

Web検索語による意識推定の試み (検索語解析ツールの構築と評価)

平 俊男・武藤 武士

A preliminary inference of intentions on the web with searched words
— Case studies by an analyzer for web-searched words —

Toshio HIRA and Takeshi MUTOH

WWW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、WWW利用者の意識を推定する手法を提案し、検索語解析ツールの構築を行った。また、ケーススタディにより本校機械工学科webページへのアクセスにみられる検索語を示し、ツールの有効性の評価を行い、さらにツールの応用可能性について議論を行った。

1. はじめに

World-Wide-Web (WWW)に限らず、情報の発信にあたっては何が要求されているかを把握することが必要不可欠である。一般にwwwにおいては、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行われているが、ニーズの十分な把握は行えていないのが現状である。これに対して著者らは、wwwの利用形態が、当初のリンク主体のネットサーフィンからサーチ主体の情報検索へと変貌しつつあることに注目し、サーチエンジンの利用における検索語(以下、web検索語)を把握することで、アンケートのような設問の枠内に限られがちな回答だけでなく、www利用者の実際的な関心を探る可能性を示した¹⁾。

本報では、web検索語の解析によるwww利用者の意識推定と解析ツールの構築について考察を行い、ケーススタディによって解析ツールの評価を試みる。また、機械工学科設置のwwwサーバ上に構築したツールによる集計・推定結果からツールの他の応用可能性について議論する。

2. Web検索にもとづくアクセス

Webページに含まれる情報へのアクセスには、ブラウザ上に表示されているwebページ上のリンクをたどっていく場合と、探したい情報を記述する検索語をサーチエンジンに入力し提示されたリンクをたどる場合とに大別され、ここではそれぞれ誘導型、検索型と名付ける。一般にwebページの多くは前者を想定しており、この場合、

情報の分類やリンクの名称付けの成否がアクセスを誘導するための重要な因子となる。また、他のwebページからのリンクの存在が、アクセス数を左右することになる。たとえば典型的な学校紹介のwebページでは学科構成、スタッフ紹介等、階層的にカテゴリ化された情報が掲載されている。www利用者は、学校一覧が掲載されたwebページなどからリンクをたどってその学校のwebページにアクセスすることになる。

一方、goo (<http://www.goo.ne.jp>)等の全文検索型サーチエンジンでは、webページに含まれている情報は単語等をキーに再組織化されており、階層的なカテゴリに意味は存在しない。このような状況においては、webページ作成者の意図とは無関係に、検索時に定められる評価値によってページの価値が定められることになる。

機械工学科で公開しているwebページに対するアクセス(99年1月18日～9月2日の227日間)のうちHTMLで記述された文字情報のみについて見ると、総アクセス約50300件のうち、本校他学科のwebページの学科紹介リンクを経由したものが約700件、本校トップページのリンクを経由したものが約1800件であるのに対して、サーチエンジンによる検索結果をもとにした外部からのアクセスは約25400件であった。なお機械工学科webページ内のリンクを参照するアクセスは約20500件である。すなわち、ほぼ半数のアクセスが情報検索型のwww利用に基づくものであり、これらのアクセスを無視できない状況であることがわかる¹⁾。

¹⁾ ここではアクセスの件数のみを比較しているが、もちろん性質の把握が必要である。一般に誘導型では当該webページから関連する他のwebページへとアクセスが広がることが多いのに対して、検索型では当該webページのみが読まれる傾向が強い。

3. Web検索語解析ツール

3.1 Web検索語 http (Hyper Text Transfer Protocol) を処理するwwwサーバは、一般に自サイトで保持するwebページへのアクセスについて参照元リンクを記録する。たとえば、NCSA HTTPd NCSA/1.5.1の場合では、`referer_log`と名付けられたファイルに、ローカルページtoへのアクセスがページfrom内のリンクをたどることにより行われたことをfrom->toの形式で記録する(図1)。図1(a)の例では、“/memo/ob.html”へのアクセスが、“http://www.mech.nara-k.ac.jp/”で表わされるページ中のリンクを参照して行われたことを示している。多くのサーチエンジンでは、参照元リンクfromのURL (Uniform Resource Locator) に検索語を含めるので、www利用者がサーチエンジンを経た場合には、図1(b)(c)に示す形式となる。ここに示したサーチエンジンの場合、文字列“MT=”と“&”との間にみられる“Nara+National+College+of+Technology+Mechanical+Engineering”が検索文字列を表し、“->”以降の“/”がアクセスされたページを表している。

