

文書検索システム ExtractRequest における 用語分析マップによるフィードバックの評価

篠原靖志

(財)電力中央研究所

文書検索システム ExtractRequest は、キーワードの配置によってキーワードの優先度を視覚的・即応的に制御する検索機能と、利用者が選択した文書からキーワード候補を抽出して用語分析マップとして提示する補助機能とを持つ。用語分析マップでは、文書中に現れる用語の適合性および用語のデータベース中での使用頻度というごく基本的な情報を視覚的・即応的に利用者にフィードバックすることで、利用者が対象データベースでの対象テーマにかかわる用語の利用のされかたを把握して、適切なキーワードを選択することを支援する。本稿では、本システムのフィードバックの効果についての実験結果について述べる。実験により、用語分析マップにより、検索結果が改善する効果が確認された。

Evaluation of Keyword Selection Support by Map of Candidate Keywords in a Document Retrieval System ExtractRequest

Yasushi Shinohara

Communication and Information Research Laboratory
Central Research Institute of Electric Power Industry
2-11-1 Iwado-kita, Komae-shi, Tokyo, 201 Japan

ExtractRequest is a document retrieval system with the support of users to recall appropriate query keywords. It feeds back basic information of words visually and spontaneously such as word frequency and word relevance based on the user's input on the relevance of several documents to promote the user to grasp the usage of words related to the theme in the target database. This paper describes the basic features of ExtractRequest and the evaluation of its functions by the experimental comparison with conventional document retrieval systems

1. はじめに

一般の利用者が従来のデータベース検索システムを利用し難い大きな理由として、検索に適

切なキーワードを探し出すことが難しい点が挙げられる。

そこで、われわれは、検索条件入力用の視覚的で即応的なインタフェースによって、検索の

試行錯誤を行いやすくすると同時に、利用者に対象データベースでの用語の利用状況の情報をフィードバックすることで、適切なキーワードの選択を促す用語分析機能を持つ文書検索システム ExtractRequest を開発した[1]。

本稿では、ExtractRequest の設計の考え方、システムの概要を述べた後、各機能の効果についての実験・分析を行った結果について述べる。

2. ExtractRequest の設計の考え方

現在もっとも広く使われている文書検索システムとして、利用者が与えたキーワードと文書中の用語との照合による全文検索システムがある。サーチャなどの専門家のみでなく、検索の専門家でない人でも容易に利用できるためには、われわれは、以下の2つの問題があると考えた。

第1に、利用者は、自分の検索したい文書群に共通に現われる用語を検索キーとして入力する必要がある。しかし、検索したい文書に現われる用語を事前に適切に想定することは多くの利用者にとっては困難である。検索の過程で、適切なキーワードを見つけることを支援する必要がある。

第2に、複数の関連の深い用語の OR 操作、AND 操作などを行なう必要がある場合がしばしば生じる。AND、OR などによる論理型操作は形式的で直観的ではないため、使いこなすことが困難である。

第1の問題、即ち、適切な検索用のキーワードが事前にはよく分からないということから、この第2の問題は特に重要となる。適切なキーワードが不明であるために、最終的に欲しいものが見つかるまでの検索の過程では、検索の結果を見て、適切なキーワードを推定し、適当にキーワードを加えたり、入れ替えたり、キーワードの組み合わせの条件を変えたりするという、試行錯誤を繰り返す必要がある。

従来の論理型のキーワード検索は、ゆっくり

考えて複雑な条件を記述して使うには向いているが、検索結果を見て条件を何度か追加、修正しながら目的の検索結果に近づいていくような使い方は難しい。われわれは、検索の非専門家を対象にした場合、後者の試行錯誤的使い方が自然であり、そのために、感覚的かつ迅速に条件が入力、修正できる記述方法とユーザインタフェースが必要と考える。

そこで、ExtractRequest では、キーワード加重による検索方式を採用し、このキーワードの追加、削除、加重（優先度）の変更を、視覚的インタフェースを介して、利用者が直接操作することができるようにしている。さらに、利用者の直接操作は、即座に、検索結果に反映させることで、即応性を実現している[1]。

このような視覚的・即応的インタフェースを用いることで、利用者による感覚的かつ迅速な条件の入力・修正が可能となり、第2の問題は解決すると期待できるが、第1の問題、すなわち、適切なキーワードをみつける過程の支援の問題は依然として残る。

