

10-開-10

データベース構築促進及び技術開発に関する報告書  
地域ホームページに係わるリンク集整備および自動巡回システムと  
全文検索エンジンを利用した効果的検索機能の開発

平成11年3月

財団法人 データベース振興センター

委託先 ミネルバ株式会社



KEIRIN

00

この事業は、競輪の補助金を受けて実施したものである。

## 序

データベースは、わが国の情報化の進展上、重要な役割を果たすものと期待されている。今後、データベースの普及により、わが国において健全な高度情報化社会の形成が期待される。

さらに海外に対して提供可能なデータベースの整備は、国際的な情報化への貢献および自由な情報流通の確保の観点からも必要である。しかしながら、現在わが国で流通しているデータベースの中でわが国独自のものは1/3にすぎないのが現状であり、わが国データベースサービスひいてはパランスある情報産業の健全な発展を図るためには、わが国独自のデータベースの構築およびデータベース関連技術の研究開発を強力に促進し、データベースの拡充を図る必要がある。

このような要請に応えるため、(財)データベース振興センターでは日本自転車振興会から機械工業振興資金の交付を受けて、データベースの構築および技術開発について民間企業、団体等に対して委託事業を実施している。委託事業の内容は、社会的、経済的、国際的に重要で、また地域および産業の発展の促進に寄与すると考えられているデータベースの構築とデータベース作成の効率化、流通の促進、利用の円滑化・容易化などに関係したソフトウェア技術・ハードウェア技術である。

本事業の推進に当って、当財団に学識経験者の方々に構成されるデータベース構築・技術開発促進委員会（委員長 東海大学教授 上條史彦氏）を設置している。

この「地域ホームページに係わるリンク集整備および自動巡回システムと全文検索エンジンを利用した効果的検索機能の開発」は、平成10年度のデータベースの構築促進および技術開発促進事業として実施した課題の一つで、当財団がミネルバ株式会社に対して委託実施したものである。この成果が、データベースに興味をお持ちの方々や諸分野の皆様方のお役に立てば幸いである。

なお、平成10年度データベースの構築促進および技術開発促進事業で実施した課題は次表のとおりである。

平成11年3月

財団法人 データベース振興センター

平成10年度 データベース構築・技術開発促進事業委託課題一覧

| 分 野   | 課 題 名 |                                                       | 委 託 先                |
|-------|-------|-------------------------------------------------------|----------------------|
| 社 会   | 1     | 生命保険商品情報データベースプロトタイプ構築                                | (株) 平和情報センター         |
|       | 2     | ビジネス法定文書のサンプルフォームデータベースの構築                            | ジャスネットコミュニケーションズ (株) |
|       | 3     | グループウェアによるISO 9001品質文書管理データベースツールの構築                  | 日本電子開発 (株)           |
| 地域活性化 | 4     | 地域対応型の農業技術情報データベース構築                                  | アシストマイクロ (株)         |
| 技 術   | 5     | 登録ジオコートに基づく世界データベースシステムの開発                            | アジア航測 (株)            |
|       | 6     | 複数の認証局間における認証データベースの有効利用に関する調査研究                      | (株) イフ・アドバタイジング      |
| 地域振興  | 7     | 十勝毎日新聞記事データベース構築                                      | (株) ケーシーズ            |
|       | 8     | 付加価値データベースシステム構築に関する調査研究                              | (株) インタークラフト         |
|       | 9     | 新規産業創出のための業種を融合した企業情報のデータベースの構築                       | (財) 相模原市産業振興財団       |
|       | 10    | 地域ホームページに係わるリンク集整備および自動巡回システムと全文検索エンジンを利用した効果的検索機能の開発 | ミネルバ (株)             |
|       | 11    | 古墳・遺跡の指標となる土器データ検索パイロットシステムの作成                        | 凸版印刷 (株)             |
|       | 12    | 出雲古代遺跡デジタルデータベース構築                                    | (株) 出雲王国             |
|       | 13    | インターネットを利用した四国地域の研究者データベースの構築                         | (財) 四国産業・技術振興センター    |
|       | 14    | 長崎県観光写真素材データベースの構築                                    | 長崎メディアミックス協同組合       |
|       | 15    | 沖縄伝統空手・古武道国際人材リソースデータベース                              | (株) アイエムアイコーポレーション   |

## 目次

|                               |    |
|-------------------------------|----|
| 1. 背景と目的.....                 | 2  |
| 2. 開発体制とスケジュール.....           | 3  |
| 2. 1 開発体制.....                | 3  |
| 2. 2 スケジュール.....              | 3  |
| 3. 検索対象サイトの調査.....            | 4  |
| 3. 1 検索対象サイトの調査.....          | 4  |
| 3. 2 検索対象サイトの決定と連絡.....       | 5  |
| 3. 3 リンク集の公開と連絡.....          | 6  |
| 4. 検索システムの開発.....             | 7  |
| 4. 1 システム構成.....              | 7  |
| 4. 2 動作環境.....                | 8  |
| 4. 3 検索キーワード入力ページ.....        | 9  |
| 4. 4 自動巡回ロボット.....            | 10 |
| 4. 4. 1 巡回方法.....             | 10 |
| 4. 4. 2 更新の有無.....            | 11 |
| 4. 4. 3 リンク追尾.....            | 12 |
| 4. 4. 4 収集ページの制限.....         | 15 |
| 4. 5 検索エンジン.....              | 17 |
| 4. 5. 1 既存の検索エンジン namazu..... | 17 |
| 4. 5. 2 namazu の改造.....       | 17 |
| 5. システムの運用.....               | 18 |
| 5. 1 他の検索エンジンとの比較.....        | 19 |
| 5. 2 利用状況.....                | 20 |
| 5. 3 巡回結果.....                | 21 |
| 6. まとめと今後の課題.....             | 22 |
| 7. 添付資料.....                  | 23 |
| 7. 1 検索対象サイトと URL.....        | 23 |
| 7. 2 画面サンプル.....              | 27 |
| 8. 参考文献.....                  | 32 |

## 1. 背景と目的

インターネットの普及とともに、簡単に多くの情報を得られるようになってきた。その反面、広範囲の中から必要となる情報を探すための手段は、ますます複雑化してきている。

この問題は WWW(World Wide Web) によるインターネットの普及が始まった時から指摘されており、このニーズから Yahoo! Japan(Yahoo) や infoseek 等の広範囲の検索を目的とした商用の検索エンジンが広く普及してきた。

しかし、Yahoo のようなディレクトリサービス型の検索エンジンは事前にページの登録とキーワードの抽出を行うため、対象が限られてしまう。また、時々刻々と変化する情報を網羅することができない。

また、infoseek のように広範囲なホームページを対象とするロボット型の検索エンジンは、大量のデータ更新を行う必要があり、複数のサーバを使って処理を行わなければならない、大規模なシステムが必要となる。それに加え、利用者は、限られた地域や分野に限定した情報を探したい時は、目的とする情報の絞り込みに複雑な手順を要することが多い。

本システムの目的は、これらの問題を解決するアプローチとして、東海北陸地域の情報を利用者にわかりやすくまとめて提供するクリアリングハウスとしての機能を、新しい技術を用いることにより開発することである。

## 2. 開発体制とスケジュール

### 2. 1 開発体制

データベースの専門家による委員会を組織化し、意見交換を行いながら、開発を行った。

#### (1) 委員会

|        |               |       |
|--------|---------------|-------|
| 委員長    | 中京大学教授        | 廣木 守雄 |
| 委員     | 東海北陸データベース懇話会 | 新井 昭二 |
| ”      | ミネルバ 株式会社     | 安藤 真介 |
| オブザーバー | 中部通商産業局産業企画部  | 嶺村 英逸 |

#### (2) ミネルバ 株式会社

企画・プロジェクト管理、調査、システム開発／運用、 検索対象サイト調査

注：検索対象サイト調査については東海北陸データベース懇話会の協力を  
頂いて実施した。

### 2. 2 スケジュール

| 項目              | 7月          | 8月          | 9月          | 10月         | 11月 | 12月                                         | 1月          | 2月          | 3月 |
|-----------------|-------------|-------------|-------------|-------------|-----|---------------------------------------------|-------------|-------------|----|
| 委員会設置           | .....<br>—— | .....       |             |             |     | <div>凡例</div> <div>予定 .....<br/>実績 ——</div> |             |             |    |
| リンク先ホームページ調査、整備 |             | .....<br>—— | .....       |             |     |                                             |             |             |    |
| リンク先ホームページ追加調査  |             |             |             | ——          | ——  | ——                                          |             |             |    |
| 調査・技術研究         |             | .....<br>—— | .....       |             |     |                                             |             |             |    |
| システム設計          |             |             | .....<br>—— | ——          |     |                                             |             |             |    |
| ソフトウェア開発        |             |             |             | .....<br>—— | ——  | ——                                          |             |             |    |
| テスト             |             |             |             |             |     | .....<br>——                                 |             |             |    |
| テストシステム試行運用     |             |             |             |             |     |                                             | .....<br>—— | ——          | —— |
| 報告書作成           |             |             |             |             |     |                                             |             | .....<br>—— | —— |

図 2-1 実施スケジュール

### 3. 検索対象サイトの調査

#### 3. 1 検索対象サイトの調査

東海北陸地区に関連し、企業経営に関して有効な情報を公開するホームページを調査した。

企業経営に関する情報を対象としたのは、この情報が地域に関連する情報として有効であり、利用者にアピールしやすいためである（図3-1参照）。

公的機関を中心に100サイトを目標として対象サイトの選定作業を実施した。企業経営に関して有効な情報である”補助金”，”政策”，”融資”，”経営指導”，”資格”，”採用”，”講習会”の内容を公開しているかどうかを調査したが、内容や更新頻度を吟味した結果、第1回目では68サイト（約150サイト中）にとどまった。そこで、市町村等の自治体まで範囲を拡大し、選定作業を行った結果、最終的に93サイト（約450サイト中）を検索対象サイトとして選定した。

安易に東海北陸地域の自治体をすべて選択すれば、簡単に100サイトを超えてしまうことになる（約300サイト存在）が、実際には観光情報などを掲載する自治体等が多く、市町村等の自治体から対象サイトとして選定したのは16サイトとなった。