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RFC1738²⁾で規定されたエンコーディングに則り記述される。たとえば“奈”はEUCの文字コードセットでE0C6hと表現されるが、URL中では“%C6%E0”と記述されることになる。したがって図1(c)に含まれる文字列“%C6%E0%CE%C9”をデコードすると、検索語は“奈良”であったことが分かる。

3.2 Web検索語解析ツール 記録されている参照元リンクのURLに含まれている様々な検索語に対して、ここでは、a)頻度、b)文脈、c)アクセスページについて注目し、解析ツールの仕様を検討する。もちろん、検索の性

```
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+Natio
nal+College+of+Technology+Mechanical+Engineeri
ng&SM=MC&WTS=ntt&DE=2&DC=10&_v=2&act.search.x=
41&act.search.y=9 -> /
```

(b) サーチエンジン経由(1バイト文字による検索語)

```
http://www.goo.ne.jp/default.asp?MT=%C6%E0%CE%
C9%B9%E2%C0%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) サーチエンジン経由(2バイト文字による検索語)

図1 記録される参照元リンク

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www利用者総体の検索傾向を知ることはできない。ここに示す手法による解析結果には、ページ作成者の意識とページ検索者の意識との積集合が現われることになる²⁾。

a) 頻度情報 記録されている検索語の頻度は、その検索語が示す概念へのwww利用者の関心の強さを指す直接の指標となる。本ツールでは、web利用者が指定した検索文字列から、まず、空白文字(' '), 記号('&', '+')等をキーに文字列を分解し、それぞれを検索単語と呼び、これらの頻度を数える。

たとえば、「燃費」、「燃費 車」、「車&低燃費」、「燃費のよい車」の検索文字列があった場合には、「燃費」×2、「車」×2、「低燃費」×1、「燃費のよい車」×1が検索単語および頻度となる。

b) 文脈情報 単語単位で頻度を数える場合、様々な文脈で使われうる一般的な単語(前項の場合「車」)が多くなるなど、本来の検索意図が不明瞭となる。また、一般に日本語の文章は欧米語とは異なり、単語単位への分解は意味解析なしでは困難となる。ここでは、前項で抽出した単語をキーに文字列を照合し、その単語が含まれている検索文字列を全てとりあげ検索文脈と呼ぶ。この時、文字列中の単語位置も考慮し、単語が文字列中の先頭に位置する頻度も数えることで、意図推定の補助とする。

前項の例では、検索単語「燃費」に注目した場合には、「燃費」、「燃費 車」、「車&低燃費」、「燃費のよい車」の4つの検索文字列全てが「燃費」に関連する文脈であるのに対して、「車」に関連する文脈は3つである。さらに、位置情報を考慮した場合には、「燃費」が先頭に位置している文字列が3つなのに対して、「車」が先頭にあるのは1つのみである。したがって、これらの検索文字列からは、www利用者の関心度が「燃費」>「車」であり、検索時には「燃費」に重きをおいていたことが推定できる。

もちろん、自然言語処理の観点からいえば、本来「燃費のよい」と「低燃費」は同じ意味を指さねばならない。しかしながら、ここでは意味解析を全く行わず簡単に意識推定を行うための一つの近似として、文字列のマッチングのみに基づく文脈情報を用いるものとする。

c) アクセスページ 検索型のアクセスでは、www利用者とページ作成者の意図に相違が生じることがある。したがって、検索時に用いられた検索語をページ毎に把握

²⁾ あるサイトで公開されている情報が、あらゆる分野を一様に網羅している場合、web検索語の解析によりwww利用者総体の意識が、www利用者による検索語があらゆる分野に関して一様に分布している場合、そのサイトで公開されている情報の傾向が分かることになる。しかしながら、このような状況は現実的ではなく、積集合としての傾向のみをここでは取扱う。