利用者は、試行錯誤的な検索過程で現れるさまざまな検索条件での検索結果を見て、適切なキーワードを推測をする必要があるが、この推定は容易でない。例えば、あまりに一般的な用語をいれると検索するテーマに関連しない文書が多数現われるし、あまりに特定の用語を入力すると、ごく一部の文書群しか収集できない。また、同一テーマに関する用語でも対象とするデータベースによりその利用が異なり、適切な用語選択を行なうためには、対象データベースでの用語の使用傾向が反映される必要がある。

ここでは文書中の用語とのキーワード照合による全文検索システムを対象としているので、適切なキーワードは、既に見つかった好ましいと思われる文書に現れていると期待される。

そこで、ExtractRequest では、利用者に見つけたい文書に関連すると思われる文書を検索結果の中から複数選択してもらい、それらの文書に共通して現れる文字列を、検索用キーワー

ドの候補とする。また、キーワードの適切さを判断するための補助情報として、選択文書でのキーワードの出現状況（キーワードの適合度）および、DB 中でのキーワードの出現頻度の2つの情報が有効と考え、2つの情報を2軸とする平面に、キーワードの候補を配置する。

適切なキーワード選択などについては、適合フィードバック(relevance feedback)から派生する様々な研究がなされてきた[2,3]。適合フィードバックは、利用者が検索結果の文書に対して、その適合度を判定することで、適合した文書をより検索しやすいように、検索式を自動的に修正する方式である。このような自動化は、一方で、検索過程がブラックボックス化することで、利用者が検索された結果の特性を把握・制御しづらい可能性がある。また、欲しい文書に関連しているか否かという適合性の判断も、曖昧性を持ったり、検索の過程の中で変化をする。特に、適切な文書が見つからない場合は、許容の範囲は緩くなり、場当たり的となる。

適合フィードバックにおける検索条件の変更候補（キーワード）を利用者に示して、その選択を促すなど、適合フィードバックをよりインタラクティブにする試みがなされている[4]が、試行錯誤的に検索を進めている過程では、利用者の揺れる適合度の判定によって直接、検索式の変更を行うよりも、適合度判定情報を基に利用者が対象テーマについてのデータベースの特性を把握することを支援することが、欲しい情報を速く見つけ出すことに役立つ可能性がある。

ExtractRequest では、利用者からの検索結果文書に対する適合度の判定情報を基に、利用者に検索式（キーワード加重方式の場合、キーワードと加重）を変更するために必要な基礎的情報を視覚的に、かつ、即座に提示することで、利用者が直観的に検索に適した用語を把握していくことを支援しようとしている。

3. ExtractRequest のシステム概要

ExtractRequest は、検索画面と用語分析画面の2つからなる。

検索画面は、キーワードの優先度を設定する検索条件部と検索結果表示部からなる画面である。キーワードの優先度（加重）は、検索条件入力部にキーワードを配置することで決定される。左端に配置されるほど、優先度が高く、大きい正の加重を持つ。右端に配置されるほど優先度は低く、大きい負の加重を持つ。

利用者がキーワードを入力すると、キーワードの頻度に応じて、頻度の少ないものほど高い優先度を持つように自動的に配置される（自動配置機能）。加重は、頻度の指数オーダで減少させている。利用者は、検索結果を見て、キーワードの配置を変えることで、優先度を適宜変更することができる。キーワードが入力されたり、キーワードの配置が変更されると、即座に、検索結果が更新される。

検索結果部は、キーワードを含む文書が、文書に含まれるキーワードの加重の和によって降順に整列されている。このため、優先度の高いキーワードを多く含む文書ほど、上位に現れる。検索結果は、文書のタイトルによる表示と、含まれる用語のリストによる表示の2つのモードを持ち、簡単に切替えることができる。

用語分析画面は、キーワード候補となる文字列を提示する。検索画面の結果表示部で利用者が適当と考える文書を複数選択して、用語分析ボタンを押すことで切り替わる。

キーワード候補となる文字列は、選択された文書群の2つ以上に共通に現れる最長共通文字列のうち、一定出現頻度以下のものである。従来の適合性フィードバックでは、文書ベクトル空間を構成するキーワードとして使用されない可能性が高い、「第二電電」と「電電公社」の共通部「電電」などもキーワード候補となる。