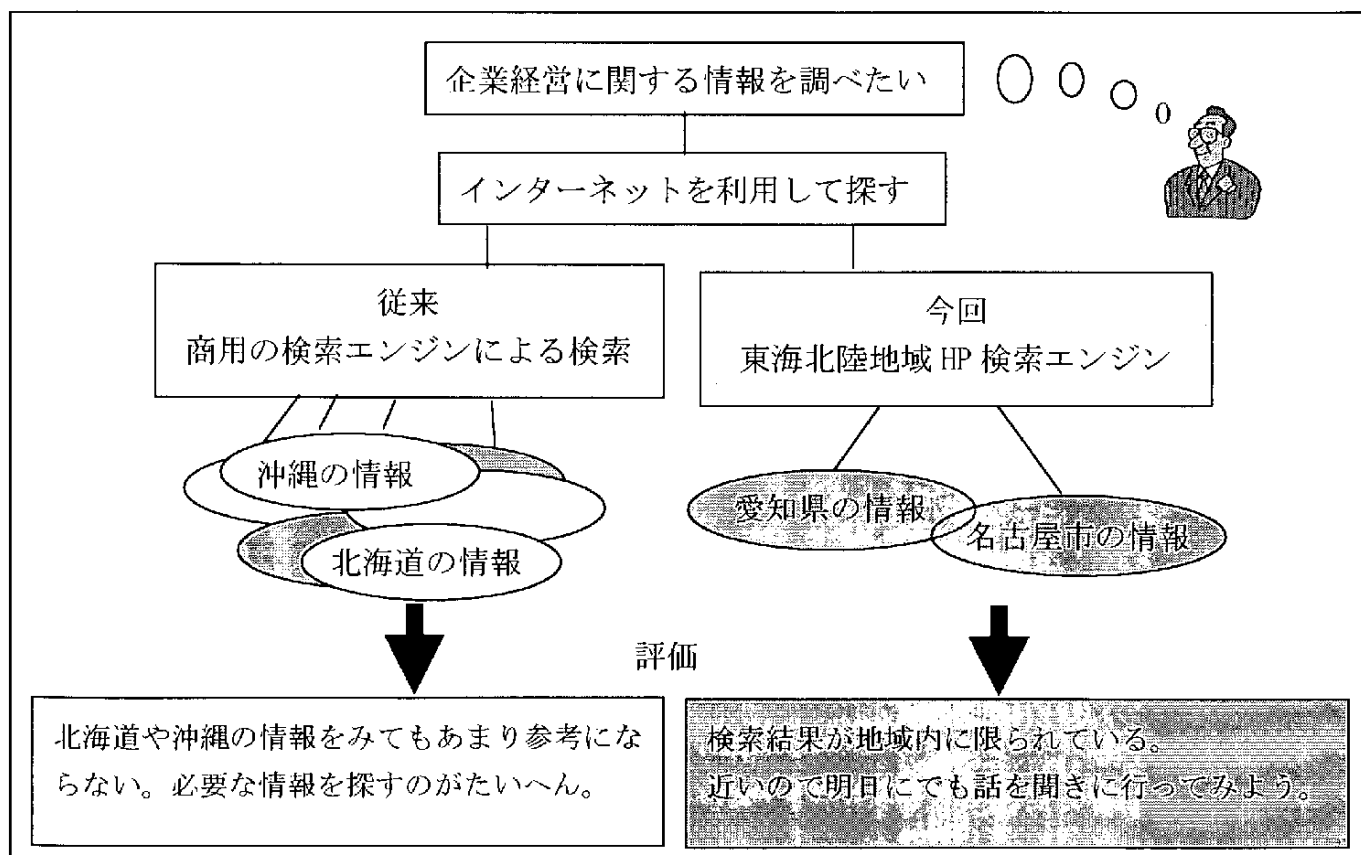
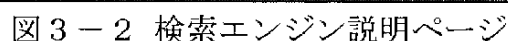


図3-1 地域に関連した企業経営に関する情報の有効性



愛知県庁、岐阜県庁、名古屋商工会議所に直接訪問し、本開発作業の説明を行うとともに協力をお願いした。その他のサイトについては電話、FAX、電子メールを利用して連絡を行うと共に、ホームページによる説明も併せて行った(図3-2参照)。各サイトへの連絡方法と反応は表3-1の通りである。

| 対象サイトへの連絡方法 |    | 連絡に対する反応     |    |
|-------------|----|--------------|----|
| 直接連絡        | 4  | 了解           | 4  |
| 電話、FAX      | 17 | 了解           | 17 |
| E-Mail      | 73 | 了解の返事        | 15 |
|             |    | 了解と判断        | 58 |
|             |    | リンクの拒否       | 1  |
| 連絡したサイトの合計  | 94 | 了解があったサイトの合計 | 93 |



### 3. 3 リンク集の公開と連絡

登録対象としたサイトを分野別に分け、リンク集をホームページ上で公開した(図3-3参照)。さらに、登録先のサイトに対して検索エンジンを公開したことについて電子メールにより通知した。

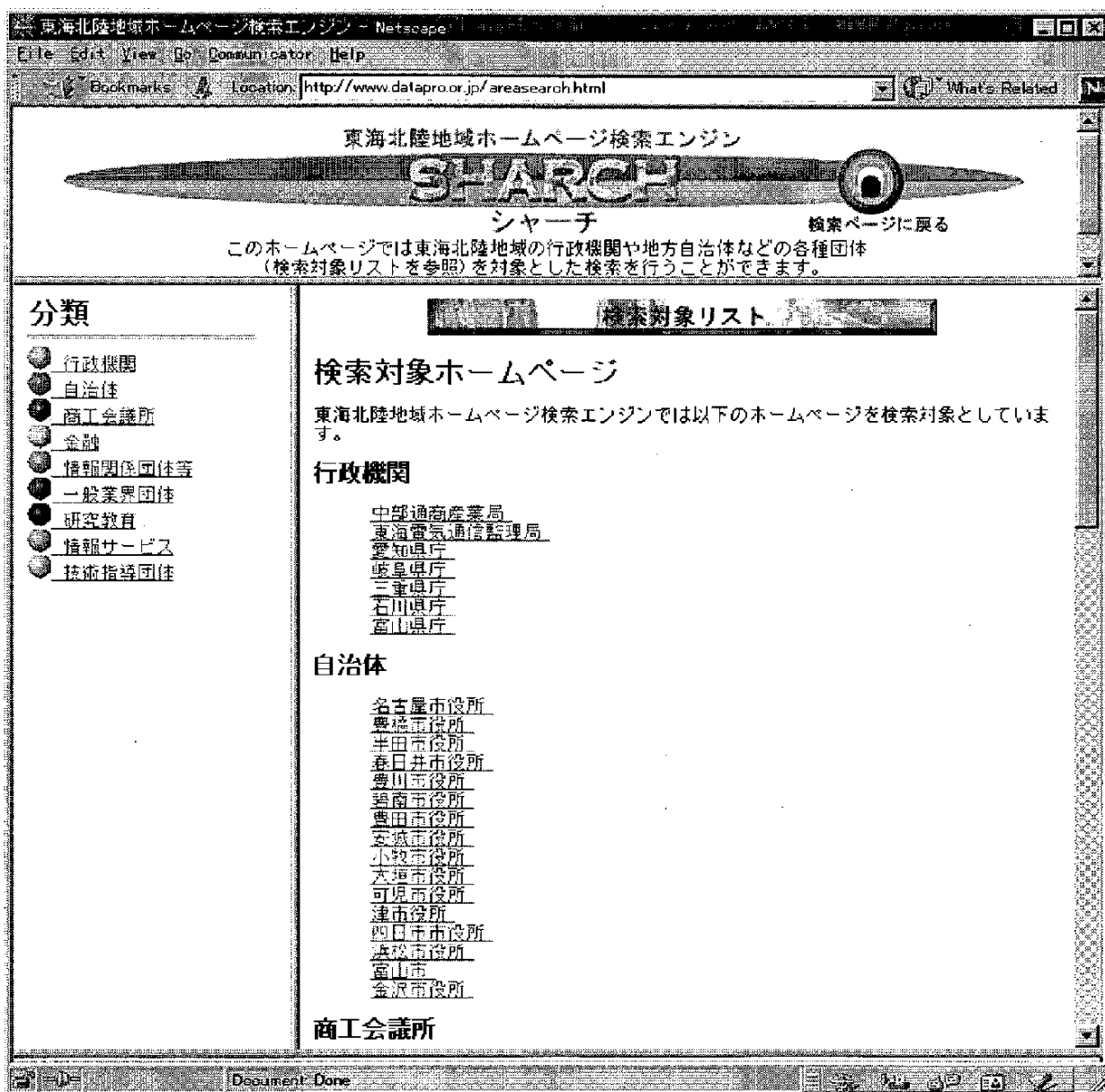


図3-3 リンク集ページ

#### 4. 検索システムの開発

##### 4. 1 システム構成

検索システムの構成は図 4-1 の通りになる。大きく 3 つの機能に分けられ以下の  
ことを行う。

(1) 検索キーワード入力ページ

- ・ユーザが入力した検索キーワードを検索エンジンに渡す。

(2) 自動巡回ロボット

- ・検索対象サイトのホームページから情報を収集する。

(3) 検索エンジン

- ・自動巡回ロボットが収集した情報から検索を効率良く行うための  
インデックスを生成する。
- ・検索キーワード入力ページを通して渡されたキーワードを元にインデックス  
から結果を求め、WWW サーバを通じて表示を行う。

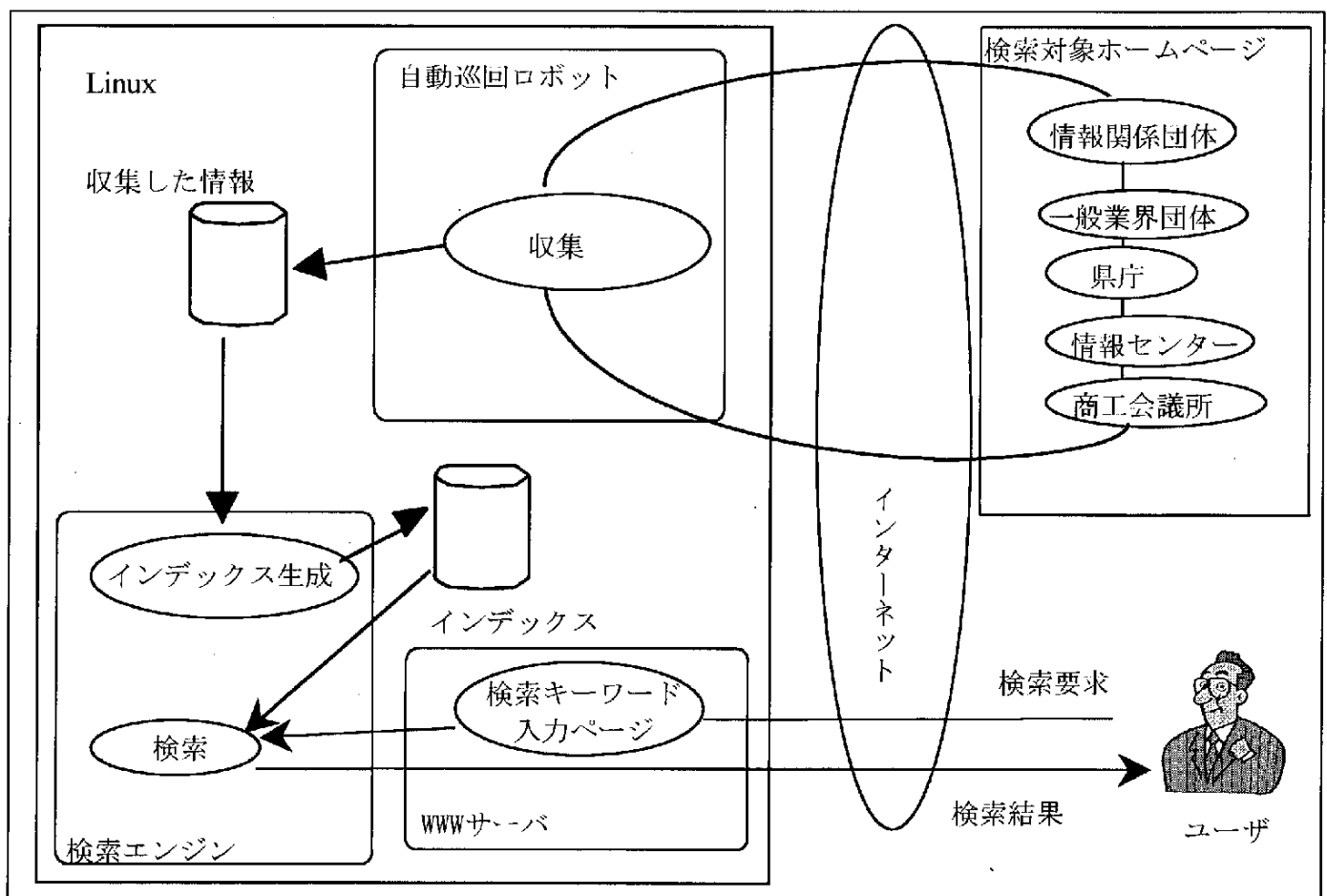


図4-1 システム構成

#### 4. 2 動作環境

検索システムが動作するマシンと必要なソフトウェアは表4-1の通りである。マシンの性能はごく一般的な IBMPCAT 互換機であり、必要なソフトは OS を含めてフリーソフトを使用した。このように検索システムの動作環境は入手が容易で汎用的な構成になっている。

