

NHK デジタル教材のアクセスログ分析・ポートフォリオ収集システムの開発[†]市川尚^{*1}・森山了一^{*2}・宇治橋祐之^{*3}・稲垣忠^{*4}・鈴木克明^{*1}岩手県立大学ソフトウェア情報学部^{*1}・NHK 学園高等学校^{*2}NHK エデュケーショナル^{*3}・東北学院大学教養学部^{*4}

本研究では、NHK が提供するデジタル教材を協力校の授業で利用した際のアクセスログの分析、及びポートフォリオの収集を行うシステムを開発した。アクセスログ分析においては、アクセスログの集計を行うだけでなく、アクセスされたコンテンツを端末ごとに時系列で参照できるようにし、デジタル教材の利用ポートフォリオと併用することで、教材の利用場面を把握できるようにした。

キーワード：デジタル教材、アクセスログ、ポートフォリオ、システム開発、Web

1. はじめに

NHK が2001年度から提供を開始したデジタル教材は、「テレビ」(番組そのもの)、「クリップ」(番組に関連した短い動画)、「ホームページ」(番組内容の解説やクイズ・リンク集等)、「掲示板」(番組利用校同士の情報交換の場)から構成されており、その評価の試みが報告されてきた(稲垣ほか 2004)。評価は、デジタル教材が学校現場でどのように活用され、どのような学習効果を挙げているのかの検証を目的として、イメージマップや情報活用能力テスト、教師へのインタビューや質問紙調査、アクセスログ分析など、多面的に行われている。

本研究では、デジタル教材の活用実態を調べるために、教材の利用傾向を把握するアクセスログ分析機能と、教材が利用された授業の文脈を把握するポートフォリオ

収集機能を有するシステムを開発した。製作者だけでなく教師にも有用な情報を提供するシステムを目指した。

2. アクセスログ分析

アクセスログ分析は、Web サイトへのアクセス数の集計や導線検索(ページ移動の追跡)によって、ユーザの動向(接続環境や時間帯など)やサイトの問題点(アクセスの少ないページなど)を把握し、サイトの改善に役立てられている(鎌田・竹内 2002)。従来までの教材評価でユーザの動きを追跡する場合には、教材に履歴記録機能を組み込むか観察する必要があったが、Web 上の教材の場合は、自動的にサーバにアクセスログが残るため、それを評価に利用することができる。なお、アクセスログは、Web サーバやプロキシサーバが標準的に出力しているログファイル(アクセス元のホスト、参照日時、参照 URL 等の情報が記載)を指すこととする。

Web 教材の評価にアクセスログを用いた研究には、例えば白井・斉藤(1997)が、Web 上の電子教科書のアクセスログから各ページのアクセス数を、リファレンスログから各ページの参照回数を集計し、アクセスの少ないページや参照回数の少ないリンクに対して改善を行っている。真嶋ほか(2001)は、アクセスログはユーザの全体的な動向は分析できるが、教材の洗練を行うにはユーザの背景も把握する必要があるとし、利用目的や職業等をアンケートとして入力させ、その情報を拡張ログとして、アクセスログと共に分析に用いている。

アクセス数を集計した量的な分析によって、デジタル教材の各コンテンツの利用傾向をつかむことができ

2004年4月5日受理

[†] Hisashi ICHIKAWA^{*1}, Ryouichi MORIYAMA^{*2}, Yuuji UJIHASHI^{*3}, Tadashi INAGAKI^{*1} and Katsuaki SUZUKI^{*1}: Development of the System for Analyzing Access Logs and Collecting Portfolios of NHK's Digital Material

^{*1} Faculty of Software and Information Science, Iwate Prefectural University, 152-52, Takizawa aza Sugo, Iwate, 020-0193 Japan

^{*2} NHK Academy of Distance Learning, 2-36, Fujimidai, Kunitachi, Tokyo, 186-8001 Japan

^{*3} NHK Educational Corporation, 6-3, Kamiyama-cho, Shibuya-ku, Tokyo, 150-0047 Japan

^{*4} Faculty of Liberal Arts, Tohoku Gakuin University, 2-1-1 Tenjinzawa, Izumi-ku, Sendai 981-3193 Japan

るが、授業でどのようにデジタル教材が活用されたかを知る上では、量的な集計よりも導線検索のように、ユーザがどのようなコンテンツを見て回ったのかを示す必要があると考えられる。これまでのアクセスログ分析は、主に製作者のために行われてきたものであるが、例えば教材を自由に使いわける活動をした場合に、個々が何を参照したかの情報は、利用者（教師）のためにも役に立つ情報である。そこで、本システムでは時系列で各端末がアクセスしたコンテンツ名を表示するようにした。また、真嶋ほか（2001）がアンケートを入力させたように、アクセスログだけで教材を利用した文脈を判断することは難しく、何の目的でどこを利用したのかの情報を得る必要がある。そこで、アクセスログ分析とデジタル教材の授業利用記録（ポートフォリオ）を併用することで、教材の利用場面を詳細に把握できるようにした。