用語分析マップは、これらのキーワード候補となる用語を、選択した文書群での出現状況とデータベースでの出現状況との2軸を持つ2次元平面に配置したものである。

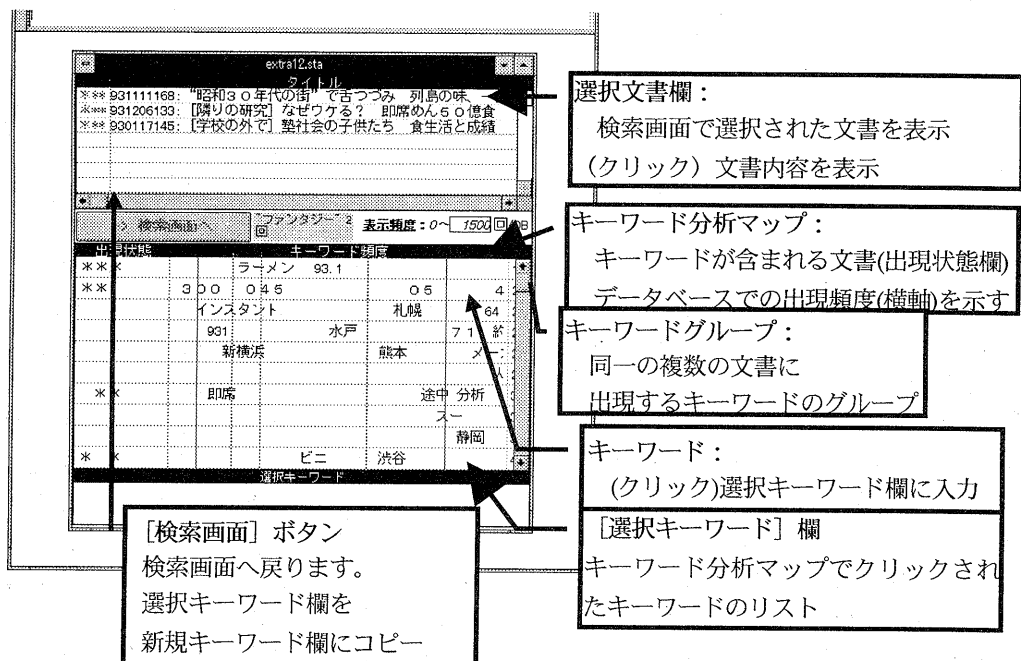
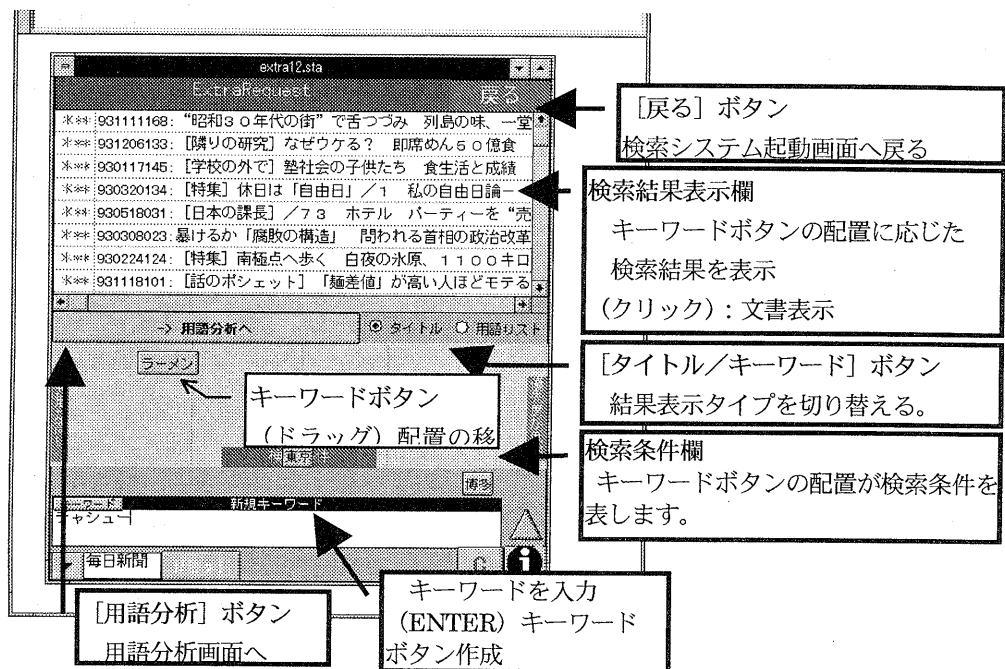


図1. ExtractRequest の2つの画面 (上: 検索画面 下: 用語分析画面)