表 4-1 運用に使用したマシンの性能と必要なソフト

|           |                                                             |
|-----------|-------------------------------------------------------------|
| CPU       | Pentium 130MHz                                              |
| Memory    | 64Mbyte                                                     |
| HDD       | 2.0 Gbyte                                                   |
| OS        | Linux2.0.33                                                 |
| 運用に必要なソフト | WWW サーバ (apache1.3.3)<br>perl 5.004<br>namazu1.3.0.0 (改造有り) |

#### 4. 3 検索キーワード入力ページ

- (1) 検索キーワード入力ページは初心者にもわかりやすく使ってもらうことを考えて、単純なインターフェースを採用した。  
(デザイナーの協力を得て、わかりやすいページを作成した)。
- (2) 利用者が検索システムを呼ぶ時にわかりやすくするため、SHARCH という愛称を設けた。
- (3) 他に以下のページを設けた (ページの内容は添付資料の図 7-3 ~ 図 7-5 参照)。
  - ・ 検索エンジンの概要説明
  - ・ 御意見・ご感想
  - ・ 登録

図4-2 検索キーワード入力ページ

#### 4. 4 自動巡回ロボット

自動巡回ロボットは登録サイトの先頭ページからリンク先のページをたどりながら情報を収集するソフトウェアである。登録サイトの情報を効率良く収集するために

- ・インターネット上のトラフィック
- ・WWW サーバの負荷
- ・同一サイトの情報のみを収集
- ・登録サイトが自動巡回の対象から外すページを指定する手段の提供

について考慮したソフトウェアの開発を行った。

##### 4. 4. 1 巡回方法

- ・日々更新されるホームページの内容を迅速に反映させるために登録サイトを1週間に1回巡回し、内容の更新を行うこととした。
- ・登録サイトのサーバの負荷を軽減させるためにデータのダウンロード時に5秒間のインターバルを設け、サーバの占有を避けるようにした。
- ・効率的に情報の収集を行うためにインターネットの回線が空いている時間に情報の収集を行うようにした(午前0時～午前9時)。

#### 4. 4. 2 更新の有無

トラフィック軽減のため情報を収集した日付を管理し、以下の方法により WWW サーバの機能を利用して、更新の有無を検知するようにした。

- ・ ホームページの情報をダウンした日付を記録。
- ・ サーバへの要求時にダウンした日付を渡すことにより、その日付より新しいページのみ受け取る。

また、存在しないページの情報削除し、情報収集後のインデックス化時に対象から外れるようにすることで、存在しないページがいつまでも検索結果として表示されないようにした。

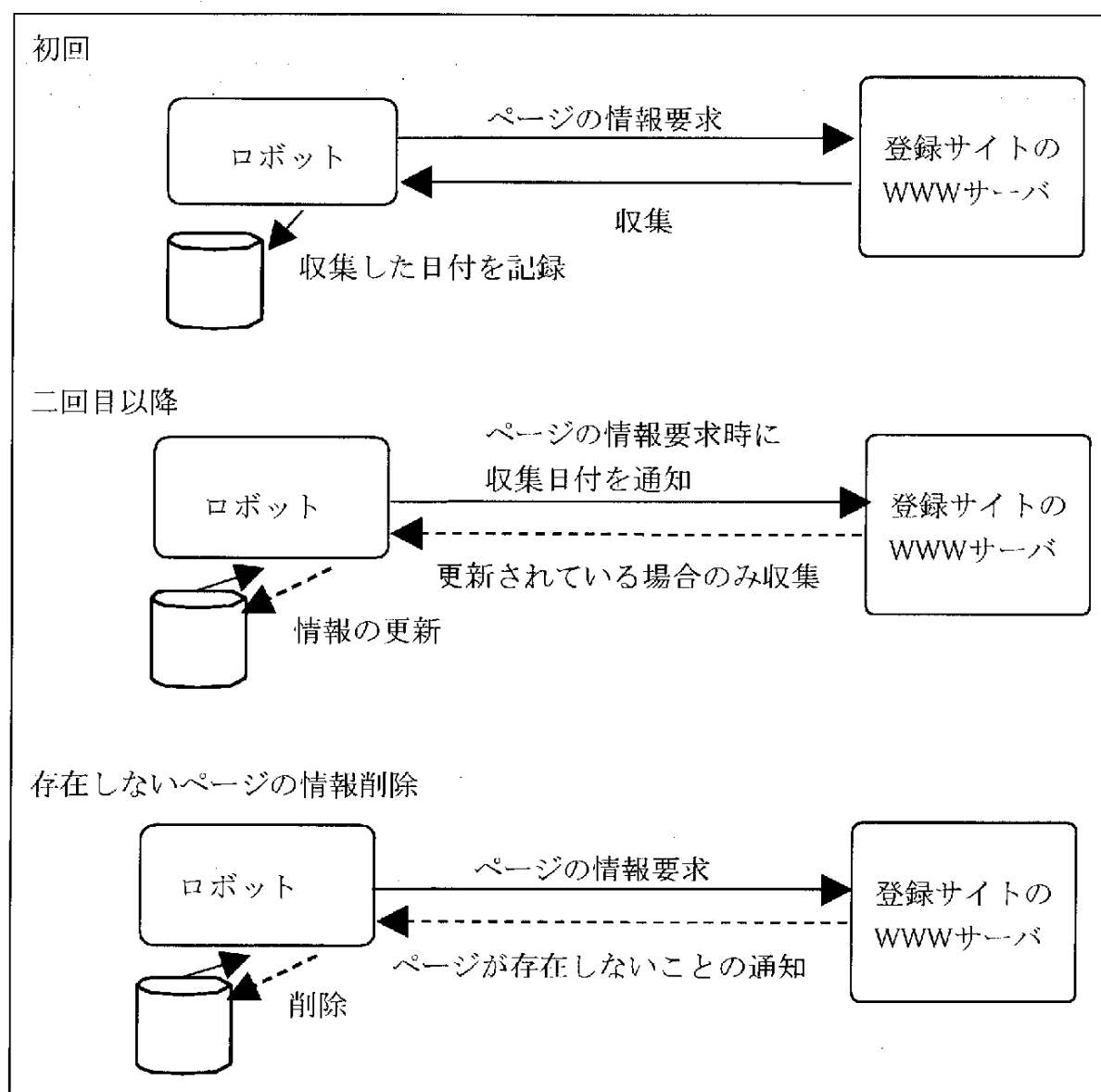


図4-3 更新の有無の検知と存在しないページの削除

#### 4. 4. 3 リンク追尾

##### (1) リンク先情報の取出し

本開発では、一般的なホームページで利用されている以下の記述からリンク先の情報を取出し、追尾を行うようにした。

- ・アンカータグ

<A HREF="リンク先">

- ・フレームタグ

<FRAME SRC="リンク先">

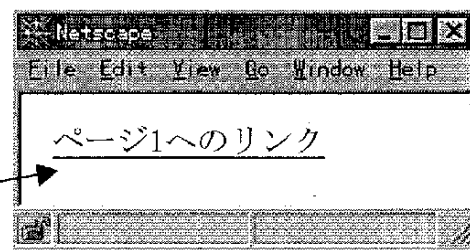
- ・AREA タグ

<AREA HREF="リンク先">

アンカータグ(通常のリンク指定)

<A HREF="page1.html"> ページ1へのリンク

page1.htmlへのリンク



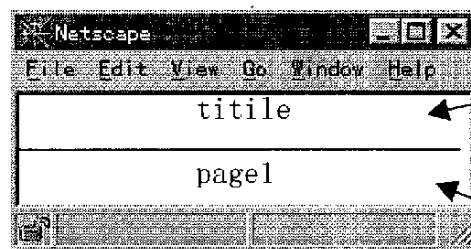
フレームタグ(ページを分轄して表示する時の指定)

<FRAMESET>

<FRAME SRC="title.html">

<FRAME SRC="page1.html">

</FRAMESET>



title.htmlを表示

page1.htmlを表示

AREA タグ(イメージ上の指定した位置をクリックした時にリンク先へ飛ぶ指定)

<IMG USEMAP="#indeximg" SRC="image.gif">

<MAP NAME="indeximg">

<AREA SHAPE="rect" COORDS="10, 10, 100, 100" HREF="page1.html">

</MAP>

イメージ中の一部がpage1.htmlへのリンク

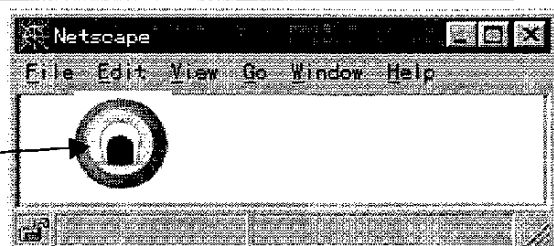


図4-4 自動追尾の対象とするタグ



## (2) 追尾対象の選択

取出したリンク先が同一サイトのページかどうかを判断し、同一サイトのページだった時のみ追尾を行うようにした。

- ・サーバから指定されている場合

“http://” から始まる指定。

対象サイトの サーバとディレクトリの比較を行い一致する時のみ同一サイトとして追尾の対象とする。

- ・トップディレクトリから指定されている場合

“/” で始まる指定。

対象サイトのトップディレクトリとの比較を行い、一致する時のみ同一サイトとして追尾の対象とする。

- ・カレントディレクトリから指定されている場合

上記以外の指定。

同一サイトとして追尾の対象とする。

自動巡回サイトの URL http://aaa.bbb.ccc/dir1/

サーバ : aaa.bbb.ccc

ディレクトリ : /dir1/

<A HREF = “http://aaa.bbb.ccc/dir1/file1.html”>

サーバとディレクトリが一致するので追尾の対象とする。

<A HREF = “http://ddd.bbb.ccc/dir1/file1.html”>

サーバが一致しないので追尾の対象とはしない。

<A HREF = “/dir1/file1.html”>

ディレクトリが一致するので追尾の対象とする。

<A HREF = “/dir2/file1.html”>

ディレクトリが一致しないので追尾の対象とはしない。

<A HREF = “file1.html”>

同一サイトとして追尾の対象とする。

図4-5 同一サイト内のページかどうかの判断例

### (3) 複数サイトを同一サーバで公開する場合への対処

同一の WWW サーバで複数サイトのページを公開していて、1つのサイトが別のサイトを含んでいる場合がある。このような時には他のサイトを含んでいるサイトにあるホームページを追尾する時にそこに含まれるサイトのホームページを対象としないように制限をかけた。

