

Web アクセスに見る奈良高専像 —機械工学科・情報工学科の場合—

平 俊男 武藤 武士*

奈良工業高等専門学校 機械工学科 (hira@mech.nara-k.ac.jp)
*奈良工業高等専門学校 情報工学科 (mutoh@info.nara-k.ac.jp)

概要 webによる情報発信へのフィードバックとしてアンケート等が考えられるが、その回答は設問の枠を超えにくく、発信者にとって必ずしも有効な情報とは限らない。本報では、web検索に用いられる検索語に注目することで、web利用者の関心を探る手法を提案し、ケーススタディとして本校機械工学科、情報工学科 web ページへのアクセスにみられる検索語を示した。発信者の想定外なものも含む幅広い検索語がみられ、これを参考とすることで、より充実した情報発信が可能となると考えられる。

1. はじめに

現在、webを手段とした様々な情報の発信が広く行われており、高等専門学校もその例外ではない。情報の発信にあたっては、どのような情報が要求されているかを把握することが必要不可欠となる。これに対して、アクセスカウンタによる閲覧件数の集計やアンケートによるフィードバックなどが考えられるが、十分な把握は行えていないのが現状である。

一方、webの爆発的な普及に伴って、goo¹等の検索ツールによって知りたい情報を掲載しているwebページをキーワード検索する利用法が一般的となってきた。著者らは、webのサーバであるhttpdがアクセスの参照元を記録するreferer_logに注目し、検索ツールを経由するアクセスの際に残される検索語の抽出・集計を試みた。本手法では、アンケートのように設問の枠内に限られがちな回答だけでなく、web利用者が実際に関心を持っている多様なキーワードが明らかになると考えられる。

2. referer_log に記録される検索語

図1は、NCSA HTTPd NCSA/1.5.1のreferer_logの一例である。このファイルには、ページtoへのアクセスはページfrom(URL表記)内のリンクをたどることにより行われたことがfrom->toの形式で記録されている。図1(a)の例では、“/mech/ob.html”へのアクセスが、“http://www.mech.nara-k.ac.jp/”で示されるページを参照して行われたことを示している。閲覧者が、検索ツールによりページを探し出しアクセスした場合には、図1(b)(c)に示すように、fromのURLは検索語を含む形式となる。ここに示した検索ツールgooの場合には、“MT=”以降“&”までの文字列、“Nara+National+College+of+Technology+Mechanical+Engineering”が検索キーワードを表し、“-”以降の“/”が閲覧されたページを表す。

検索が漢字等の2バイト文字を含む検索語によって行われた場合、fromはRFC1738⁽¹⁾で規定されたエン

コーディングに則り記述される。例えば‘奈’はEUCの文字コードセットで0xE0C6と表現されるが、URL中では“%C6%E0”と記述される。従って図1(c)に含まれる検索語“%C6%E0%CE%9”をデコードすると、“奈良”であることが分かる。

ここでは、webアクセスの解析に関して検索語のみに注目する。もちろん、検索の性質上用意されているwebページに含まれていないキーワードは拾い得ないので、不特定多数のweb閲覧者総体の検索傾向を知ることはできない。しかし、ここで試みる手法では、実際にアクセスした特定の閲覧者の検索語に注目することで、実際の関心度を知ることが可能になると考えられる。

3. 機械工学科へのアクセスに用いられた検索語

3.1 機械工学科 referer_log の概要 ここでは、機械工学科 (<http://www.mech.nara-k.ac.jp/>) 以下のwebページへのアクセスをとりあげる。対象となるページは、学科概要、関連行事等の新着情報、教育課程(シラバス)、スタッフ紹介(教官・技官)、他学科等へのリンク、アンケート、専攻科概要、教官・技官の作成によるページ(経歴等)からなる。なお、学生が個人的に作成したページは存在しない。集計の対象としたのは、webページ開設(1996年6月)から1998年5月23日までに記録されたreferer_logである。

`http://www.mech.nara-k.ac.jp/ -> /mech/ob.html`
(a) 通常のリンク参照の例

`http://www.goo.ne.jp/default.asp?MT=Nara+National+College+of+Technology+Mechanical+Engineering&SM=MC&WTS=ntt&DE=2&DC=10&_v=2&act.search.x=41&act.se`
`arch.y=9 -> /`

(b) 検索ツール経由の場合(1バイト文字による検索語)

`http://www.goo.ne.jp/default.asp?MT=%C6%E0%CE%9%B9%E2%0%EC+%B5%A1%B3%A3&SM=MC&WTS=ntt&DE=2&DC=10&_v=2&act.search.x=26&act.search.y=8 -> /mech/ob.html`

(c) 検索ツール経由の場合(2バイト文字による検索語)

図1 referer_log に記録される文字列

¹ <http://www.goo.ne.jp/>

表1 機械工学科に対する頻出検索語 (15 回以上)