横軸は、文書データベース全体の中での用語の出現頻度（文書本数）の対数である。用語の出現頻度は、厳密な出現頻度ではなく、推定値を使用する。

縦軸は、選択した文書群での用語の出現状況を示し、「*_*」は1番目と3番目の選択文書に用語が現れていることを示す。出現状況は、多数の文書に現れるほど、上位に並べられている。（厳密には、適切な文書を選択のみでなく、不適切な文書も選択も許しているため、適切な文書により多く出て、不適切な文書に現れない出現状態ほど上位に並ぶ。）

利用者は、用語分析画面上の用語をマウスでクリックすると、最下段の選択キーワード欄にキーワードがプールされ、検索画面ボタンで、検索画面に移動すると、検索条件部にキーワードが配置される。

本システムは、サーバ・クライアント方式を取り、サーバは DEC ALPHA Station (330 MHz) 上で、クライアントは Windows95 上で稼動する。100M バイト、5 万件程度の大規模データベースでも、検索は大半 1 秒～数秒以内に終了する。用語分析も数本の文書なら瞬時に、50 文書程度を選択しても数秒以内に終了する。

4 評価実験の目的と方法

われわれは、ExtractRequest を他のいくつかのシステムと比較実験することで、各機能の有効性を検討した。

本章では、評価実験の目的と方法について述べる。

実験の目的は2つある。

第1は、ExtractRequest 全体での検索効率の検討である。このため、代表的な従来型の検索システムとして WAIS-J、および、AND-OR 型検索システムとの比較を行った。

第2は、視覚的・即応的インタフェースの効果と、用語分析マップによるフィードバック

の効果の検討である。このため、ExtractRequest から用語分析機能を使用できなくしたものを VisualWeight と名づけて、VisualWeight と用語分析機能を使用できる ExtractRequest との比較を行う。ExtractRequest の用語分析では不適切な文書を選択しても良いが、初めての利用者には、「○」「×」「無印」の選択となり煩雑となると予想したので使用できないようにした。

実験では、検索の非専門家である男女 11 名を対象とした。使用した文書群は、文書件数 600 件の情報検索システム評価用標準問題集(以下、BMIR-J1)と、毎日新聞東京版朝刊 1 年分 (93 年) 5 万件、約 100M バイト (以下、大規模 DB) である。用語マップで表示する用語の最高頻度は、BMIR-J1 で 500 回、大規模 DB で 1500 回とした。

問題は、BMIR-J1 の問題 60 問の内 15 問を選択した。大規模 DB については、BMIR-J1 を参考に 36 問を作成した。BMIR-J1 では、A,B,C の 3 段階にランクづけされており、A がもっとも関連が深い。大規模 DB 用の正解文書は、被験者の検索結果から複数の評価者で抽出し、BMIR-J1 と同じく、A,B,C の 3 段階で整理した。

1 問当たりの制限時間は、5 分 (一部 10 分) とし、その間数回の検索をして良いとした。

5 分析に用いた評価基準

われわれは実験の評価基準に従来の検索精度 (precision) と呼び出し率 (recall) 以外の新たな評価基準、重み付き得点率を考案し利用している[5]。

検索精度は「検索結果中の正解文書数/検索結果の文書数」、呼び出し率は「検索結果中の正解文書数/データベース中の全正解数」と定義される。また、全体的な検索結果の評価には 2 つの合成関数が使われることも多い。われわれは、これらの評価基準には 2 つの問題がある

と考える。

1. 正解文書がA, B, Cのようにランク付けされている場合が想定されていない。
2. 検索結果中での正解文書の順位が考慮されていない。正解が先頭に現れても末尾に現れても値に変化がない。

2は、じっくり検索結果の全体を吟味するような検索では重要ではないが、本システムのように、利用者が検索結果を見ながら頻繁に再検索を行うようなシステムでは、検索結果の上位に正解が現れることは重要である。また、WAIS-Jや本システムなど、結果中の順位を変える検索方式（レーティングシステム）では、順序を考慮した評価基準が妥当である。

われわれの使用する指標「重み付き得点率」は下記で定義される。

重み付き得点率＝

重み付き得点／完全正解の重み付き得点

ここで、

重み付き得点＝

$\sum_i$ （順位 i の得点 $\times$

第 i 番目の回答のランクの得点）

- ・ 各順位の得点は、上位1～5位は16点、6位～20位は21－順位、21位以上は0点とした。

- ・ ランクA,B,Cの得点は各々3,2,1点

完全正解＝

完全正解をランクの高い順に並べたもの(降順)