することは作成者にとって重要なものとなる。ここでは指定したwebページに対して用いられた検索語、指定した検索語をもとにアクセスされたwebページを表示する機能を解析ツールに持たせることとする。

3.3 解析ツールの実装 あるホストに対するwebアクセスに関する参照元リンクの記録(referer_log)を対象に、前節に述べた検索語解析機能は、文字列処理との親和性のよいPerlにより記述したスクリプトにより実現する。ただし、単語情報、文脈情報についてはデータベース化を行い、定期的にreferer_logを参照し自動的に更新しておくことで処理の高速化を図る。また、これらの情報を人間可読形式としたHTMLファイルと同じホストに生成することとする。特定のwebページ、検索語についての対話的なreferer_logの解析はcgi(Common Gateway Interface)により実現し、web閲覧とのシームレス化を図るが、wwwサーバとなるホストの負荷を軽減するために処理を分散化し、cgi処理専用のホストを用意する。このホストへは、更新されたreferer_logおよびデータベースが定期的に転送される。

4. 機械工学科設置の解析ツールによるケーススタディ

4.1 構築したweb検索語解析ツールの概要 ここでは、機械工学科(<http://www.mech.nara-k.ac.jp/>)以下で公開されているwebページへのアクセスを対象とする。wwwサーバとするホストはSun MicroSPARC II(SunOS 4.1.4-JL)であり、httpdにはApache-1.2.6を用いている。参照元リンクは、referer_log_moduleを用い、NCSA互換形式でreferer_logファイルに記録している。検索語解析ツールはPerl(5.003)スクリプトにより記述した。また、cgi処理を行うホストはDell Dimension XPS(Linux 2.0.36)であり、cgiはPerl(5.004_04)スクリプトにより記述している。検索語は一日に一度データベース化され、wwwホストからcgiホストへ転送される。

検索単語および文脈の頻度は、検索語のデータベース化と同時に更新される

<http://www.mech.nara-k.ac.jp/~hira/search.html>

によりwww上で確認できるようにしている。このページには、新出検索文字列および累積検索単語の頻度ランキング、検索単語毎の文脈情報等が示される。また、ページ毎に検索語を確認するためのフロントエンドページとして

<http://www.mech.nara-k.ac.jp/~hira/keycheck.html>

を、検索単語毎に文脈およびアクセスされたページを確認するためのフロントエンドページとして

<http://www.mech.nara-k.ac.jp/~hira/keyword.html>

を用意している。これらのフロントエンドページから発行されるcgiリクエストには、cgiホストが応答する。

4.2 公開されているwebページの概要 機械工学科wwwホスト(<http://www.mech.nara-k.ac.jp/>)が保持しているページは、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・研究室)、他学科等へのリンク、アンケート、専攻科概要等からなる。さらに、授業の課題として学生が作成した一般的話題を含むページを準備している。これらの総ページ数は260であり、そのうち、学生が作成したものが約130、教育課程関係が約30、研究室紹介のページは12である。

4.3 ケーススタディ 構築した解析ツールによるケーススタディの期間は、1999年1月18日から8月18日までの7ヶ月間である。この間にみられた検索文字列は23520件であり、平均110件/日の検索型アクセスがあったことがわかる³。検索文字列を単語単位に分解した場合、31145の検索単語が見られ、ひとつの文字列には平均1.3単語が含まれている。

a) 検索語頻度による解析 表1に単語単位の累積頻度が上位10位までの検索単語および頻度を示す。頻度には累積値、1日間の頻度、2週間おきの順位変動を含んでいる。なお、ここではHTMLファイルを示さないが、以降に示す結果は、全て8月18日に生成された

<http://www.mech.nara-k.ac.jp/~hira/search.html>

および派生するページ上で閲覧可能なもののみによっており、2次の加工は行っていない。ただし、本報では企業等の営利団体名および商品名、タレント等の個人名は省略し、記号<>で囲み記している。

