

**e-ビジネス情報技術講座
第7回
検索エンジン**



講師: 片岡 信弘

教科書 第7章

ポイント

- 検索エンジン2種
 - ◆ ディレクトリ型とロボット型
- 検索結果表示順序の重要性
- 順序を決定する仕組み
- SEO
 - ◆ サービスエンジン・オプティマイゼーション

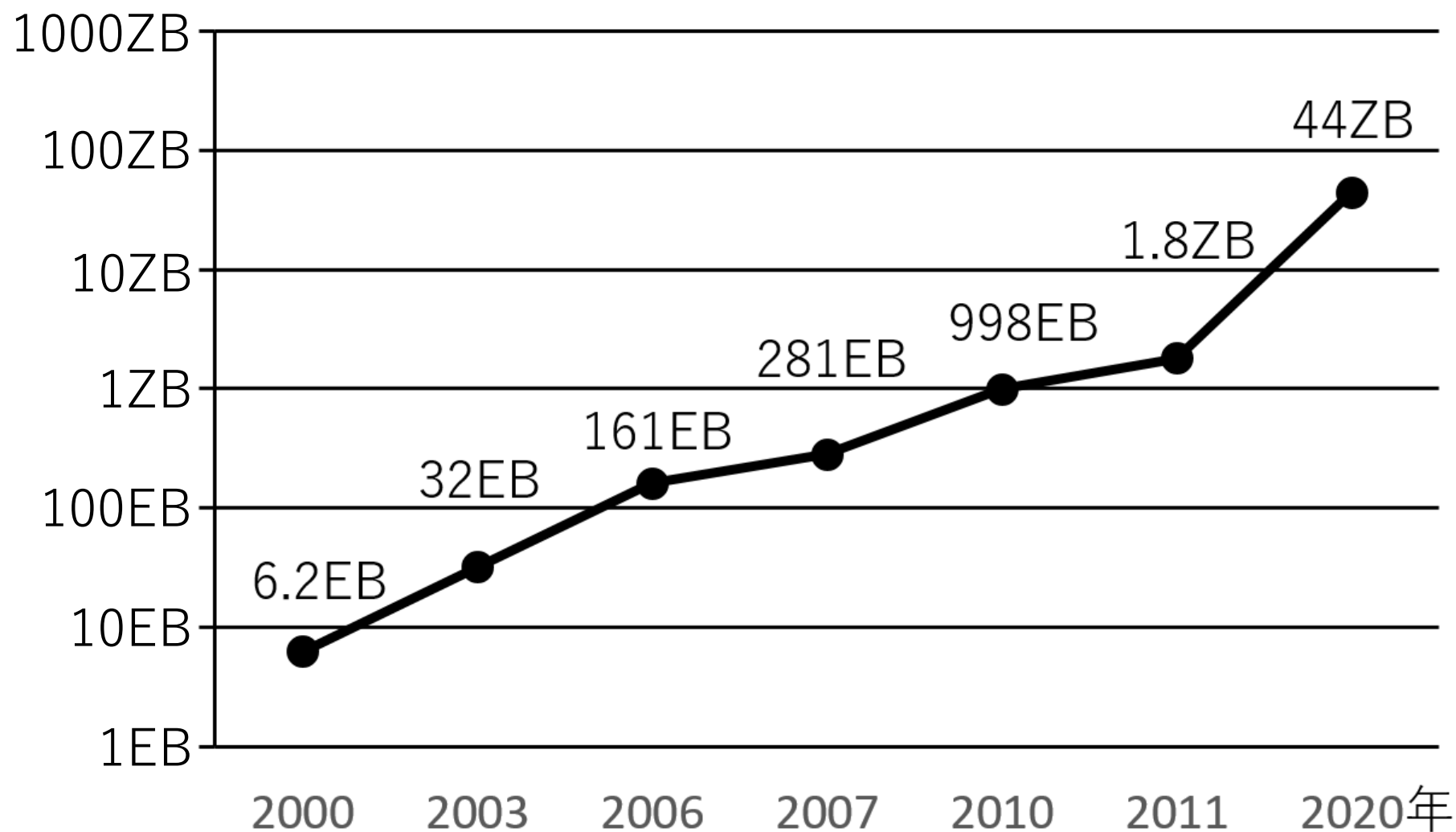
7.1 情報爆発と検索サービス

日々増大するWebサイトから必要な 情報
を見つけ出すには検索エンジンが必須

インターネット上の情報量の増大