6. 実験結果の分析

6. 1 全体的検索効率

1 試行の中での数回の検索の中で最大の重み付き得点率の平均を表1に示す。これを見ると、ExtractRequest, VisualWeight は、従来型の検索システム WAIS-J,AND-OR に比べて、平均20ポイントほど高い。この差は、VisualWeight と ExtractRequest の差に比べ、大きいことから、VisualWeight と ExtractRequest に共通の視覚的・即応的イン

タフェースの効果が大きいと推測される。

VisualWeight と WAIS-J および AND-OR の比較については、文献[5]に詳しい。

6. 2 用語分析マップの効果

ExtractRequest と VisualWeight を比較すると、BMIR-Jと大規模データベースでは、重み付き得点の優劣が逆転しており、一概には、どちらが良いか言えない。ただし、上記のExtractRequest のデータには、用語分析を使用しなかった場合、および、使用しても用語分析マップから用語を選択しなかった場合が含まれる。

表1 各システム別重み付き得点の平均

重み付き得点	AndOr	WAIS-J	Visual Weight	Extract Request
BMIR-J1	40	43	67	50
大規模DB	39	40	57	62.

表2 用語選択を行った場合

	重み付き得点
BMIR-J1	72
大規模DB	64

そこで、用語分析マップを作成して、その中から用語を選択した場合に限定して、重み付き得点の平均をみると、BMIR-J1、大規模DB いずれの場合においても、ExtractRequest は VisualWeight より5～7ポイント高い値を持つ。このため、用語分析マップからのキーワード選択は、検索結果の改善には、有効であったと判断される。

ExtractRequest を使用した171試行のうち、用語分析マップの使用は104試行で全試行の60%、さらに、その内、用語分析マップから用語を選択したのはそのうちの74%、79試行（全試行の46%）であった。

このように、用語分析マップの結果を利用する機会が少なかった点が、VisualWeight と

ExtractRequest の間で、差が少なかった大きな原因と考えられる。

6. 3 用語分析マップでの提示候補の評価

本節では、用語分析機能を使用した 107 試行で作成された 174 個の用語分析マップについて、提示されたキーワード候補がどの程度有効なものとなっているかを分析する。

検索結果から選択される好ましい文書数は、平均 4.3 本（最頻値は 3 本）である。これらから生成される候補となる用語数は、平均 38 語（最頻区間 10～19（10 単位）、中央値 26 語）である。利用者は、これらの語の中から、適切な用語を選択する必要がある。

候補となる用語を、自動配置機能を使って、直前の検索条件に付け加えた場合、重み付き得点が向上する語（以下、向上語と呼ぶ）数を調べると、平均 5.7 語である。ただし、37 回（用語分析マップ使用回数 174 回の 21%）では、向上語が含まれていない。一方で、10 語以上ある場合も、28 回（用語分析マップ使用回数 174 回の 16%）ある。用語分析マップ中の候補語中の向上語の割合は平均 2 割である。

向上語についてその効果を見ると、向上幅（重み付き得点の差）でみると平均 8 ポイント、最大の向上幅は平均 14 ポイントである。ただし、全候補語でみると、向上幅の平均は、マイナス 39 ポイントであり、無作為に、候補を選択した場合、大きく得点が低下すると予想される。

用語分析マップの位置と向上語の位置の関係を分析すると、向上語が、縦軸の第 1 語群（選択した文書群の中でもっとも多数の文書群に現れる語群）に現れる確率は、54%（95 回）、第 1 群または第 2 群に現れる確率は、67%（138 回）あり、縦軸による用語の優先順位付けは、有効であると推測される。

6. 4 利用者の選択の評価

上記分析によると、用語分析マップでは、検索条件に加えることで重み付き得点率が向上する向上語が 2 割程度含まれる。ただし、無作為に候補とされた用語を選択すると、重み付き得点率は大きく低下する。そこで、利用者の選択がどの程度、有効かを分析した。

