理の基礎」の講義の受講生ならびに、すでに一度履修したことがある学習者合計 30 名に対し、5 段階のアンケート形式で評価を行った。

表 4. 教育支援機能の評価

1. bookmark機能は、復習を行うさいに役に立ちましたか？				
全く役に立たなかった	役に立たない	普通	役に立った	とても役に立った
0	2	5	11	15
2. 比較・分析機能を使うことで振り返りを行うことが出来ましたか？				
全く出来なかった	出来ない	普通	出来た	よく出来た
7	11	6	6	3
3. 学習履歴表示機能は、自分の学習の振り返りに役に立ちましたか？				
全く役に立たなかった	役に立たない	普通	役に立った	とても役に立った
0	1	9	11	12

6.2 教育支援機能の評価結果

(1)bookmark 機能は、26 名の被験者から「役に立った」「とても役にたった」との意見をもらうことが出来た。「役に立たなかった」と回答した被験者の意見としては、bookmark 機能が利用できなかったという意見があった。これは、bookmark 機能を利用するには、ブラウザの JavaScript を on にしていないと利用できない。そのため off に設定になっている場合やブラウザが対応していない場合など利用することができない。このことから、事前に JavaScript を on にしてもらうことや対応ブラウザの利用、off の場合は注意をポップアップするなどの工夫が必要だと考えている。

(2)比較・分析機能については、「出来た」「よく出来た」と回答した被験者より「出来ない」「全く出来ない」と回答した被験者のほうが多かった。これは比較・分析機能は、あくまで自身が気になるキーワードを比較・分析するための機能であるため、気になるキーワード自体は被験者が自身で選択しなければならない。しかし、被験者の中には、気になる場所はあるが、それを何というキーワードで比較・分析しているかが分からない被験者がいた。そのため結果的に、利用されないケースがあった。また、意見として、入力するキーワードの候補をいくつか表示してほしいという意見があった。そこで、この点について改善していく必要があると考えている。

(3)学習履歴表示機能は、23 名の被験者から「役にたった」「とても役にたった」との意見をもらうことが出来た。意見として、表示部分の工夫を

してほしい回答があった。これは、現在は bookmark 機能の色による重要度から、表示結果に色を付け表示しているのだが、同じ形で表示しているため、区別しづらいことによるものだと考えている。そこで、表示する形も各色によって、変更することで区別しやすくする必要があると考えている。

7. まとめ

本システムの概要、協調フィルタリングの評価ならびに検証、教育支援機能の評価を行った。今後の予定として、閲覧履歴情報の収集とともにコンテンツ推薦の精度の向上について検討する必要があると考えている。また、教育支援機能についても、学習者の意見を取り入れつつ、改良していく予定である。

現在、本研究と平行して、オンラインテストでの協調フィルタリングを利用した問題配信や関連しているコンテンツの推薦方法の検討などが進められている。また、制作した学習教材やコンテンツの XML 化についても検討を行っている。

謝辞

本研究を通してご指導をいただいた、和田教授、土肥講師ならびに評価実験やアンケートに協力していただいた数多くの被験者の皆様に感謝いたします。本研究は、科学研究費補助金（基盤研究（C）課題番号 18500731「Eラーニングにおける学習教材レコメンデーション技術に関する研究」）の支援による。

参考文献

- [1]仲林 清. e-Learning の要素技術と標準化. 情報処理学会誌, Vol43, No4, pp.401-406(2002)
- [2]高橋 泰樹, 松澤 俊典, 山口 未来. フィルタリングを活用した教育支援システム AIRS, 情報処理学会 全国大会(2006.3)
- [3]高橋 泰樹, 松澤 俊典, 山口 未来. 履歴データを活用した教育支援システム AIRS, PC CONFERENCE(2006.8)
- [4]高橋 泰樹, 松澤 俊典, 山口 未来. 履歴データを活用した教育支援システム AIRS, 第 5 回情報技術科学フォーラム(2006.9)
- [5]浪平 博人. 動的視覚化法: 論理的内容の教育における新しい技法 パーソナルコンピュータ学習者利用技術協会 論文誌, Vol.16 No2(2005)
- [6]大杉 直樹 門田 暁人. 協調フィルタリングに基づくソフトウェア機能推薦システム, 情報処理学会論文誌 Vol.45 No.1(2004)
- [7]大谷紀子, 情報検索におけるベクトル空間モデルの応用, 武蔵野工業大学環境情報学部