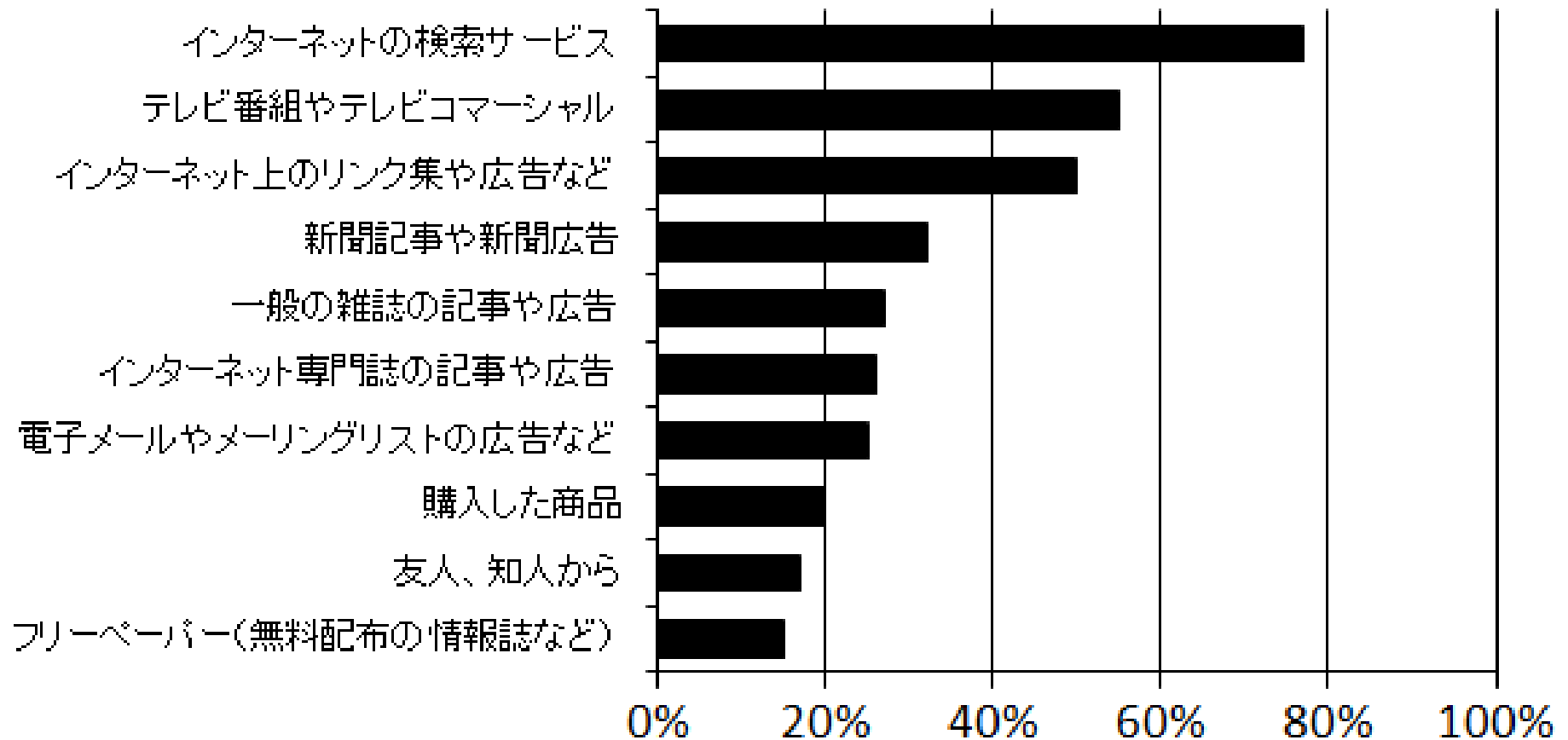
- Webサーバ数 2011年 **3億4千万台**
- Webサーバ上のデータ
 - ◆ 2006年 180EB (Exabyte : 10^{18} byte)
 - ◆ 2011年 1,800EB
 - ◆ 2016年 18ZetaB 5年で10倍
 - ◆ ムーアの法則 1.5年で半導体の集積率2倍
- データの特性
 - ◆ 体系立ってはいない
 - ◆ 日々更新されている
- 人々はどのようにして必要な情報を入手しているか