用語分析で用語選択を行った前後の重み付き得点の変化を見ると平均 5.9 ポイントの増加が

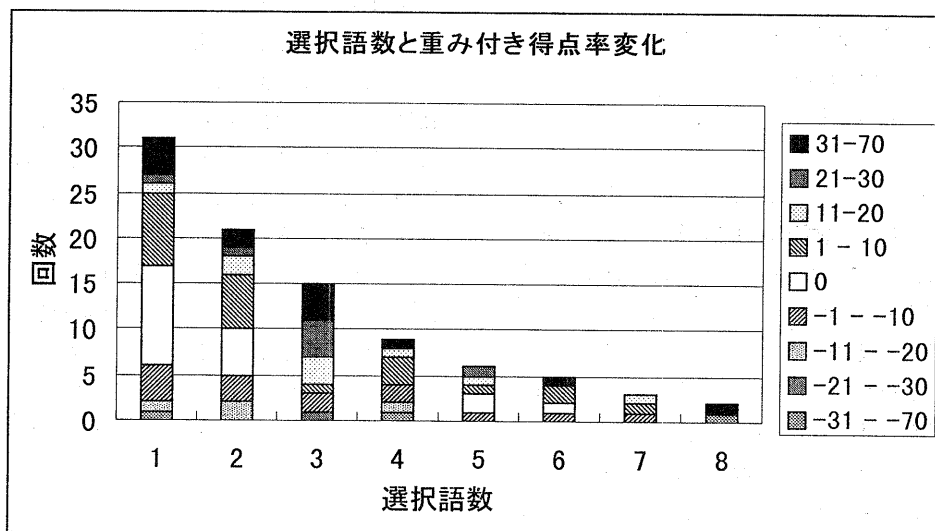


図 2. 選択語数と重み付き得点率の変化

見られる(用語選択を行わなかった場合を含めると3.6ポイント)。

向上語でない候補が選択される場合もあるため、向上語を選択した場合に期待される平均8ポイントには達していないが、利用者は用語分析マップから向上語を選択する能力があることが分かる。

ただし、選択が行われない、または、選択が不適切で、重み付き得点が向上しない場合も全体の2割(用語選択をおこなわなかった場合を除くと3割)ある。

利用者は、分析マップから複数の候補を選択してから、次の検索を行う。この際に、選択したキーワードの個数と向上幅の関係を示したのが図2である。3個程度のキーワードを選択した場合に、向上幅が高くなっている。多数のキーワード(5個以上)を選択しても向上幅は大きくならない。これは、検索の制御がうまく行えていない可能性が強い。誤った制御としては、関連性の低いキーワードが選択されている場合と、多数のキーワードの相互配置が不適切な場合がある。

7 まとめと今後の課題

本稿では、視覚的かつ即応的インタフェースと、用語分析機能を持つ ExtractRequest について、各機能の効果についての評価実験の結果について述べた。

今回の評価実験から、視覚的かつ即応的インタフェースは、試行錯誤的な検索を行いやすくすることで、検索効率の向上に特に有効であることが確かめられた。また、用語分析マップについては、視覚的かつ即応的インタフェースほど顕著な効果は持たないが、検索結果の改善に有効である点が確かめられた。

近年、relevance feedback を活用したシステムは多数出現しているが、ExtractRequest で使用した用語分析マップは、①その時点での

利用者の(アドホックな)適合度評価のみを使用する、②利用者が主体的にキーワード候補を選択する、③キーワード候補の由来について明示的にフィードバックする、④データベースでの出現頻度(推定値)以外、大域的情報を使用しない点などに特徴がある。しかし、他のインタラクティブな relevance feedback 手法と比較した場合の性能、特性の違いについては、今後検討が必要であろう。

現在、ExtractRequest の用語分析機能は、検索条件画面で、キーワードの優先度以外に、キーワードの綴りの正確さも制御しながら検索を行える SearchSpace[6] と統合し、さらに、一般のWEBブラウザでできるように開発した文書検索システム「みつけるぞう」に組み込まれている。

参考文献

- [1] 篠原靖志, "ExtractRequest—利用者への情報開示に基づく検索要求抽出", 情報処理学会SIG-HI, Vol.96, No.21, pp.69-74, 1996.
- [2] David Ellis, New Horizons In Inforamtion Retrieval, 1990.
- [3] Peter Ingwersen, Information Retrieval Interaction, 1993.
- [4] Koeneman, J., et. al., 'A Case of interaction: A study of interactive information retrieval behavior and effectiveness, Proc..of CHI'96, pp.205-212 (1996).
- [5] 堤富士雄, 篠原靖志: キーワードを2次元平面に配置する文書検索システム, コンピュータソフトウェア, 1998 (掲載予定).
- [6] 堤富士雄, "キーワードの2次元配置により検索条件をあらわす全文検索システム: Search Space" in proc of Workshop on Interactive Systems and Software, 1994.