検索語	頻度
P (T社のハイブリッド乗用車) ²	127
奈良工業高等専門学校	49
Linux+v (ノートパソコン)	35
スターリングエンジン	32
振動解析	29
V (ノートパソコン)	21
T社 (自動車メーカ)	19
リアモーターカー	18
機械設計	17
振動	17
ロボット+H社 (自動車メーカ)	15
大阪府立工業高等専門学校	15
新素材	15

表2 機械工学科に対する検索語 (奈良高専等)

カテゴリおよび検索語	頻度 (計 121)
・奈良高専	56
奈良高専: 49, 奈良高専+スターリング等	
・高専	29
高専+専攻科等: 20, 高専: 6, 編入等	
・他高専	16
大阪府立工業高等専門学校: 15	
・制度等	7
学位授与機構: 6	
・機械工学・機械工学科	6
・奈良	5
・大和郡山市	2

3.2 集計 検索ツールを参照元とする機械工学科へのアクセスは 1078 件であるが、一部の検索ツールでは前章に述べた形式で検索語を記録しないため、検索語の不明なものが 23 件あった。表 1 は 15 回以上の頻度で現われた検索語を示している。ただし「奈良高専」と「奈良工業高等専門学校」のような検索語は同一視して扱い、企業等の営利団体名および商品名は省略した。

表 2~6 に、検索語を高専関連、学科目関連、技術・工学関連、本校機械工学科関連、その他に大別し、それぞれについて分類したものを示す。表中「・奈良高専」のように表記した項目はカテゴリとそのカテゴリに対する頻度である。表 2~4 では、5 回以上の頻度で現われた検索語は頻度も含め、2 回~4 回の頻度で現われたものは検索語のみ、1 回しか現われなかったものは省略して示している。カテゴリをまたぐ連言型の検索語に対しては、先頭語を分類の対象とした。

3.3 考察 表 1 に示すように、検索語には新製品などの技術的トピックスだけでなく、奈良高専などの学校概要にかかわるもの、振動や設計などの機械工学の基礎的分野にかかわるものも多くみられる。このことから、検索ツールを用いる web 利用者の層は厚く、多様な背景からアクセスが行われていることが分かる。また、表 2 からは、「奈良高専+スターリング」など限定的な検索

² 他に比べ非常に多い 50/週以上の頻度がみられた。しかし、同期間中に「インターネットで P が当たる」というキャンペーンが行われていた模様であり、この検索語の取扱いには注意が必要である。他の検索語についても同様の可能性は否定できないが、時間当たりの検索数については特に異常はみられなかった。

表3 機械工学科に対する検索語 (学科目)

カテゴリおよび検索語	頻度 (計 289)
・熱工学	50
スターリンエンジン: 35, エンジン: 5, 熱工学, 熱伝導, 蒸気等	
・振動工学	47
振動解析: 29, 振動: 17, 消音, 逆位相等	
・設計	43
機械設計: 17, 概念設計: 5, 歯車, 設計, 工程設計, 潤滑, 設計法等	
・機械工作・加工	30
機械工作実習: 6, 工作機械, 旋盤, 溶接, ガラス切削, 機械工作, 機械工作法, 工具等	
・材料	29
新素材: 15, 材料強度, 金属材料, CBN 等	
・材料力学	23
材料力学: 13, たわみ: 5, ひずみエネルギー等	
・製図	11
機械製図: 8, その他	
・計測工学	10
計測工学: 7, 形状測定等	
・工業英語	10
・生産工学	8
生産管理: 6, 生産システム等	
・塑性工学	6
塑性力学: 5	
・制御工学	4
機械制御等	
・電気工学	4
コイル等	
・情報処理	4
C言語等	
・エネルギー変換	4
・流体工学	3
・基礎力学	2
・機械工学実験	1

表4 機械工学科に対する検索語 (技術・工学)

カテゴリおよび検索語	頻度 (計 204)
・ロボット	55
ロボット+H社: 15, 自立歩行ロボット: 9, ロボット+歩行: 7, 二足歩行ロボット: 5, ロボット, 人間型ロボット, 自律分散ロボット等	
・環境問題	45
ハイブリッド: 9, 排気ガス: 8, 二酸化炭素, リサイクル, 原子力, リーンバーン, 燃費, 環境+自動車等	
・鉄道等	32
リアモーターカー: 18, 鉄道総合研究所: 11, 高速鉄道等	
・自動車・バイク・自転車	28
バイク+Y社: 6, エアバッグ: 5, 電気自動車, バイク, 新型バイク, 電動自転車, パンク, 自転車等	
・技術一般	25
スラップスケート: 7, クローン, 仮想現実, 遺伝子的アルゴリズム等	
・資格	8
工業英検: 5, Professional Engineer 等	
・宇宙	6
ボイジャー, マーズパスファインダー等	
・未来・社会	5
未来のエンジン等	