3. 設計・開発

3.1. システムの概要

本システムは、デジタル教材の活用実態を調べるために、アクセスログの分析や、ポートフォリオの記録・検索が可能となっている。対象者はデジタル教材の製作者と授業で利用する教師であり、「おこめ」、「川」、「びっくりか」、「ふしぎいっぱい」、「にんげん日本史」の5番組に対応している。システムの構成を図1に示す。

本システムの構成は、ログ分析、ポートフォリオ収集、ログ取得の3つのモジュールに分けられる。協力校のプロキシサーバのアクセスログがNHKのサーバに1日1回収集されており、ログ取得モジュールが自動的にログファイルを取得し、ログ分析が可能な形にデータを変換する。各協力校のプロキシサーバのアクセスログを用いることで、LAN内の各端末がどのコンテンツを参照したかの情報を得ることができる。また、ポートフォリオ収集モジュールは、ポートフォリオの記録や検索を提供する。本システムを利用するにはログインをする必要があり、協力校ごとにパスワードが設定されている。

3.2. ログ分析モジュール

3.2.1. アクセス集計

アクセス集計では、アクセスログの集計結果を表とグラフで表示する。デジタル教材の利用傾向を様々な視点から捉えられるように、年度の総アクセス数、月や時間ごとの総アクセス数、放送回ごとのアクセス日数、コンテンツごとのアクセスランキングなどを表示するようにした（表1）。図2は利用ログの概況の画面である。

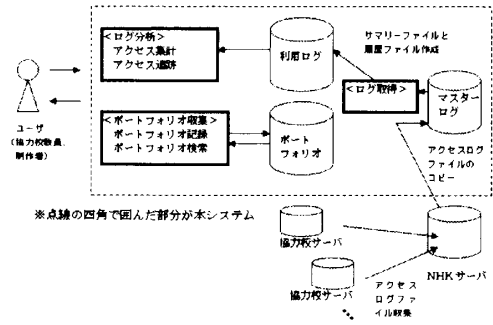


図1 システム構成図

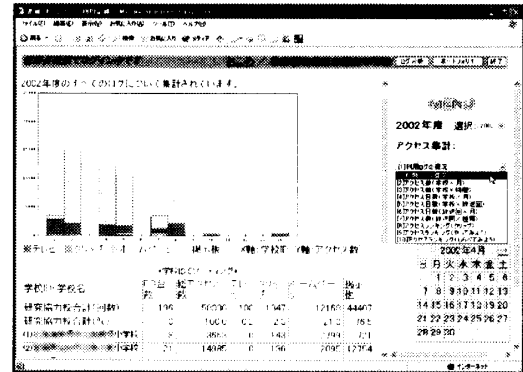


図2 利用ログの概況の画面

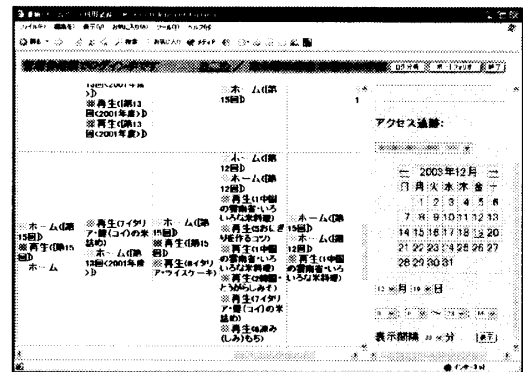


図3 アクセス追跡の画面

3.2.2. アクセス追跡

アクセス追跡は、デジタル教材をいつどのような順序でどんなコンテンツにアクセスしたのかがわかるように、時系列で各端末がアクセスしたコンテンツを表示するようにした（図3）。メニューにあるカレンダーの日付をクリックすることで、その日一日の追跡画面が表示される。追跡画面は、縦軸に時間、横軸にアクセスしたホスト名（IPアドレス）を配置した表で構成し、各セルの中にアクセスされたコンテンツ名が表示される。再生したクリップ名や、過去の年度のもはその年度が、括弧書きで付記される。また、コンテンツ名の前にある四角の色で種類（テレビ、クリップ、