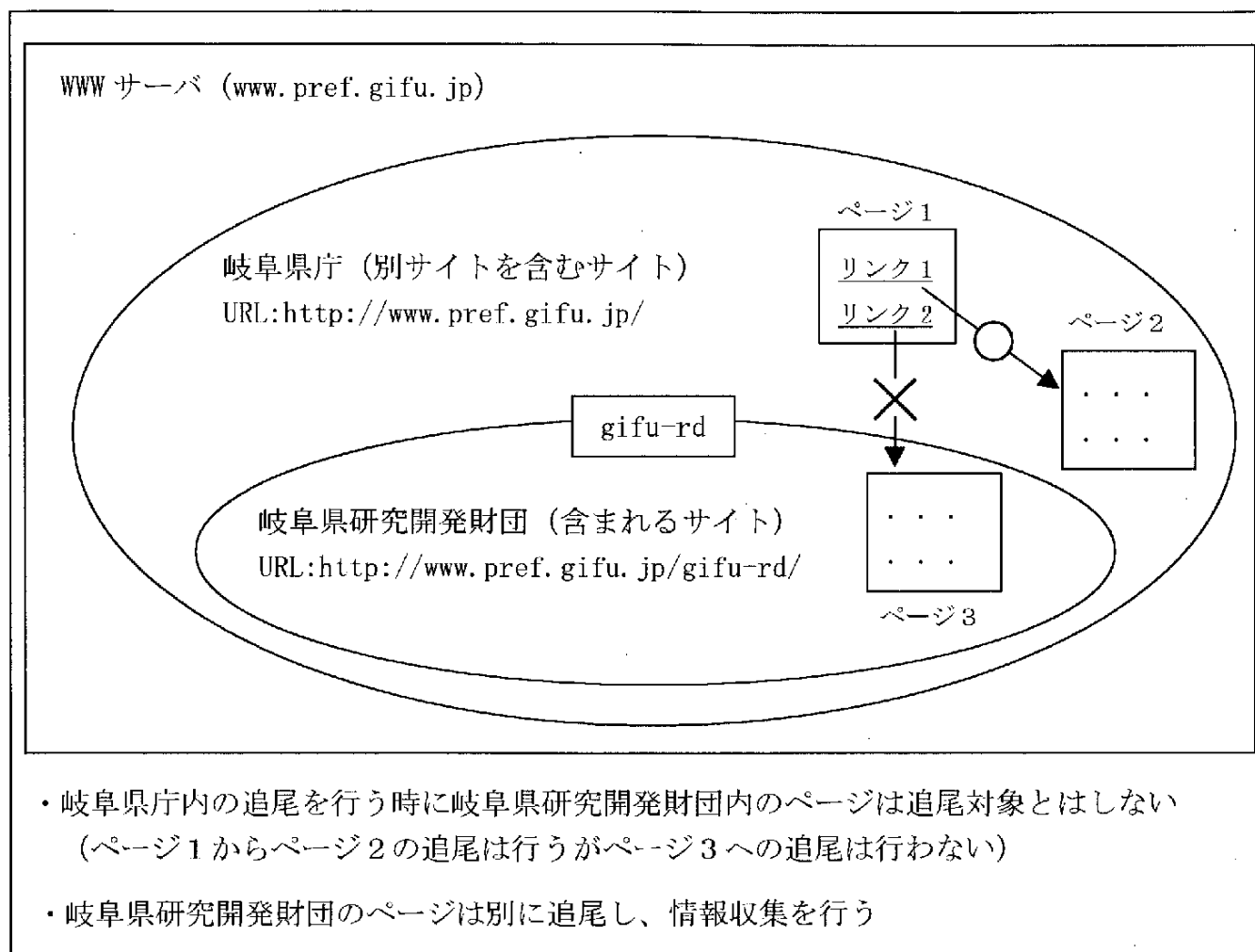


図4-6 複数サイトを同一サーバで公開する例

#### 4. 4. 4 収集ページの制限

ホームページを公開するサイトがロボットによる自動巡回の対象としたりたくないページを指定する方法として、W3C(World Wide Web Consortium) が推奨している robots.txt による制限と META タグによる制限を実装した。

##### (1) robots.txt による制限

WWW サーバのルートディレクトリに robots.txt というファイルを用意し、巡回対象から外すページ、ディレクトリの指定を行う。

”Disallow:” に続く記述とドキュメントルートからのフルパスが前方一致するページとディレクトリが巡回対象から外れる。。

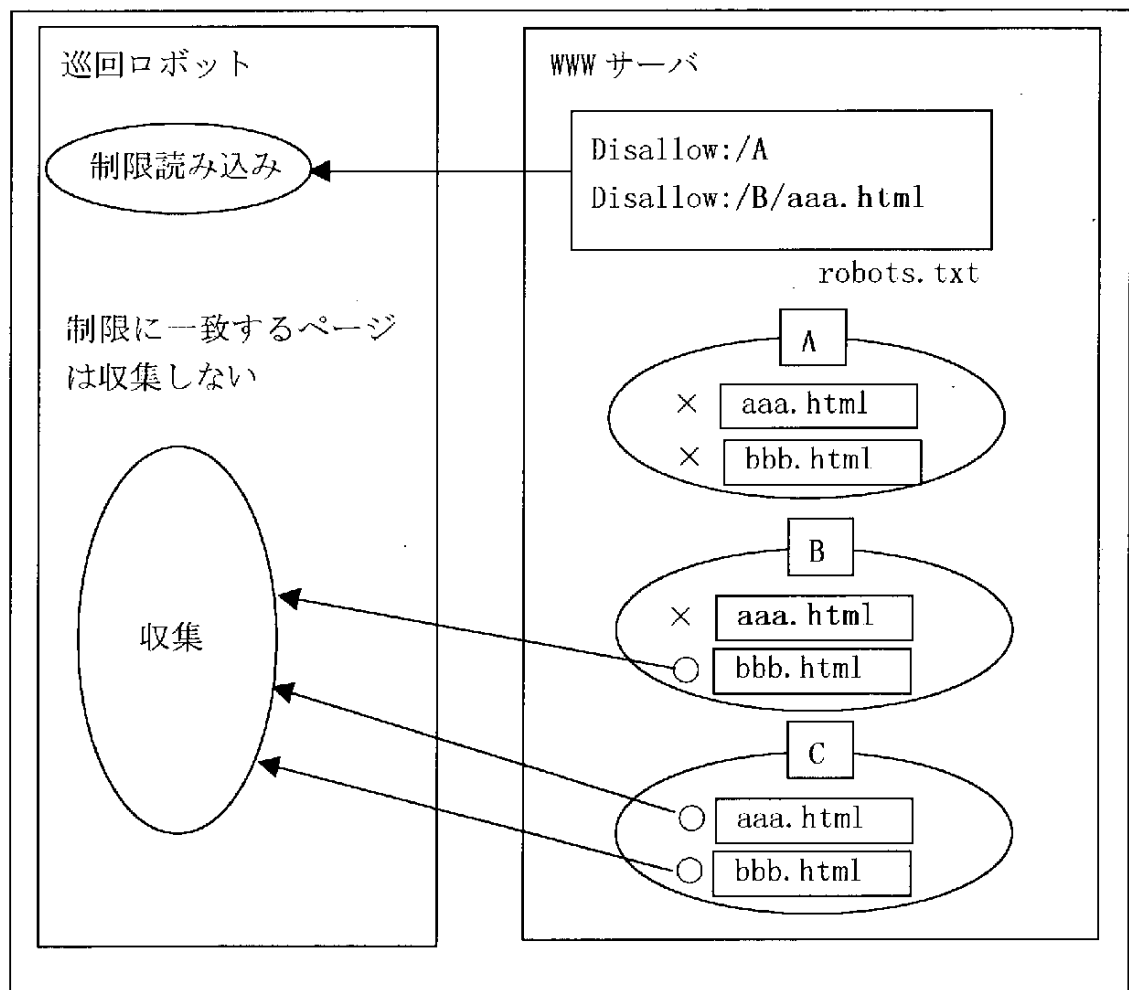


図4-7 robots.txt による制限

## (2) META タグによる制限

ページのMETA タグ内で NAME=robots の記述と” CONTENT=” の後の記述を用いてページ自体とページからのリンク先をロボットの巡回対象からはずすことができる。

表 4 - 2 CONTENTE=に記述する除外対象の指定方法

|          |                            |
|----------|----------------------------|
| NOINDEX  | ページを巡回対象からはずす。             |
| NOFOLLOW | ページからのリンクを巡回対象から外す。        |
| NONE     | NOINDEX と NOFOLLOW の両方を示す。 |
| INDEX    | ページを巡回対象として良い。             |
| ALL      | ページとリンク先の巡回自由。             |