図7.1 Webサイトのデータ量の推移 P80



総務省：平成27年版 情報通信白書 より作成

図7.2 Webサイトの認知経路の調査結果P80



Yahoo! JAPAN: 第26回 インターネット利用者
アンケートーウェブサイトの認知経路 による

検索サービスの種類

- カテゴリ別に分類された情報の探索
 - ◆ **ディレクトリ型**検索エンジン
- キーワード入力による検索
 - ◆ **ロボット型**検索エンジン

7.2 ディレクトリ型 検索エンジン

ディレクトリ型検索エンジンは世の中のWebサイトを分類整理したもの

ディレクトリの構築方式

- サイトをエディタ(専門家)が**木構造ディレクトリ**に登録
 - ◆ **カテゴリ階層**(大分類, 中分類, 小分類)に分類
 - ◆ 検索結果として質の高い情報が得られる
- Yahoo!ディレクトリ(1994年開始)など初期の検索サービスで主流
- ネット上のページが増大につれて, **特定の専門家**だけでは包括的なディレクトリ**登録維持が困難**
- **Open Directory プロジェクト**(ODP)
 - ◆ **ボランティア**のエディタによってディレクトリを整理する活動を継続
 - ◆ 各国語のディレクトリが整備
 - ◆ **日本語**のディレクトリは**約19万件**のサイトが登録

Yahooカテゴリ <http://dir.yahoo.co.jp/> 2018年3月29日終了

探すキーワードを入力 (例: ニュース) ウェブ全体 Yahoo!カテゴリ全体 PR

新着オススメサイト 2015年05月19日 (火) [もっと見る](#)



全自動DIGA(ディーガ) - Panasonic NEW!
女優・綾瀬はるかと俳優・西島秀俊がCM初共演! オリジナル版のほか、各地の方言を披露した地域限定版も。



いたずらΩオウム タイハクオウムの生活雑記
タイハクオウムの“ろう君”と、その飼い主の日常を描いた4コマ漫画ブログ。KADOKAWAから書籍化も決定!



VERENA
人生の景色を美しく。
大和地所レジデンス 新社名キャンペーン
パリ旅行やヨーロッパの
素敵な賞品が当たる!



新型 iPhone
アクセスをチェック



スマートフォン向け Yahoo!カテゴリ
スマホサイトに
アクセスしよう!



**YouTubeで
動画検索**
「ユーチューブ」で配信中の
動画のまとめサイト大集合。



**日本全国の
地域情報**
47都道府県のローカル情報を
厳選してまとめました。



**ショッピングと
サービス情報**
買い物・グルメから価格比較
まで、お得な情報満載。



新聞
主要5大紙や、芸能・スポーツ
新聞のオンライン版。



競馬・競輪・競艇
レースの会場や開催の状況、
予想・結果をチェック。



18+ 大人のカテゴリ
アダルトな大人のサイト集。
※18歳未満は入場できません



エンターテインメント



メディアとニュース

芸能人 グラビアアイドル 音楽 占い 新聞 芸能、スポーツ新聞 2chまとめ
面白画像 AKB48 コミック、アニメ 動画 テレビ ラジオ 雑誌 週刊誌



箱根・大涌谷情報 NEW!
ライブ配信やTwitterなど、箱
根・大涌谷に関する情報はこち
ら。



浅田真央 (フィギュアスケー

図7.3 カテゴリーによる検索 P81



[ウェブ](#) [画像](#) [グループ](#) [ニュース](#) [ショッピング](#) [地図](#) [もっと見る »](#)

Google 検索

[表示設定](#)
[ディレクトリヘルプ](#)

カテゴリ別Google!