表1 アクセス集計の項目

	選択項目	説明	表示方法
1	利用ログ概況	協力校ごとのパソコン台数（ログのホスト名から判断）、総アクセス数、テレビ・クリップ・ホームページ・掲示板へのアクセス数を表示する。	積み重ね棒グラフ、表
2	アクセス数 （学校×月）	協力校ごとのアクセス数を月ごとに集計したものを表示する。	折れ線グラフ、表
3	アクセス数 （学校×時間）	協力校ごとのアクセス数を時間ごとに集計したものを表示する。	折れ線グラフ、表
4	アクセス日数 （学校×月）	協力校ごとのアクセス日数を月ごとに集計したものを表示する。アクセス日数は、1日に1度でもアクセスがあればカウントされる。	折れ線グラフ、表
5	アクセス日数 （学校×放送回）	協力校ごとのアクセス日数を放送回ごとに表示する。放送回の集計は、ログのURLから放送回を特定できるコンテンツのみを対象とした。	折れ線グラフ、表
6	アクセス日数 （放送回×月）	放送回ごとのアクセス日数を月ごとに表示したものであり、各放送回のコンテンツが何月に参照されたかを把握することができる。	表
7	アクセス数 （放送回×種類）	放送回ごとのアクセス数を種類（テレビ、クリップ、ホームページ）ごとに表示する。掲示板は放送回を特定不能なので対象からはずした。	積み重ね棒グラフ、表
8	アクセスランキング （クリップ）	クリップごとのアクセス数をもとに、参照数が多かったもの上位30位を表示する。	表
9	アクセスランキング （やってみよう）	「川」以外の番組で提供、やってみようで提供されているゲームごとのアクセス数をもとに、参照数が多かったもの上位30位を表示する。	表
10	アクセスランキング （しらべてみよう）	「おこめ」のみ提供。しらべてみようでアクセスした用語の参照数が多かったもの上位30位を表示する。	表

ホームページ、掲示板）を表すようにした。端末ごとに分けて表示しているため、個別（もしくはグループ別）で端末を利用した場合でも、各々が参照したコンテンツを特定することができる。授業のポートフォリオと見比べることで、デジタル教材を利用した授業活動をより詳細に把握できる。

3.3. ログ取得モジュール

ログ取得モジュールは、ログ分析モジュールで利用するデータの準備を行う。まずNHKのサーバから毎日定期的に前日のアクセスログを自動的に取得してくる。取得後にURLとコンテンツの対応関係を記述したファイルとマッチング処理を行い、同時に参照日時、アクセス元のホスト名を取得し、アクセスの集計結果を保存したサマリーファイル（アクセス集計機能に利用）と、1日ごとに参照したコンテンツ名や時刻を保存する履歴ファイル（アクセス追跡機能に利用）を生成する。

マッチング処理用のファイルには、筆者らがホームページの構成を事前に調査し、手動でURLとコンテンツの対応を記述した。記述内容は、ひとつのコンテンツに対して、対象URL（正規表現で記述）・分類（テレビ・クリップ・ホームページ・掲示板）・タイトル（対象URLに対応したコンテンツ名）・サブタイトル・放送回フラグ・ランキングフラグである。ページが更新された場合には、書き換えが必要となることもある。対象URLは正規表現で記述しているため、URLとコンテンツの対応に法則性があるものやディレクトリで整理されているものは、まとめて記述できる（例えば放送回の番号がURLに含まれる場合は、番号部分を抜

き出して放送回やタイトルに利用できる）。逆に法則性がないとコンテンツ毎に記述する必要がある。放送回とランキングのフラグは、それぞれが特定可能だった場合に、プログラムに処理をさせるために用いる。マッチング処理用のファイルは年度ごとに記述され、最新の年度で見つからない場合は、ひとつ前の年度に戻って該当するコンテンツがないかを解析する。また、これらの処理はURLを基に行っているため、CGIやFlashでURLに反映されない場合には、放送回やタイトル等を特定することができない。

3.4. ポートフォリオ収集モジュール

3.4.1. ポートフォリオ記録

ポートフォリオ記録は、教師が授業のポートフォリオを登録するための機能である。図4にポートフォリオの登録画面を示す。入力項目は、番組を利用した授業実施日、記入者名、放送タイトル、利用したデジタル教材カテゴリ、利用形態、利用内容、関連学習活動、実施後の感想・コメントとなっており、授業中に利用した教材を添付資料としてアップロードすることもできる。登録したポートフォリオは、閲覧・修正・削除が可能である。

3.4.2. ポートフォリオ検索

ポートフォリオ検索は、他校のポートフォリオを検索・閲覧する機能である。検索方法には、放送タイトル別一覧と協力校別一覧の2種類を用意した。

3.5. 開発環境

本システムは、Webサーバ上でPerl（CGI）により稼働し、ブラウザからすべての機能を利用できる。グラフ出力にはPerlのGD-Graphモジュールを用いている。