語がみられ、身近なところに web 閲覧者が存在することが推察される。

学科目毎に分類した表 3 では、エンジン関連の検索語を熱工学に含めているが、学科目としての基本的語彙に関しては 10 件程度である。これを考慮すれば、振動、

表 5 機械工学科に対する検索語 (機械工学科全般)

カテゴリ	頻度 (計 90)
・個人名 (教官・技官・非常勤講師)	60
・出身校等	26
・学会 (機械学会等)	3
・公開講座	1

表 6 機械工学科に対する検索語 (その他)

カテゴリ	頻度 (計 331)
・商品名 (自動車等)	135
・商品名 (コンピュータ・OS 等)	88
・企業名	57
・書名等 (一般書)	45
・書名等 (専門書)	6
・その他 (分類不可能, 不明等)	43

表 7 主なページへの検索ツール経由のアクセス数 (総数 1078)

ページ	アクセス数	ページ数
/ (機械工学科トップページ)	12 (1.1%)	1 (0.7%)
syllabus/以下	469 (44%)	26 (19%)
~hira/1m/以下	413 (38%)	40 (30%)
~ac-wada/, ac-mech.html	55 (5.1%)	2 (1.5%)
	計 949(88%)	計 69(51%)

設計, 加工, 材料といった実務的分野に対する関心が多くみられることが分かる。表 4 に示した技術・工学関連の検索語では, ロボットや環境問題に関心が集まっている。一方で, 一般に広く話題となったクローンなどの科学的トピックスに関するアクセス数は少ない。

表 5 の教官名に対する検索では, 傾向として 30 代前半までのフルネームに対するものが多くみられた。著者の経験では同年代の知人等が検索してくるケースが多く, web 検索の利用者には 30 代前半までの者が多いことが推察される。表 6 では自動車等の商品名について突出的な検索がみられるが, このほとんど (127 件) は 'P' に関するものであり (脚注参照), 実際上は, あるコンピュータと OS との組合せに関する検索が多い。

検索対象となったページと実際に検索されたページ数との比較を主なページについて行ったものを表 7 に示す。http://www.mech.nara-k.ac.jp/ 以下の全ページ数 135 に対して, 約半数のページのみで web 検索のほぼ 9 割に対応していることが分かる。特に, シラバスに関してはページ数が 26 と全体の約 2 割であるが, 検索結果によるアクセスでは 44% を占めている。

4. 情報工学科へのアクセスに用いられた検索語

4.1 情報工学科 referer_log の概要 情報工学科では情報工学科のサーバ³と実習室のサーバ⁴, さらに武藤研究室のサーバ⁵が一般向けのサービスを行っている。今回の集計では, これらを区別せずに行った。サービスの主な内容は以下の通りである。

³ http://www.info.nara-k.ac.jp/

⁴ http://herb.info.nara-k.ac.jp/

⁵ http://sentinel.info.nara-k.ac.jp/

表 8 情報工学科の頻出検索語 (40 回以上)

検索語	頻度
情報処理技術者試験	260
ソート (アルゴリズム)	222
xgraph	138
CASL(COMMET)	129
高専 (奈良高専を除く)	96
a2ps-j	94
IE5.0	76
ウインドウマネージャ	69
双子	61
図書館 (読書案内, 文献, 司書)	55
制御工学 (フーリエ変換, 標本化定理)	54
UNIX マガジン	50
電子工学 (アナログ回路, 衛星通信)	49
mp3	48
FTP	44
奈良高専	44

表 9 情報工学科の検索語 (奈良高専など)

カテゴリおよび検索語	頻度
他高専	94
奈良高専	44
金魚	13
奈良	9
大和郡山	7

- 学科情報: 教職員名やシラバス, 学生や教官が作成したページ
- コンピュータの利用方法: 各ソフトウェアの設定方法や利用方法など
- 奈良高専同窓会のページ
- 情報処理技術者試験模擬試験
- 整列アルゴリズム可視化ツールのページ
- UNIX マガジンの目次検索

集計の対象としたのは web ページ開設時から 1998 年 6 月 30 日までに記録された referer_log である。

4.2 集計 検索ツールを参照元とする情報工学科へのアクセスは 3412 件である。表 8 に 40 回以上の頻度で現れた検索語を示す⁶。集計方法は機械工学科の集計方法に準じた。ここでは, 情報処理技術者試験に関する検索が群を抜いて多く, 模擬試験のサービス⁽²⁾によるものと考えられる。アルゴリズムに関する検索も, 整列アルゴリズム可視化ツールを公開している関係と考えられる。表 9 ~ 13 は機械工学科と同様の分類を行ったものである。

⁶ ここで用いた Perl のコードを http://sentinel.info.nara-k.ac.jp/search/search.cgi として公開しているので興味のある方は参照されたい。