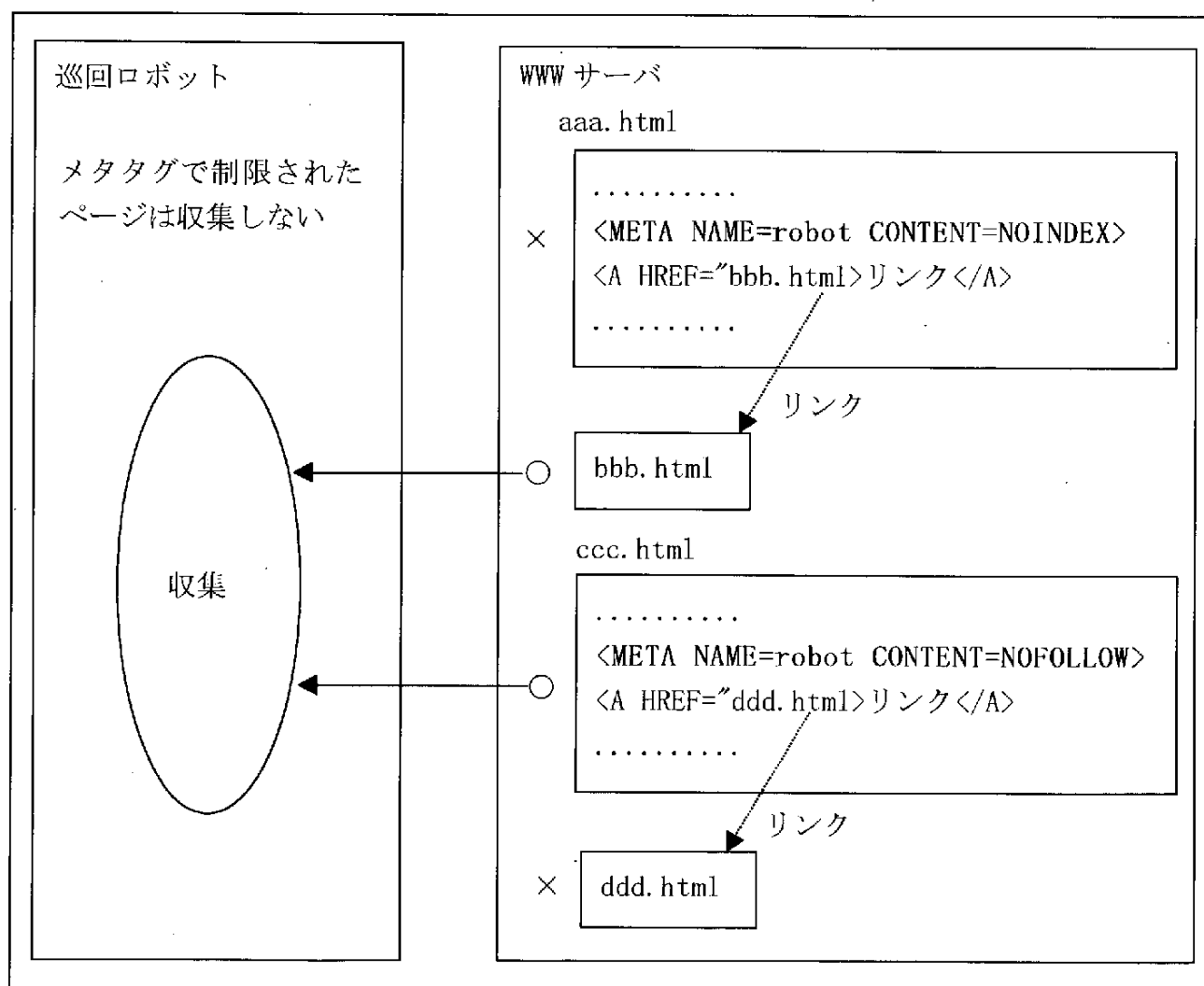


図4-8 META タグによる制限

## 4. 5 検索エンジン

自動巡回ロボットにより収集した情報を用いて検索を行う時に、そのまま文字の検索を行うと非常に時間がかかってしまう。そこで、文書中に存在する単語をインデックス化し、検索効率の向上を図る必要がある。本開発ではフリーで高機能なソフト `namazu`（”参考文献2）”参照）を利用して検索エンジンの構築を行った。

### 4. 5. 1 既存の検索エンジン `namazu`

`namazu` は手軽に使えることを第一に目指した小中規模の WWW 全文検索システムを構築することができるフリーのソフトウェアである。CGI として動作させることが可能であり、インデックスを分轄して作成し、まとめて検索する機能を持っている。これらの機能は地域検索で複数サイトを対象とする目的と合致していたため、検索エンジンの機能は `namazu` を利用して実装を行った。

### 4. 5. 2 `namazu` の改造

`namazu` はインデックス毎に含まれるキーワードの数をカウントして表示する機能を持っているが、その表示は内部的に利用しているインデックス名でインデックス全体の検索を行った時に検索結果の上部にキーワードの数には関係なく表示されるのみであった。

本開発では以下の機能を実装し、使いやすい検索インターフェースを提供するために `namazu` の改造を行った。

- ・ キーワードを多く含むサイトから順番に表示を行う。
- ・ 表示された情報からのリンクによりサイト毎の検索結果を表示する。
- ・ 対象サイトのトップページへのリンクを示す。

## 5. システムの運用

登録サイトからの情報収集を行いインターネット上に検索エンジンを公開した。キーワード入力ページを中日新聞の WWW サーバ上にある東海北陸データベース懇話会のホームページに設け、実際の情報収集と検索機能はミネルバのサーバで行う構成とした。

公開 URL は <http://www.datapro.or.jp/areasearch.html> である。

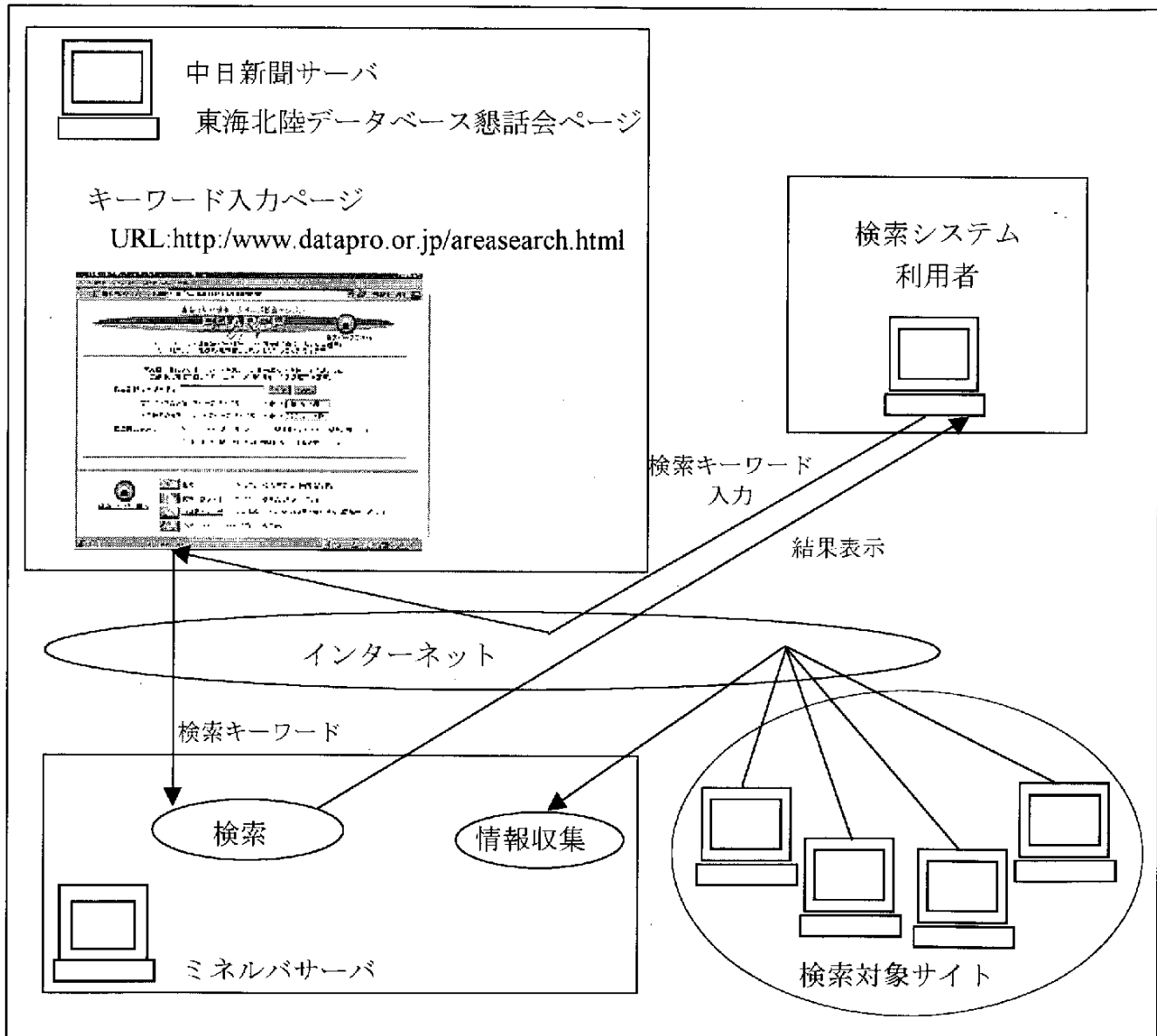


図 5 - 1 東海北陸地域ホームページ検索の運用形態

## 5. 1 他の検索エンジンとの比較

運用を開始した SHARCH と他の検索エンジンとの比較を行った。検索キーワードは”補助金”と”募集”のAND検索を行ったが、Yahoo Japan では登録ページに該当するページが存在しなかったため”補助金”のみによる検索を行った。検索を行うためのクライアントはミネルバのネットワークに繋がる PC を利用して行った。結果は表の通りである。

表 5-1 検索結果

| 検索エンジン       | 所要時間 | 該当ページ数 | 東海北陸地域のページ数 | 存在しないページ数 |
|--------------|------|--------|-------------|-----------|
| SHARCH       | 3 秒  | 127    | 127         | 0         |
| Infoseek     | 3 秒  | 246    | 28          | 10        |
| Yahoo! Japan | 3 秒  | 16     | 3           | 0         |

所用時間に余り差は見られなかった。

該当ページ数は SHARCH と Infoseek の差は少なかった。Infoseek の検索結果は全国のページを示しているため、東海北陸地域の情報に絞り込むと該当するページ数は SHARCH が示す該当ページ数よりかなり少なくなる。これは意外な結果であったが、SHARCH の情報収集が詳細に渡り、行えていることを示している。しかし、この結果はユーザにさらなる情報の絞り込みを要求することになる。これに対しては登録サイトの協力を頂き、検索に有効なページを収集対象としてもらうことが、より使いやすい検索手法の提供に繋がる。巡回ロボットには登録サイトとの連携に必要な収集ページの制限機能が実装されている（4. 4. 4 参照）。

一方、Yahoo については予想以上に該当ページ数が少なかった。しかし、Yahoo の検索結果として表示されるページはキーワード、カテゴリ共に厳選されたものなので、信頼度は高いと言える。

存在しないページ数は表示された結果をたどった時にサーバから削除されたことにより、WWW ブラウザにエラーが表示されるページの数である。SHARCH は 1 週間に 1 回の頻度で更新を行っているため、この数は検索結果の数に対してかなり少なくなる。これは対象サイトを限定することで、巡回時間を短縮することによって実現できた効果である。