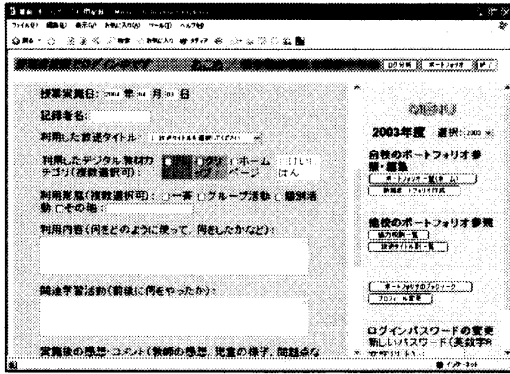


図4 ポートフォリオ入力画面

4. 運 用

本システムは、2002年度末から運用を開始し、2002年度分のアクセスログを一括して集計した。対象校は22校であり、結果は稲垣ほか(2004)で報告している。

2003年度は、12月18日から運用を開始した。協力校がデジタル教材を利用したアクセスログは本システムで自動取得される。協力校の教師には実践時のポートフォリオ記入を依頼した。アクセスログが取得できた協力校は整備の都合もあり、「おこめ」の1校であった。

その1校のポートフォリオへの記録は、授業1回分であったが、グループごとにデジタル教材(テレビとクリップが中心)で調べる活動であり、各活動に対して1件のポートフォリオ(計9件)が記録されていた。

アクセス集計の「利用ログ概況」は、運用開始日から3月末までに計21台の端末から、6199回のアクセスがあった。内訳は、テレビが259回(4.2%)、クリップが171回(2.8%)、ホームページが2304回(37.2%)、掲示板が3465回(55.9%)であった。掲示板の利用が最も多かったため、他校との交流が盛んだったことが伺える。「アクセス数(学校×時間)」から、アクセスの多い時間帯は15時と12時と10時であり、授業だけでなく昼休みや放課後にも利用されていた。「アクセス日数(放送回×月)」を見ると、以前に放映された番組のコンテンツにもアクセスされていることがわかった。

アクセス追跡からは、ポートフォリオに利用したと記載されていた番組以外にも、アクセスしたコンテンツを特定できた。例えば、放送回が2003年第11回の番組「現代おこめ事情」を視聴したと記載のあったグループは、2001年第8回「こめ作りの歴史」、2002年第18回「田んぼの持つ力」も視聴していた。このように、自由にグループ等で活動をさせた場合に、何を参照したのかを把握するためにアクセス追跡は有効であった。

また、この実践では過去の関連したコンテンツへのアクセスが多く、過去のコンテンツを簡単に参照できるような工夫がデジタル教材に必要であることが示唆された。

一方で、アクセス追跡では、どのグループがどの端末を利用したかの対応づけが困難であった。今回はポートフォリオがグループ活動ごとに記載されていたため、ある程度推定することはできたが、今後の課題として残る。また、ポートフォリオに授業時間の項目が無く、アクセス追跡画面でどこまでが授業なのかが不明であった。授業時間の項目を設け、その時間のみアクセス追跡が参照できるようにするなど、ポートフォリオとアクセスログ分析をより連携していく必要がある。アクセス集計では、集計期間が設定できなかったため、例えば学期や単元における集計などができるようにする必要もある。

5. お わ り に

本研究では、NHK デジタル教材のアクセスログ分析、及びポートフォリオの収集を行うシステムを開発した。運用を通して、アクセス集計による全体的な傾向の把握だけでなく、ポートフォリオとアクセス追跡によって、どのような授業の文脈でどんなコンテンツにアクセスしたのかを把握できることが示唆された。また、製作者のためだけでなく、教師のリフレクションや他校の記録を授業の立案に役立てることも考えられる。今後の課題としては、本システムの教師や製作者による評価、ポートフォリオとアクセス分析機能の連携等が挙げられる。

参 考 文 献

- 稲垣忠, 岡本恭介, 市川尚, 鈴木克明, 宇治橋祐之, 小平さち子, 黒上晴夫(2004) デジタル学習環境における教材設計—NHKデジタル教材を対象にした評価研究の取り組みから—。教育メディア研究, 10(2): 15-22
- 鎌田英広, 竹内実(2002) ウェブサイトリデザイン—アクセス解析を活用した戦略的サイト設計—。ソフトバンクパブリッシング株式会社, 東京
- 白井靖敏, 斉藤暢久(1997) インターネット上の電子教科書の開発とその分析—高校生向け「街角物理探検」—. 日本教育工学雑誌, 21(Suppl.): 53-56
- 真嶋由貴恵, 宗陽一郎・島田恭宏, 塩野充, 橋本禮治(2001) インターネットのWebサイトを利用した看護生涯教育システムの開発。日本教育工学雑誌, 25(2): 149-157

(Received April 5, 2004)