表 10 情報工学科の検索語 (学科目)

カテゴリおよび検索語	頻度
アルゴリズムとデータ構造	222
制御工学 (フーリエ変換, 標準化定理)	54
電子工学 (アナログ回路, 衛星通信)	49
プログラミング (C 言語, Pascal)	33
画像処理	32
デジタル回路 (Parthenon, spice)	21
人工知能 (人工生命)	19
オートマトン理論	19
論理回路	14
計算機ネットワーク	11
情報倫理	9
計算機工学	1
情報理論	1

表 11 情報工学科の検索語 (情報技術)

カテゴリおよび検索語	頻度
ソフトウェア (xgraph: 138, a2ps-j: 94, IE5.0: 76, その他: 62)	370
情報処理技術者試験	260
ネットワーク (FTP: 44, xv: 28, mailinglist: 17, その他: 54)	143
CASL/COMMET	131
UNIX マガジン	50
DTP (PostScript: 27, LaTeX: 18)	45
OS	32
エミュレータ	18
ニューラルネットワーク	8
遺伝的アルゴリズム	8

表 12 情報工学科の検索語 (情報工学科全般)

カテゴリおよび検索語	頻度
教官名	39
学生名	35

4.3 考察 表 8 に示したように, 情報処理関係の検索語が多く, 特別なサービスを行っているページにアクセスが集中している. また, 検索語にはコンピュータの利用技術に関連するものが多く, 現状のインターネットの利用形態の多くは技術情報の検索であることを現わしていると考えられる. 表 9 で, 他高専名によるアクセスが多いのは, 奈良高専同窓会のページで各高専の同窓会ページへのリンクが行われているため, ここでの高専名称が検索語と一致するためである.

科目別分類 (表 10) を見ると, シラバスが不完全ながら用意されているにもかかわらず機械工学科とは異なり, 検索語としては多くない. これは他大学・高専の情報工学系の学科などでも, 同様の情報提供がされているために, 検索を行ったとしても本学の情報までたどり着けないケースが多いためだと考えられる.

情報技術関連用語に関する検索は表 11 に示したように, かなりの割合を占めている. 特にソフトウェアやネットワークに関する検索が多い. 表 13 に示した検索語は機械工学科とは様子が異なり, 芸能人名等も多い.

表 13 情報工学科の検索語 (その他)

カテゴリおよび検索語	頻度
個人名 (芸能人等)	222
個人名 (不明)	133
漫画, アニメーション	39
herbal	27
映画・番組・書名	26

これは, 主に学生が作成したページが検索されている結果である.

5. Web 検索の分析によるフィードバックの指針

本校他学科等を参照元とするアクセスと比較して, 検索ツールを参照元とするものも多く (機械 3888 : 1078, 情報 5223 : 3412), 検索によるアクセスが無視できない状況であることが分かる.

このような状況の中で, 機械工学科にはシラバスに掲載されているような実務的傾向の強い単語, 情報工学科には情報技術関連の単語が検索によく用いられる関心の強いものであることが明らかとなった. しかしながら, 例えば機械工学科シラバスでは 10 行程度の講義概要のみ, 情報工学科の CASL/COMMET の検索に対応している情報は情報実験の手引だけであるなど, 関心度が高くとも簡単な説明程度の情報も多い.

ここに示した手法によって, web 利用者の関心は決して一部の先端的情報関連技術だけでなく, 学科目関連の基礎的項目等にも広がっていることが明らかとなった. このような知見を web ページの作成へフィードバックしていくこと, 例えば CASL/COMMET に対応している情報実験の手引きについては, 情報処理技術者試験のページ⁷と関連づけて解説を行うことなどにより, 他サイトとは異なる有効な情報発信が可能となると考えられる.

6. おわりに

現在の web には, 以前のようなリンク主体のネットサーフィンよりも, サーチエンジンで関連付けられた膨大なデータに基づく情報検索のイメージの方がふさわしい. ここでは, ページ作成者の意図とは無関係に, 検索時にあたえられる評価値によってページの価値が決められるという点で, 特異な状況が見られる. いったん作成されたページが外部からどのように見えるかは, もはや作成者の予測外の範囲である. このような利用形態に対して, 本報で示した分析手法は, web ページに独自性を与えるための有効な手がかりの一つであると考えられる.

参考文献

- (1) T. Berners-Lee, L. Masinter, and M. McCahill, "Uniform Resource Locators," <ftp://ds.internic.net/rfc/rfc1738.txt>, Sec.2.2, (1994).
- (2) 武藤武士, 榊原和彦, 成田統一, "WWW を利用した選択式問題テスト採点システムの試作", 情報処理教育研究会予稿集, (1996), 69.

⁷ <http://herb.info.nara-k.ac.jp/InfoExam/>