[アート](#)

[音楽](#), [パフォーミングアート](#), [文学](#), ...

[ニュース](#)

[テレビ](#), [ラジオ](#), [新聞](#), ...

[地域](#)

[アジア](#), [ヨーロッパ](#), [北アメリカ](#), ...

[オンラインショップ](#)

[食品](#), [家庭・園芸](#), [ファッション](#), ...

[ビジネス](#)

[食品](#), [製品・サービス\(産業向け\)](#), ...

[家庭](#)

[料理](#), [ガーデニング](#), [住居・マイホーム](#), ...

[ゲーム](#)

[ビデオゲーム](#), [テーブルゲーム](#), ...

[レクリエーション](#)

[アウトドア](#), [ペット](#), [自動車](#), ...

[社会](#)

[宗教・精神世界](#), [団体](#), [政治・行政](#), ...

[コンピュータ](#)

[ソフトウェア](#), [インターネット](#), ...

[健康](#)

[体調・症例](#), [医学](#), [動物](#), ...

[科学](#)

[自然科学](#), [社会科学](#), [人文科学](#), ...

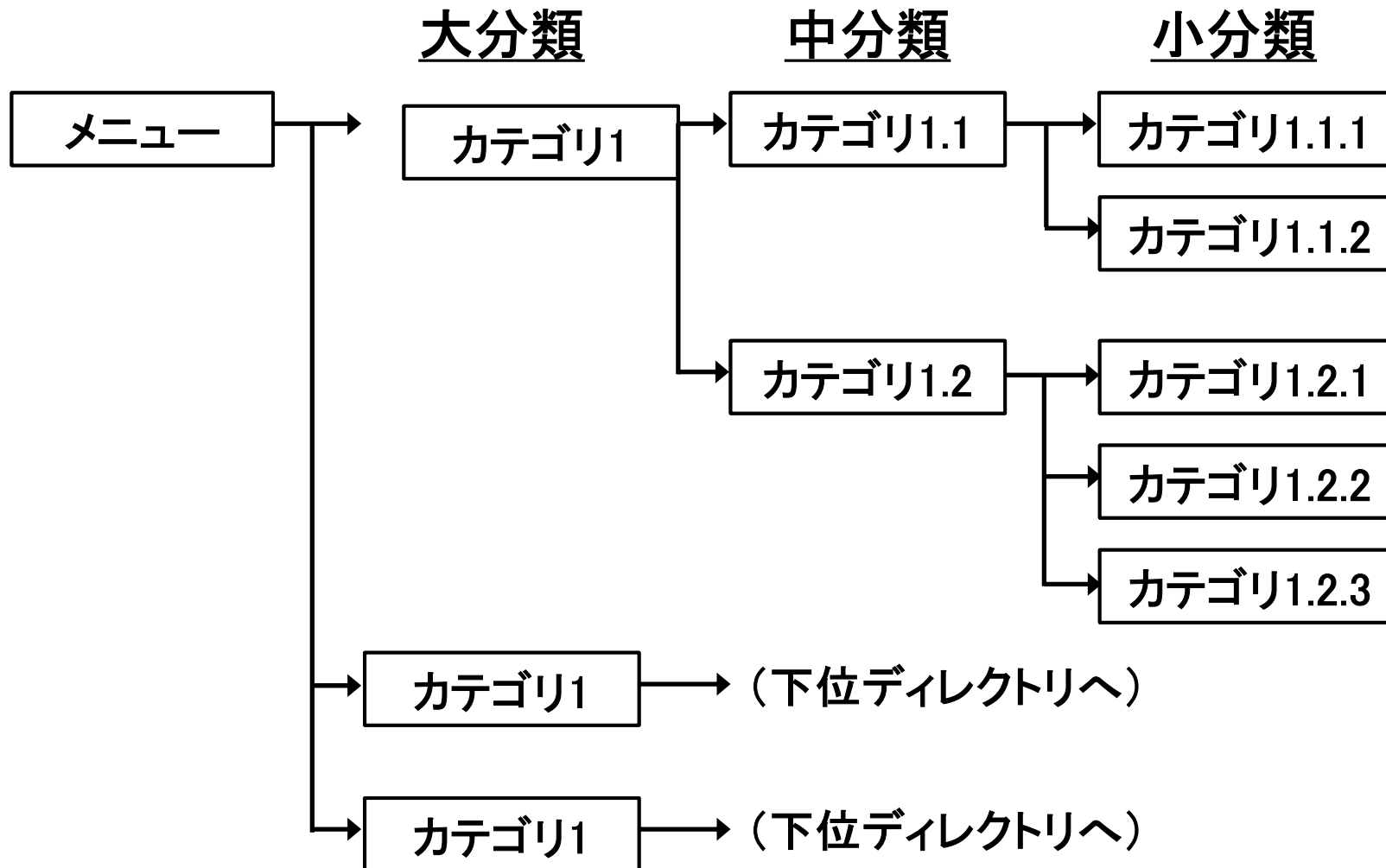
[スポーツ](#)