## 5. 2 利用状況

SHARCH を公開してからの利用状況は以下の通りである。

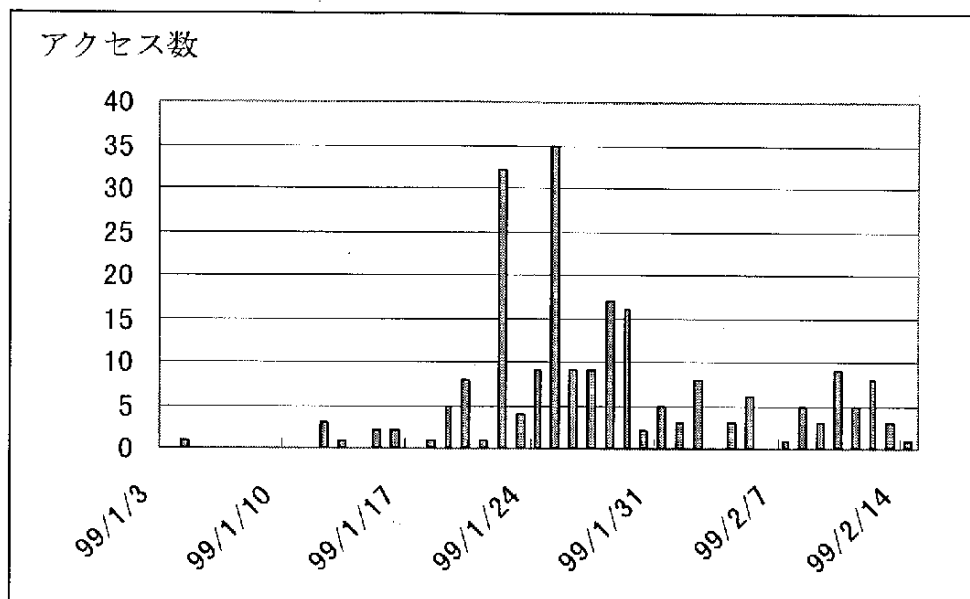


図 5 - 2 アクセス数

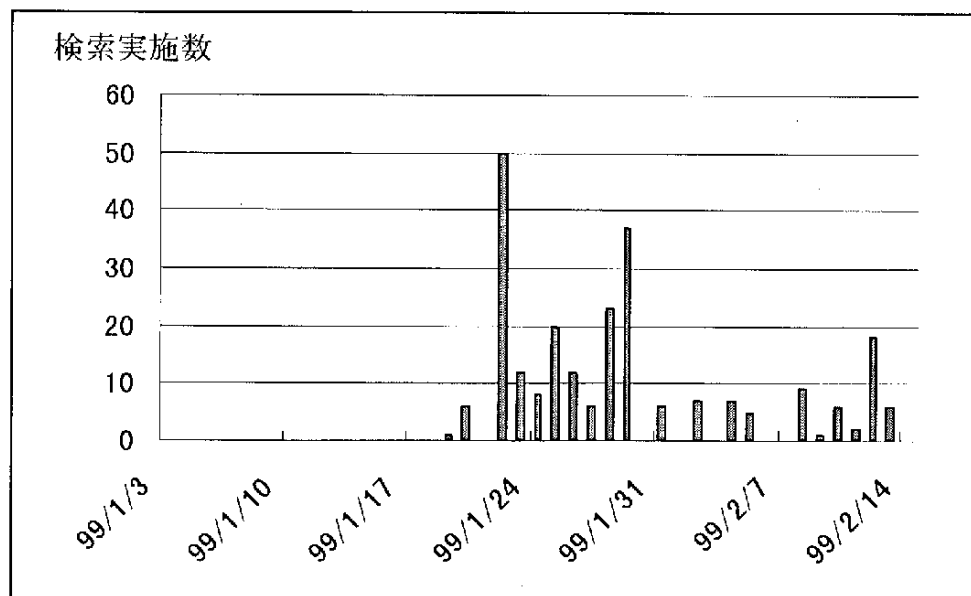


図 5 - 3 検索実施数



### 5. 3 巡回結果

自動巡回ロボットの情報収集状況は以下の通りである。

表 5－2 自動巡回ロボットの情報収集状況

|                   |                 |
|-------------------|-----------------|
| 情報収集に要した時間        | 20 時間           |
| 収集ページ数            | 59,179 ページ      |
| インデックス化に要した時間（初回） | 75 時間           |
| インデックスの合計サイズ      | 163 M バイト       |
| インデックスの更新に要した時間   | 17 時間/週（平均）     |
| ホームページの更新状況       | 5,421 ページ/週（平均） |

表 5－3 各キーワードによる検索該当ページ数

| キーワード | 該当ページ数 |
|-------|--------|
| 補助金   | 766    |
| 政策    | 1,582  |
| 融資    | 1,071  |
| 経営指導  | 102    |
| 資格    | 1,419  |
| 採用    | 1,516  |
| 講習会   | 888    |

## 6. まとめと今後の課題

今回、登録サイトとして調査、選択を行った東海北陸地域の公的機関のホームページリンク集はクリアリングハウスとしての機能を十分果たすものと期待される。また、利用者が、簡単な操作で検索を行える機能を整備することができた。

本検索エンジンを使用して頂いた方からは、” 関連機関だけを対象としているということで、一般のエンジンより使い勝手は良い” や” 登録サイトが公開しているページの状況がわかりやすい” 等の意見をもらうことができた。

しかし、今回の試験運用では公開からの期間が短く、一般ユーザへの通知が十分では無いことから、利用者からの意見や検索実施数が少なかった。このため、正確な利用情報を得ることができていない。利用情報の分析結果や利用者からの意見を反映するには、少なくとも1年程度の運用継続が必要と考えられ、それにより、更なるシステムの充実を図ることが可能となる。

今回の開発により、検索対象の絞り込みを行い、必要な資源の削減と情報収集に必要な時間の減少を図ることが、有効であるという例を示すことができた。今後はこのように対象範囲を限定し、利用者にとって本当に必要な情報を提供することができるよう仕組みがますます重要になってくると思われる。結果として、検索対象サイトが、本来情報を必要とする人に、公開する内容を提供する機会を増やすことにつながる。

今回は、東海北陸地域という地域に限定した情報を対象とした。しかし、本システムは以下のように特定の分野やネットワークを対象とするような検索にも応用が可能である。

- ・各部署や関連会社に散在する文書の統合検索システム
- ・自社に関連する製品情報検索システム
- ・専門分野に特化したサイトだけを検索するシステム（例：風力発電）
- ・同じ事に興味を持つコミュニティー（例：リサイクル）

今後は、このようなイントラネットや他地域への展開を目指していく予定である。

## 7. 添付資料

### 7. 1 検索対象サイトと URL

表 7-1 検索対象一覧

| 分野    | 地域  | 団体名       | URL                                                                                           |
|-------|-----|-----------|-----------------------------------------------------------------------------------------------|
| 行政機関  |     |           |                                                                                               |
|       |     | 中部通商産業局   | <a href="http://www.chubu.miti.go.jp/">http://www.chubu.miti.go.jp/</a>                       |
|       |     | 東海電気通信監理局 | <a href="http://www.tokai-ht.mpt.go.jp/">http://www.tokai-ht.mpt.go.jp/</a>                   |
|       |     | 愛知県庁      | <a href="http://www.pref.aichi.jp/">http://www.pref.aichi.jp/</a>                             |
|       |     | 岐阜県庁      | <a href="http://www.pref.gifu.jp/">http://www.pref.gifu.jp/</a>                               |
|       |     | 三重県庁      | <a href="http://www.pref.mie.jp/">http://www.pref.mie.jp/</a>                                 |
|       |     | 石川県庁      | <a href="http://www.pref.ishikawa.jp/">http://www.pref.ishikawa.jp/</a>                       |
|       |     | 富山県庁      | <a href="http://www.pref.toyama.jp/">http://www.pref.toyama.jp/</a>                           |
| 自治体   |     |           |                                                                                               |
|       | 愛知県 |           |                                                                                               |
|       |     | 名古屋市役所    | <a href="http://www.city.nagoya.jp/">http://www.city.nagoya.jp/</a>                           |
|       |     | 豊橋市役所     | <a href="http://www.city.toyohashi.aichi.jp/">http://www.city.toyohashi.aichi.jp/</a>         |
|       |     | 半田市役所     | <a href="http://www.city.handa.aichi.jp/">http://www.city.handa.aichi.jp/</a>                 |
|       |     | 春日井市役所    | <a href="http://www.city.kasugai.aichi.jp/">http://www.city.kasugai.aichi.jp/</a>             |
|       |     | 豊川市役所     | <a href="http://www.city.toyokawa.aichi.jp/">http://www.city.toyokawa.aichi.jp/</a>           |
|       |     | 碧南市役所     | <a href="http://www.city.hekinan.aichi.jp/">http://www.city.hekinan.aichi.jp/</a>             |
|       |     | 豊田市役所     | <a href="http://www.city.toyota.aichi.jp/">http://www.city.toyota.aichi.jp/</a>               |
|       |     | 安城市役所     | <a href="http://www.city.anjo.aichi.jp/">http://www.city.anjo.aichi.jp/</a>                   |
|       |     | 小牧市役所     | <a href="http://www.city.komaki.aichi.jp/">http://www.city.komaki.aichi.jp/</a>               |
|       | 岐阜県 |           |                                                                                               |
|       |     | 大垣市役所     | <a href="http://www.ginet.or.jp/ogaki/">http://www.ginet.or.jp/ogaki/</a>                     |
|       |     | 可児市役所     | <a href="http://www.city.kani.gifu.jp/index.html">http://www.city.kani.gifu.jp/index.html</a> |
|       | 三重県 |           |                                                                                               |
|       |     | 津市役所      | <a href="http://www.info.city.tsu.mie.jp/">http://www.info.city.tsu.mie.jp/</a>               |
|       |     | 四日市市役所    | <a href="http://www.city.yokkaichi.mie.jp/">http://www.city.yokkaichi.mie.jp/</a>             |
|       | 静岡県 |           |                                                                                               |
|       |     | 浜松市役所     | <a href="http://www.city.hamamatsu.shizuoka.jp/">http://www.city.hamamatsu.shizuoka.jp/</a>   |
|       | 富山県 |           |                                                                                               |
|       |     | 富山市       | <a href="http://www.cap.or.jp/toyama/">http://www.cap.or.jp/toyama/</a>                       |
|       | 石川県 |           |                                                                                               |
|       |     | 金沢市役所     | <a href="http://www.city.kanazawa.ishikawa.jp/">http://www.city.kanazawa.ishikawa.jp/</a>     |
| 商工会議所 |     |           |                                                                                               |
|       | 愛知県 |           |                                                                                               |