表1 検索単語および頻度(99年8月18日)

検索単語	頻度 (1日間 / 順位増減)
1 <自動車商品名>	1038 (+2 / -)
2 心理テスト	782 (+10 / -)
3 ロボット	653 (+1 / -)
4 <タレント個人名>	645 (+4 / +1)
5 <工作機械企業名>	639 (+3 / -1)
6 新幹線	603 (+1 / -)
7 バイク	570 (+3 / -)
8 鉄道総合研究所	558 (+3 / -)
9 <自動車企業名>	535 (+1 / -)
10 <自動車企業名>	493 (+0 / -)

³ 実際のアクセスには、検索文字列が不明なものや、キャッシュによるもの等が含まれるため、ここで述べる件数は正確ではない。本報では相対頻度のみが重要となるので、これらの不確定要素については考慮しない。

単語頻度だけではなく、単語が文字列に含まれる文脈を考慮した場合には、〈自動車商品名〉:1062, 心理テスト:852, ロボット:1085, 〈タレント個人名〉:646となる。検索文字列に含まれる頻度の最も高かった「ロボット」に注目した場合、表2 (上位10のみを示す) に示す文脈が得られた。表中、見出し語と一致する部分は記号 - で略記している。これによればwww利用者の関心は、「ロボット」そのものにあるのではなく、「人間型」、「サッカー」、「歩行」と関連付けられた「ロボット」にあったことが推定できる。

表2 文脈に注目した頻度

「ロボット」を含む検索文字列	頻度
1 —	128
2 人間型—	83
3 <企業名カタカナ>+—	51
4 —サッカー	40
5 <企業名漢字>+—	38
6 —歩行	31
7 サッカー—	29
8 —サッカー	28
9 二足歩行—	26
10 歩行—	24

以下 255 種の検索文字列

先頭または単独で「ロボット」が含まれる: 444 語

「ロボット」を含む文字列: 1085

b) 対話的解析 前項では、本サイトで公開されている情報全てに関する検索語をもとに解析を行った。さらに本ツールでは、特定の検索語、ページについて対話的に解析を行うことが可能であり、検索語については、

<http://www.mech.nara-k.ac.jp/~hira/keyword.html>

およびページについては

<http://www.mech.nara-k.ac.jp/~hira/keycheck.html>

が準備されている。表3は「奈良高専」および「機械工学」を検索文脈に含んでいた文字列を示す。この結果からは、「奈良高専」と専攻科を関連付けた検索 (約1割) や、「機械工学」に対しては「電気自動車」や「地球環境」、「リサイクル」などの環境問題を関連付けている検索が多くみられ、www利用者の意図が推定可能であると考えられる。誘導型のwebアクセスにおいて要となる機械工学科トップページおよび専攻科トップページについての検索語を表4に示す。これらのページに対する検索型のアクセス数は少ないものの、www利用者の、「機械工学科」、「専攻科」に限らないwww情報への様々な関心がうかがえる。

表3 特定の単語を含む検索文字列

「奈良高専」を含む検索文字列	検索型アクセス数: 83
—: 64, —+専攻科: 3, —専攻科: 3, —+化学: 2, —同窓会: 2, —機械: 2, —専攻科+編入+進学する: 2, —+工学科+学科+機械工, —+機械工学, 研究紀要+—, —専攻科+進学する	
先頭に含まれる文字列: 82	
「機械工学」を含む検索文字列	29
—+電気自動車: 5, —会: 3, —+教育+中学: 2, 地球環境+—: 2, —科: 2, —+リサイクル: 2, —実験: 2, —+自動車+電気自動車: 2, —+ドリル, —+研究とは, —+バイク, —, —+エンジン, —+自動車, 奈良高専+—, 福井大学+テスト+—, 舞鶴高専+—科	
先頭に含まれる文字列: 24	

表4 特定のページに対する検索文字列

機械工学科(/index.html)	検索型アクセス数: 30
奈良高専: 8, 高専: 2, 奈良高専機械: 2, nara-k: 2, icem: 2, 技官+求人, <技官名(姓)>, 機械工学科, icem+cfid, 機械工学, 機械技術者, ロボット工学+専門学校	
専攻科(/~ac-wada/index.html)	33