[サッカー](#), [武道・武術](#), [野球](#), ...

[各種資料](#)

[教育](#), [文化施設](#), [地図](#), ...

図7.4 ディレクトリの構成 P82



オープン ディレクトリのトップ <http://www.dmoz.org/> 番外

ODP - Open Directory Project

dmoz open directory project In partnership with AOL Search.

[about dmoz](#) | [dmoz blog](#) | [suggest URL](#) | [help](#) | [link](#) | [editor login](#)

Search [advanced](#)

Arts
[Movies](#), [Television](#), [Music](#)...

Business
[Jobs](#), [Real Estate](#), [Investing](#)...

Computers
[Internet](#), [Software](#), [Hardware](#)...

Games
[Video Games](#), [RPGs](#), [Gambling](#)...

Health
[Fitness](#), [Medicine](#), [Alternative](#)...

Home
[Family](#), [Consumers](#), [Cooking](#)...

Kids and Teens
[Arts](#), [School Time](#), [Teen Life](#)...

News
[Media](#), [Newspapers](#), [Weather](#)...

Recreation
[Travel](#), [Food](#), [Outdoors](#), [Humor](#)...

Reference
[Maps](#), [Education](#), [Libraries](#)...

Regional
[US](#), [Canada](#), [UK](#), [Europe](#)...

Science
[Biology](#), [Psychology](#), [Physics](#)...

Shopping
[Clothing](#), [Food](#), [Gifts](#)...

Society
[People](#), [Religion](#), [Issues](#)...

Sports
[Baseball](#), [Soccer](#), [Basketball](#)...

World
[Català](#), [Dansk](#), [Deutsch](#), [Español](#), [Français](#), [Italiano](#), [日本語](#), [Nederlands](#), [Polski](#), [Русский](#), [Svenska](#)...

[Become an Editor](#) Help build the largest human-edited directory of the web

Copyright © 2011 Netscape

4,882,672 sites - 91,365 editors - over 1,006,912 categories

オープン ディレクトリは 2017/3/17 にクローズ



As of Mar 17, 2017, dmoz.org is no longer available.

The editors have set up a static mirror [here](#).