|  |          |                                                                                                               |
|--|----------|---------------------------------------------------------------------------------------------------------------|
|  | 名古屋商工会議所 | <a href="http://www.nagoya-cci.or.jp/">http://www.nagoya-cci.or.jp/</a>                                       |
|  | 岡崎商工会議所  | <a href="http://www.okazakicci.or.jp/">http://www.okazakicci.or.jp/</a>                                       |
|  | 豊橋商工会議所  | <a href="http://www.tsc.co.jp/~shoko/">http://www.tsc.co.jp/~shoko/</a>                                       |
|  | 瀬戸商工会議所  | <a href="http://www.taikokk.com/setocci/">http://www.taikokk.com/setocci/</a>                                 |
|  | 蒲郡商工会議所  | <a href="http://www.gfnet.co.jp/~gccci/">http://www.gfnet.co.jp/~gccci/</a>                                   |
|  | 豊川商工会議所  | <a href="http://www.tsc.co.jp/~tte/">http://www.tsc.co.jp/~tte/</a>                                           |
|  | 安城商工会議所  | <a href="http://www.anjo-cci.or.jp/">http://www.anjo-cci.or.jp/</a>                                           |
|  | 西尾商工会議所  | <a href="http://www.japan-net.or.jp/~cci2213/">http://www.japan-net.or.jp/~cci2213/</a>                       |
|  | 津島商工会議所  | <a href="http://www.synnet.or.jp/t-cci/">http://www.synnet.or.jp/t-cci/</a>                                   |
|  | 春日井商工会議所 | <a href="http://www.aichinet.ne.jp/~kcci/">http://www.aichinet.ne.jp/~kcci/</a>                               |
|  | 常滑商工会議所  | <a href="http://www.gld.mmtr.or.jp/~cci4422d/">http://www.gld.mmtr.or.jp/~cci4422d/</a>                       |
|  | 小牧商工会議所  | <a href="http://www.k-net.or.jp/~cci/">http://www.k-net.or.jp/~cci/</a>                                       |
|  | 犬山商工会議所  | <a href="http://www.gld.mmtr.or.jp/~icci/">http://www.gld.mmtr.or.jp/~icci/</a>                               |
|  | 岐阜県      |                                                                                                               |
|  | 岐阜商工会議所  | <a href="http://www.gcci.or.jp/index.htm">http://www.gcci.or.jp/index.htm</a>                                 |
|  | 大垣商工会議所  | <a href="http://www.ginet.or.jp/ogakicci/">http://www.ginet.or.jp/ogakicci/</a>                               |
|  | 中津川商工会議所 | <a href="http://www.takenet.or.jp/~ncci/">http://www.takenet.or.jp/~ncci/</a>                                 |
|  | 土岐商工会議所  | <a href="http://www.blk.mmtr.or.jp/~tokicci/">http://www.blk.mmtr.or.jp/~tokicci/</a>                         |
|  | 瑞浪商工会議所  | <a href="http://www.cci.mizunami.gifu.jp/">http://www.cci.mizunami.gifu.jp/</a>                               |
|  | 各務原商工会議所 | <a href="http://www.gcs.ne.jp/ccik/">http://www.gcs.ne.jp/ccik/</a>                                           |
|  | 可児商工会議所  | <a href="http://www.cci.kani.gifu.jp/">http://www.cci.kani.gifu.jp/</a>                                       |
|  | 三重県      |                                                                                                               |
|  | 四日市商工会議所 | <a href="http://www.miehis.co.jp/yokkaichi-cci/">http://www.miehis.co.jp/yokkaichi-cci/</a>                   |
|  | 津商工会議所   | <a href="http://www2.mint.or.jp/~tsucci/">http://www2.mint.or.jp/~tsucci/</a>                                 |
|  | 伊勢商工会議所  | <a href="http://www.ise-cci.or.jp/">http://www.ise-cci.or.jp/</a>                                             |
|  | 松坂商工会議所  | <a href="http://www.mctv.ne.jp/~kaigisho/">http://www.mctv.ne.jp/~kaigisho/</a>                               |
|  | 鈴鹿商工会議所  | <a href="http://scci.or.jp/index.html">http://scci.or.jp/index.html</a>                                       |
|  | 桑名商工会議所  | <a href="http://www.kuwana.ne.jp/chamber/">http://www.kuwana.ne.jp/chamber/</a>                               |
|  | 鳥羽商工会議所  | <a href="http://www. adesso.co.jp/toba/syokou/index.html">http://www. adesso.co.jp/toba/syokou/index.html</a> |
|  | 石川県      |                                                                                                               |
|  | 金沢商工会議所  | <a href="http://www.nsknet.or.jp/kaigisho/">http://www.nsknet.or.jp/kaigisho/</a>                             |
|  | 小松商工会議所  | <a href="http://www.kissnet.or.jp/komatcci/">http://www.kissnet.or.jp/komatcci/</a>                           |
|  | 七尾商工会議所  | <a href="http://www.noto.or.jp/nanao/cci/">http://www.noto.or.jp/nanao/cci/</a>                               |
|  | 加賀商工会議所  | <a href="http://www.kagaworld.or.jp/">http://www.kagaworld.or.jp/</a>                                         |
|  | 富山県      |                                                                                                               |
|  | 富山商工会議所  | <a href="http://www.ccis-toyama.or.jp/toyama/">http://www.ccis-toyama.or.jp/toyama/</a>                       |
|  | 高岡商工会議所  | <a href="http://www.ccis-toyama.or.jp/takaoka/">http://www.ccis-toyama.or.jp/takaoka/</a>                     |
|  | 氷見商工会議所  | <a href="http://www.ccis-toyama.or.jp/himi/">http://www.ccis-toyama.or.jp/himi/</a>                           |
|  | 新湊商工会議所  | <a href="http://www.ccis-toyama.or.jp/scci/">http://www.ccis-toyama.or.jp/scci/</a>                           |

|         |                        |                                                                                                                                   |
|---------|------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| 金融      |                        |                                                                                                                                   |
|         | 中小企業金融公庫               | <a href="http://www.jfs.go.jp/">http://www.jfs.go.jp/</a>                                                                         |
|         | 十六銀行                   | <a href="http://www.jic-gifu.or.jp/juroku/data/data.htm">http://www.jic-gifu.or.jp/juroku/data/data.htm</a>                       |
|         | 愛知県                    |                                                                                                                                   |
|         | 愛知県信用保証協会              | <a href="http://www.tokaido.co.jp/aishinpo/">http://www.tokaido.co.jp/aishinpo/</a>                                               |
|         | 名古屋市                   |                                                                                                                                   |
|         | 名古屋市信用保証協会             | <a href="http://www.cgc-nagoya.or.jp/">http://www.cgc-nagoya.or.jp/</a>                                                           |
| 情報関係団体等 |                        |                                                                                                                                   |
|         | 東海北陸データベース懇話会          | <a href="http://www.datapro.or.jp/">http://www.datapro.or.jp/</a>                                                                 |
|         | (社) 中部経営情報化協会          | <a href="http://www.aichi-iic.or.jp/co/coma/index.html">http://www.aichi-iic.or.jp/co/coma/index.html</a>                         |
|         | 中部エレクトロニクス振興会          | <a href="http://www.nmiri.city.nagoya.jp/ELESHOW/eresin.htm">http://www.nmiri.city.nagoya.jp/ELESHOW/eresin.htm</a>               |
|         | 愛知県                    |                                                                                                                                   |
|         | 愛知県産業情報センター            | <a href="http://www.aichi-iic.or.jp/">http://www.aichi-iic.or.jp/</a>                                                             |
|         | (社) 愛知県情報サービス産業協会      | <a href="http://www.aichi-iic.or.jp/enterprise/complist/AIA/">http://www.aichi-iic.or.jp/enterprise/complist/AIA/</a>             |
|         | 岐阜県                    |                                                                                                                                   |
|         | (財) ソフトピアジャパン          | <a href="http://www.softopia.pref.gifu.jp/index_J.html">http://www.softopia.pref.gifu.jp/index_J.html</a>                         |
|         | 三重県                    |                                                                                                                                   |
|         | 三重県中小企業情報センター          | <a href="http://web.iap.co.jp/MIC/">http://web.iap.co.jp/MIC/</a>                                                                 |
|         | 石川県                    |                                                                                                                                   |
|         | (財) 石川県中小企業情報センター      | <a href="http://www.ibis.or.jp/">http://www.ibis.or.jp/</a>                                                                       |
|         | 富山県                    |                                                                                                                                   |
|         | 情報ビル ((株) 富山県総合情報センター) | <a href="http://www.toyama-tic.co.jp/">http://www.toyama-tic.co.jp/</a>                                                           |
|         | 富山県マルチメディア情報センター       | <a href="http://www.mmc.pref.toyama.jp/">http://www.mmc.pref.toyama.jp/</a>                                                       |
|         | 名古屋市                   |                                                                                                                                   |
|         | 名古屋市中企業情報センター          | <a href="http://www.info-c.city.nagoya.jp/">http://www.info-c.city.nagoya.jp/</a>                                                 |
| 一般業界団体  |                        |                                                                                                                                   |
|         | (財) 中部科学技術センター         | <a href="http://www.shatchy.ne.jp/cstc/">http://www.shatchy.ne.jp/cstc/</a>                                                       |
|         | 中部映像関連事業者協会            | <a href="http://www.cri.co.jp/tyueikyo/">http://www.cri.co.jp/tyueikyo/</a>                                                       |
|         | 科学技術交流財団               | <a href="http://www.astf.or.jp/">http://www.astf.or.jp/</a>                                                                       |
|         | 愛知県                    |                                                                                                                                   |
|         | 愛知県中小企業団体中央会           | <a href="http://www.aiweb.or.jp/">http://www.aiweb.or.jp/</a>                                                                     |
| 研究・教育   |                        |                                                                                                                                   |
|         | 中部EC研究会                | <a href="http://www.infosite.ne.jp/cec/index.html">http://www.infosite.ne.jp/cec/index.html</a>                                   |
|         | (財) 中部産業・労働政策研究会       | <a href="http://www.tns.or.jp/sanseiken/">http://www.tns.or.jp/sanseiken/</a>                                                     |
|         | 岐阜県                    |                                                                                                                                   |
|         | 岐阜県研究開発財団              | <a href="http://www.pref.gifu.jp/gifu-rd/">http://www.pref.gifu.jp/gifu-rd/</a>                                                   |
|         | 名古屋市                   |                                                                                                                                   |
|         | 愛知県立名古屋高等技術専門校         | <a href="http://www.pref.aichi.jp/rodo/noukai/nagoya/syoutkai.html">http://www.pref.aichi.jp/rodo/noukai/nagoya/syoutkai.html</a> |
|         | 名古屋産業科学研究所             | <a href="http://www.nisri.moriyama.nagoya.jp/">http://www.nisri.moriyama.nagoya.jp/</a>                                           |

|        |                        |  |                                                                                                                       |
|--------|------------------------|--|-----------------------------------------------------------------------------------------------------------------------|
| 情報サービス |                        |  |                                                                                                                       |
|        | 東海リサーチリンクージ            |  | <a href="http://www.shatchy.ne.jp/trl/">http://www.shatchy.ne.jp/trl/</a>                                             |
|        | DON DON TOWN           |  | <a href="http://www.nhk-chubu-brains.co.jp/">http://www.nhk-chubu-brains.co.jp/</a>                                   |
|        | 三重県                    |  |                                                                                                                       |
|        | Human internet station |  | <a href="http://www.miehis.co.jp/">http://www.miehis.co.jp/</a>                                                       |
|        | 鳥羽パールページ               |  | <a href="http://www.adesso.co.jp/toba/index.html">http://www.adesso.co.jp/toba/index.html</a>                         |
|        | 名古屋市                   |  |                                                                                                                       |
|        | なごやUネット（名古屋都市産業振興公社）   |  | <a href="http://www.u-net.city.nagoya.jp/">http://www.u-net.city.nagoya.jp/</a>                                       |
| 技術指導団体 |                        |  |                                                                                                                       |
|        | 愛知県                    |  |                                                                                                                       |
|        | 愛知県工業技術センター            |  | <a href="http://www.airi.aichi-iic.or.jp/">http://www.airi.aichi-iic.or.jp/</a>                                       |
|        | 岐阜県                    |  |                                                                                                                       |
|        | 岐阜県工業技術センター            |  | <a href="http://www.com.rd.pref.gifu.jp/">http://www.com.rd.pref.gifu.jp/</a>                                         |
|        | 三重県                    |  |                                                                                                                       |
|        | 三重県工業技術総合研究所           |  | <a href="http://www.mie-iri.tsu.mie.jp/">http://www.mie-iri.tsu.mie.jp/</a>                                           |
|        | 石川県                    |  |                                                                                                                       |
|        | 石川県工業試験場               |  | <a href="http://www.irii.go.jp/">http://www.irii.go.jp/</a>                                                           |
|        | 富山県                    |  |                                                                                                                       |
|        | 富山県工業技術センター            |  | <a href="http://www.itc.pref.toyama.jp/">http://www.itc.pref.toyama.jp/</a>                                           |
|        | 名古屋市                   |  |                                                                                                                       |
|        | 名古屋市工業研究所              |  | <a href="http://www.nmiri.city.nagoya.jp/nmiri/index-jp.html">http://www.nmiri.city.nagoya.jp/nmiri/index-jp.html</a> |