高専: 6, 学位授与機構: 3, 奈良高専専攻科: 3, 奈良高専: 2, 機械工学+教育+中学: 2, 奈良高専+化学: 2, 資格救急救命士養成学校+養成する+本校, 高専+専攻科+教育課程, 化学技術者, 高専+専攻科, (高専+or+高等専門学校)+and+専攻科+and+学位授与+板橋高等職業技術専門学校, 高専教育+あり方, 高専+専攻科, 工学技術者倫理批評, 新居浜高専, 新居浜高専+専攻科, "高専", 専攻科, 高専専攻科, 技術	

c) 機械工学関連分野への関心 機械工学科webページでは、教育課程に関する情報としてシラバス³⁾、研究テーマに関する情報として研究室紹介⁴⁾を掲載している。教育・研究の場としての学校学科の情報公開の手段としてwwwを位置付けるにあたっては、これらのページへの関心および検索時にどのように見えているのか、等の把握が重要である。ここでは、ページ毎の検索語を示すことで、学科目、研究テーマ等機械工学関連分野への関心を調べる。表5, 6はシラバスを記載したページへの検索語、表7は研究室紹介のページへの検索語である。ただし表7の対象としたページは、1999年4月初旬に作成したものであり、累積アクセス数は、表5, 6のシラバスと単純には比較できない。

表5 シラバスwebページへの検索型アクセス(1)

機械設計製図(/syllabus/1220.html)	頻度: 150
機械設計: 53, 機械製図: 42, 製図: 8, 製図+書き方: 5, 設計製図: 3, 設計+製図: 3, 機械要素: 3, 機械+設計: 2, 一: 2, 設計書の書き方: 2, 製図の基礎: 2, 製図書き方, 機械設備+製図+設計+技術者, 大阪製図, 製図+機械+書き方, 歯車減速機+図, 機械+設計+図面, 歯車技術計算, 設計書, 機械要素+歯車, 機械要素設計, 簡単構造+設計, <教官氏名>, 設計書+書き方, 機械製図基礎, 製図+基本, 機械基礎構造, トレース+図面+製図+cad+trace+図+部品+設計+習得する, 空間認識能力, 機械設計製図+試験, 機械設計+北海道, 設計&図形+製図+機械+部品+図面+図, 機械製図+書き方, 設計製図+機械+フリーソフト, 機械設計計算+ベルト長さ, 配管フロー図の書き方+設計+製図	
振動解析(/syllabus/1213.html)	70
一: 62, 振動+解析: 2, 振動: 2, 造船機械振動, <教官氏名>, 解析手法, 振動+解析	
材料力学(/syllabus/1207.html)	57
一: 26, 断面二次モーメント 10, 曲げ応力: 3, 不静定問題: 2, 断面2次モーメント, (一), 一+and+機械, モーメントの計算, モーメント計算, 一+解析, 計算方法, モーメント+一+解析法+公式+金属+応力+耐久+疲労+強度+荷重+破壊+変形+弾性, 一+計算方法, モーメント計算, 基礎一, モール円+解説, 一+たわみ, 断面二次, ボルト+曲げ強度+応力, 断面2次モーメント+意味	
工業外国語(/syllabus/1216.html)	46
工業英検: 26, 工業英語: 17, 工業英語+, 工業英検2級, 英語+技術者にとって	
機械工作法(/syllabus/1205.html)	40
一: 15, 工作機械: 5, プレス加工: 4, 研削理論: 2, 鍛造+プレス: 2, 工作: 2, 一+切削, 工作法, 板金+加工条件, (機械加工+or+かしめ+or+ハイスピン)+not+(検索結果+or+lsi)+not+(ゴルフ+or+電子), 鍛造法, 板金+理論, 板金加工, 機械掘削, 機械加工+ブローチ, 機械工	
機械設計法(/syllabus/1217.html)	20
一: 8, 機構学: 4, 機械設計: 3, <教官氏名>, 機構学について, 動的問題, 機械+and+設計, 機械設計の基礎	
計測工学(/syllabus/1212.html)	20
一: 13, <教官氏名(非常勤)>: 4, <教官氏名(非常勤)>, <教官名(姓, 非常勤)>, 状態量	
応用設計製図(/syllabus/1221.html)	19
機械設計: 3, <教官氏名(非常勤)>: 3, 設計管理: 2, <教官氏名>: 2, 機械部品+製図, 製図の基礎知識, 設計+製図, 機械製図, トレリスベラング設計+商品値段表, jis 機械製図+面の指示記号, cad+製図+cam, 応用設計, コスト意識+設計, 機械設計+製図	
熱工学(/syllabus/1208.html)	19
一: 8, 理論熱効率: 2, <教官氏名>, 熱の性質, 熱機関工学, <教官名(姓)>, 熱機関, 基礎, 蒸気性質, 蒸気の性質, <教官氏名>	
生産システム工学(/syllabus/1214.html)	17
一: 3, 生産管理について: 3, <教官氏名(非常勤)>: 2, 生産計画+設計+策定する, 経営工学+製造工程管理技法, 生産管理+基礎知識, 生産管理手法, システム工学, 生産システム+レイアウト設計, 生産管理+管理手法, 生産管理システム+製造, <教官氏名(非常勤)>+助教	

機械工学科開講科目に対しては、設計・製図、振動解析、材料力学、工業英語、加工に関する情報が多く検索されている。振動解析および工業英語をのぞけば低学年配当の基礎的科目であり、これらに関する情報への関心

表6 シラバスwebページへの検索型アクセス(2)

流体工学(/syllabus/1209.html)	16
一: 8, 流体: 6, <教官氏名>, (流体)	
設計製作(/syllabus/1231.html)	15
??設計: 3, 設計資料, 工作実習+図面, 機械設計, 創造設計, 機械設計+製作, 設計+製作, 設計方針とは, ??設計, 機械製図, 機械設計+製図, 設計能力, iso 事例集+設計+製図+図面	
材料力学特論(/syllabus/1252.html)	14
ひずみエネルギー: 3, 一: 2, 座屈, 材料力学, 材料力学+エネルギー法, 材料力学+公式, 公式+材料力学, 設計公式, エネルギー法, 一I, 弾性ひずみ	
新素材論(/syllabus/1204.html)	14
新素材: 8, 新素材について: 3, 動向: 2, 非鉄金属材料	
機械工作実習(/syllabus/1230.html)	13
一: 3, 工作機械: 2, 機械工作: 2, 工作: 2, 工作実習, 機械加工+機械工作, <教官氏名>, 工作機械+タック	
電子工学(/syllabus/1211.html)	10
一: 3, 動物と機械: 2, <教官氏名(非常勤)>, 感覚器官, 電子回路+実例, 人間工学, アナログ+デジタル+優位性	
情報処理(/syllabus/1201.html)	8
<教官氏名>: 3, c 言語処理プログラミング, <教官氏名>, 一階微分方程式+数値+解法, c 言語+文法+プログラミング, c 言語+習得+プログラミング+習得する,	
塑性工学(/syllabus/1253.html)	8
一: 6, 塑性, 一+力学	
機械基礎力学(/syllabus/1206.html)	7
基礎力学: 4, 機械基礎: 2, 機械力学つりあい	
材料強度学(/syllabus/1203.html)	7
材料強度: 3, 一: 3, 材料強度+	
機械工学実験(/syllabus/1232.html)	6
一: 2, 実験, <教官名(姓)>, <教官氏名>, 工学関係のレポート	
エネルギー変換工学(/syllabus/1256.html)	6
一: 2, 化石燃料+エネルギー変換, エネルギー供給, エネルギー変換, エネルギー	
熱移動論(/syllabus/1255.html)	6
熱物質移動論: 3, 熱移動: 2, 熱+物質移動論	
制御工学(/syllabus/1215.html)	2
一: 2	
電気工学(/syllabus/1210.html)	2
一, 電気の基礎	
基礎材料学(/syllabus/1202.html)	1
一	

が高い傾向は既報)と同様である。振動解析、工業英語については高学年配当科目ではあるが、内容的には設計・製造における機械技術の基礎的分野を占めるものである。また、用いられている検索語についても基礎的事柄が多く、www利用者の関心は、基礎的分野に重きをおいていると推定される。

表7にみられた頻出検索語「トライボロジー」は、本科開講科目ではとりあげられていないために、シラバスへのアクセス(表5, 6)では現われていない。しかしながら、トライボロジーで取扱う潤滑問題は、機械設計分野において最も重要な問題のひとつであり、実際、web検索においても関心が持たれているテーマであることがわかる。また、他の検索語からも、材料力学、加工等につ

表7 研究室紹介webページへの検索型アクセス

検索文字列: 頻度
トライボロジー: 13, 流体潤滑: 8, 混合潤滑: 4, 滑り軸受: 3, 流体潤滑+トライボロジー: 2, トライボロジー+流体潤滑, 焼付き, 摩耗+なじみ, <教官名(姓)>, 潤滑+焼付き, 潤滑, 接触理論, 機械の寿命