## 7. 2 画面サンプル

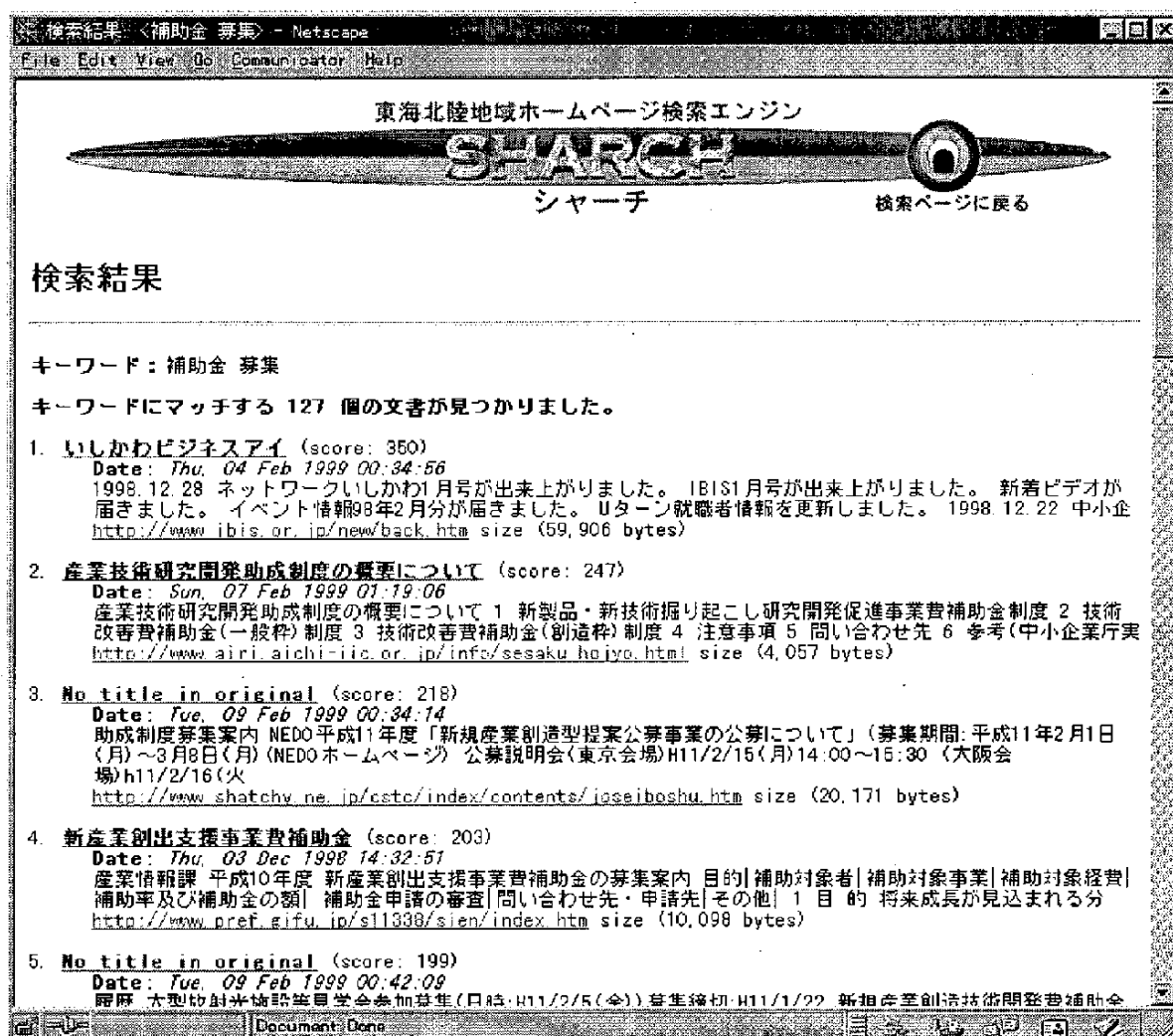


図7-1 “補助金” “募集” のAND検索結果例(ホームページ)

検索結果: <補助金 募集> - Netscape

File Edit View Go Communicator Help

東海北陸地域ホームページ検索エンジン

**SHARCH**

シャーチ

検索ページに戻る

## 検索結果

キーワード: 補助金 募集

キーワードにマッチする 127 個の文書が見つかりました。

| 各サイトの詳細                | 各サイトの先頭ページ | キーワードを含む文書数 |
|------------------------|------------|-------------|
| 岐阜県庁                   | 先頭ページ      | 41          |
| (財) 石川県中小企業情報センター      | 先頭ページ      | 39          |
| 三重県庁                   | 先頭ページ      | 9           |
| 中部通商産業局                | 先頭ページ      | 7           |
| 科学技術文庫財団               | 先頭ページ      | 5           |
| 東海電気通信監理局              | 先頭ページ      | 4           |
| 大垣市役所                  | 先頭ページ      | 4           |
| 愛知県庁                   | 先頭ページ      | 3           |
| (財) 中部科学技術センター         | 先頭ページ      | 3           |
| Human internet station | 先頭ページ      | 3           |
| 三重県工業技術総合研究所           | 先頭ページ      | 3           |
| 富山県庁                   | 先頭ページ      | 2           |
| 三重県中小企業情報センター          | 先頭ページ      | 2           |
| 愛知県工業技術センター            | 先頭ページ      | 2           |
| 名古屋市役所                 | 先頭ページ      | 1           |
| 春日井市役所                 | 先頭ページ      | 1           |
| 豊川市役所                  | 先頭ページ      | 1           |
| 常滑商工会議所                | 先頭ページ      | 1           |
| 情報ビル((株)富山県総合情報センター)   | 先頭ページ      | 1           |
| DON DON TOWN           | 先頭ページ      | 1           |

Document Done

図 7-2 ” 補助金 ” “募集” の AND 検索結果例 (サイト)



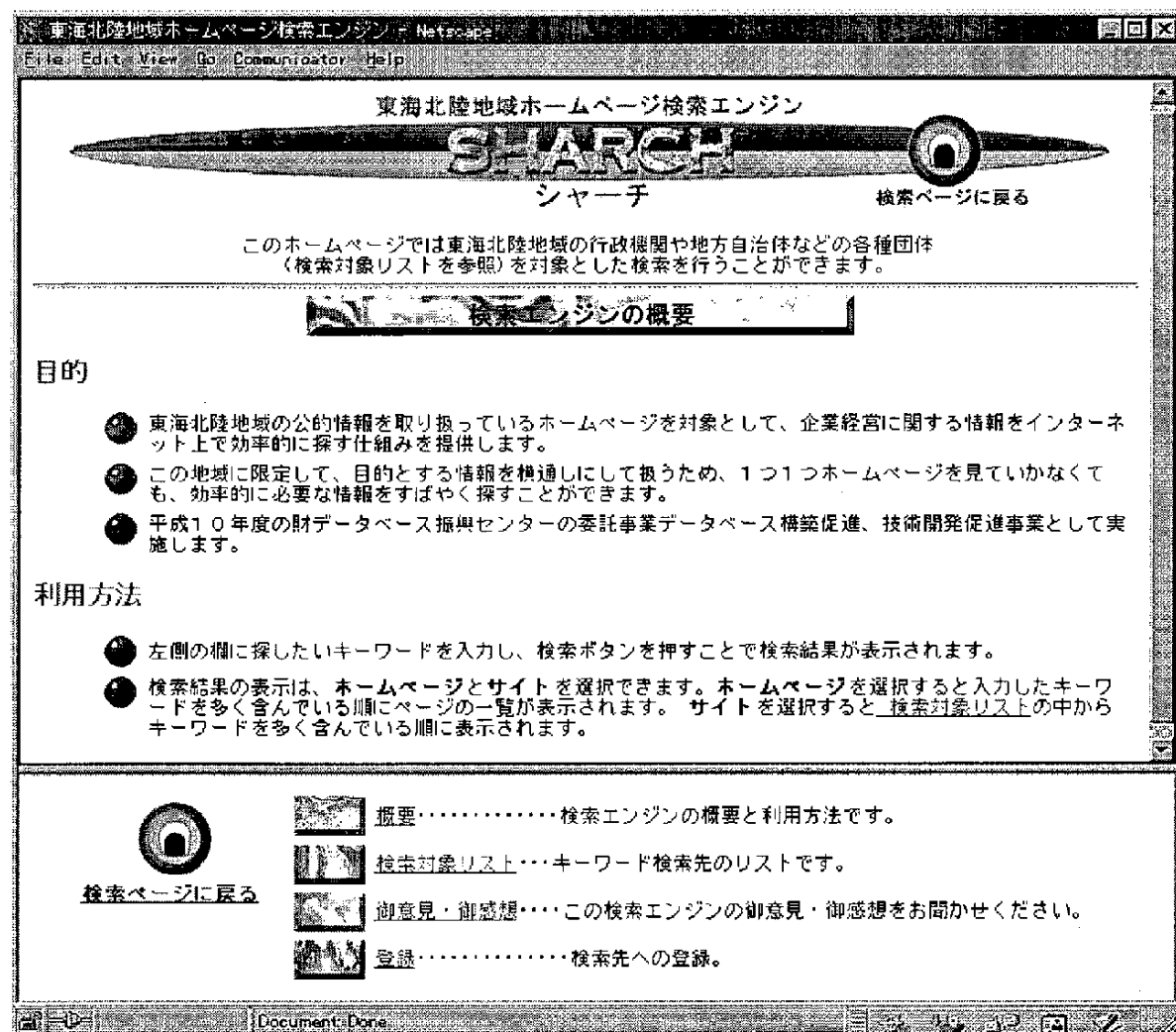


図 7-3 検索エンジンの概要ページ

東海北陸地域ホームページ検索エンジン - Netscape

File Edit View Go Communicator Help

---

東海北陸地域ホームページ検索エンジン

**SHARCH**

シャーチ 検索ページに戻る

---

**御意見・御感想**

検索エンジンをお使いになった御意見・御感想を御記入の上、“送信”ボタンを押してください。

氏名:

E-mailアドレス:

勤務先:

記入欄:

---



[検索ページに戻る](#)

 [概要](#).....検索エンジンの概要と利用方法です。

 [検索対象リスト](#)...キーワード検索先のリストです。

 [御意見・御感想](#)...この検索エンジンの御意見・御感想をお聞かせください。

 [登録](#).....検索先への登録。

Document Done

図 7-4 御意見・ご感想ページ

東海北陸地域ホームページ検索エンジン - Netscape

File Edit View Go Communicator Help

---

東海北陸地域ホームページ検索エンジン

**SHARCH**

シャーチ 検索ページに戻る

このホームページでは東海北陸地域の行政機関や地方自治体などの各種団体  
(検索対象リストを参照)を対象とした検索を行うことができます。

---

**検索対象への登録**

登録を希望される場合、下記項目を御記入の上、“送信”ボタンを押してください。

組織名 :

所在地 :

電話番号 :

インターネットURL :

担当者名 :

E-mailアドレス :

---



検索ページに戻る

 概要……………検索エンジンの概要と利用方法です。

 検索対象リスト…キーワード検索先のリストです。

 御意見・御感想…この検索エンジンの御意見・御感想をお聞かせください。

 登録……………検索先への登録。

Document Done

図 7 - 5 登録ページ

## 8. 参考文献

- 1) Web プログラミング ,Kris Jamsa Ph.D, Suleiman "Sam" Lalani, Steve Weakley  
共著, (株) スリーエーシステムズ訳, (株) 翔泳社, 1997 年 5 月
- 2) 日本語全文検索システムの構築と活用, 馬場肇 著, ソフトバンク (株) ,  
1998 年 9 月

— 禁無断転載 —

平成11年3月発行

発 行 財団法人 データベース振興センター

東京都港区新橋2丁目13番8号

新橋東和ビル5階

TEL 03-3508-2430

委託先 ミネルバ株式会社

愛知県名古屋市中区錦3丁目5番31号

協栄生命ビル10階

TEL 052-953-1821