遅れ破壊: 7, アクリル板+強度: 6, アクリル+強度: 3, 遅れ破壊試験, アモルファス金属
研削加工: 2, 鋳鉄切削, 切削加工, 工具+摩耗, 研削+黒鉛, 球状黒鉛鋳鉄, 切削加工+研究, クリープ+フィード+研削
冷間鍛造+管, 曲げ加工機, 微細穴, 鍛造+工法, ビーディング+ビード加工, ビード加工, <教官氏名>, 介護用ロボット
粗さの影響, 微細穴, 高温割れ, <教官氏名>
伸長粘度: 2, 伸長
制御工学+動特性, ヘッドスライダ
概念設計+支援, 概念設計

いて、シラバスに対するweb検索と同傾向の関心が現われていると考えられる。ここで示したweb検索語の解析による機械工学分野の教育・研究内容への関心の傾向は、企業へのアンケート結果から、設計においてよく使われる材料力学系の項目に対する要望が強いとの分析⁹⁾と対応するものであり、本手法による意識推定の妥当性を示していると考えられる。

5. 構築したツールの応用可能性

5.1 概念形成 ある概念とそれに関する語彙からなる体系をオントロジーと呼び、構造化された知識ベース構築の観点から盛んに研究されている。たとえば、表2に示したような「ロボット」に関する検索文字列および頻度情報をもとに意味処理を含む適切な言語処理を行うことで、「ロボットとは」といった概念の自動形成が可能となると考えられる。

5.2 予断を与えないアンケート 一般にアンケートによるニーズの把握には、設問者の恣意性の問題が生じる。これに対して本ツールによる意識推定は、被験者に予断を与えることなく得られたものであり、その結果には設問者の意図が入り込む余地がなく、予断を与えないアンケート手法としての利用が可能であると考えられる。

5.3 アイデアプロセッサ 本報では、検索語のうち主に多数を占めたものを対象として議論を進めてきた。少数の検索語については、一般に組合せて用いられない単語等が見られる。また、サーチエンジンの特性に依存するが、ページの内容と完全には一致しない検索語を使った場合にもアクセスが可能なることから、ページ作成者の

意図を超える語彙が得られている。たとえば「ロボット」に関する文字列には、「非ノイマン型コンピューターロボット」、「天然ロボット」などが見られたが、現段階では著者にとって想像のつかない概念である。このような意外性を持つ語彙の提示が可能な本ツールは、不確かなものも含む多様な情報提示による創造支援ツール¹⁰⁾としての利用可能性を示していると考えられる。

6. おわりに

本報では、web検索語に注目した解析ツールを構築し、www利用者の意識推定をケーススタディにより試みた。なお、ここに示したツールは、

<http://www.mech.nara-k.ac.jp/~hira/search-tool/>
において公開している。

参考文献

- 1) 平 俊男, 武藤武士, 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-, 高専情報処理教育研究発表会予稿集, (1998), pp.144-147.
- 2) Berners-Lee, T., Masinter, L., and McCahill, M., "Uniform Resource Locators," Sec.2.2, <ftp://ds.internic.net/rfc/rfc1738.txt>, (1994).
- 3) 奈良高専, 平成8年度シラバス (機械工学科), (1996).
- 4) 日本機械学会関西支部, Mechavocation 1999, (1999).
- 5) 奈良高専, 多面的評価を取り入れた教育内容・方法の改善 (国専協 教育方法改善共同プロジェクト) 最終報告書, (1998), p.123.
- 6) Chakrabarti, A. and Thomas, P.B., "An Approach to Functional Synthesis of Mechanical Design Concepts: Theory, Applications, and Emerging Research Issues," *AI EDAM*, Vol.10, No.4, (1996), pp.313-331.

ISSN 0387-1150

研 究 紀 要

第 35 号

平成 11 年度

奈良工業高等専門学校