ディレクトリ型検索エンジンの検索事例(インターネット)



[Top](#): [World](#): [Japanese](#): [ビジネス](#): [IT](#): インターネット (360)

パンくずリスト

- [Web制作](#) (162)
- [アクセス・プロバイダ](#) (80)
- [ドメイン名登録](#) (33)
- [ハウジングとホスティング](#) (69)
- [メール・プロバイダ](#)@ (27)
- [広告・マーケティング](#)@ (81)

直下のカテゴリ

こちらをご参照ください:

関連カテゴリへのリンク

- [World: Japanese: コンピュータ: インターネット](#) (910)

現在のカテゴリのページ

- [NETMAP](#) - 国際航業によるインターネット用などの地図作成サービス。
- [2021/3/6 アドミラルシステム](#) - 埼玉県川口市。ホスティング・グループウェアのASP事業など。会社概要、事業紹介。
- [オズカー電子](#) - 日本語ドメイン名の登録 転送サービス ホスティングサービスを扱う サービス案内

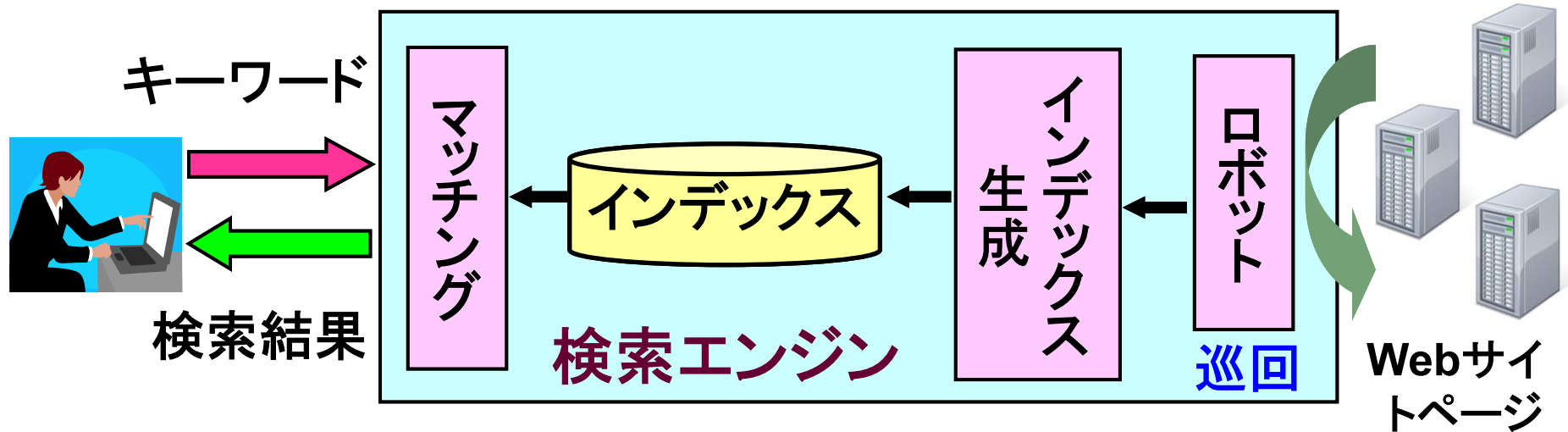
7.3. ロボット型エンジン

多数のロボットが世界中のWebサイトの
ページを検索してインデックスを作る

ロボット型検索エンジンの仕組み

- ロボット(スパイダー, クローラ)がインターネット上を巡回
 - ◆ 自動的にページインデックスを生成
- 巡回で検出したページを登録して定期的に巡回
 - ◆ ページから他のリンク先ページに移動して登録
- ロボットは頻繁に更新されるサイトは頻繁に訪問
 - ◆ 効率的な巡回を目指している
- ディレクトリ型検索エンジンは**サイト単位**分類, ロボット型は**ページ単位**検索

図7.6 ロボット型エンジン P83



検索結果の表示順位の重要性

- 企業にとって検索結果はユーザを**自社サイトに誘導**する重要な手段
 - ◆ 75%のユーザが検索結果をクリックする
- ユーザが閲覧は検索結果の**3ページ**程度
- 利用者の課題
 - ◆ 検索**要望**に**相応しい**ページを表示すること
- 企業の課題
 - ◆ 自社ページが**検索の上位**に表示されること

図7.9 ユーザが閲覧する検索結果の画面数 p85

<http://markezine.jp/article/detail/3564?p=2>より

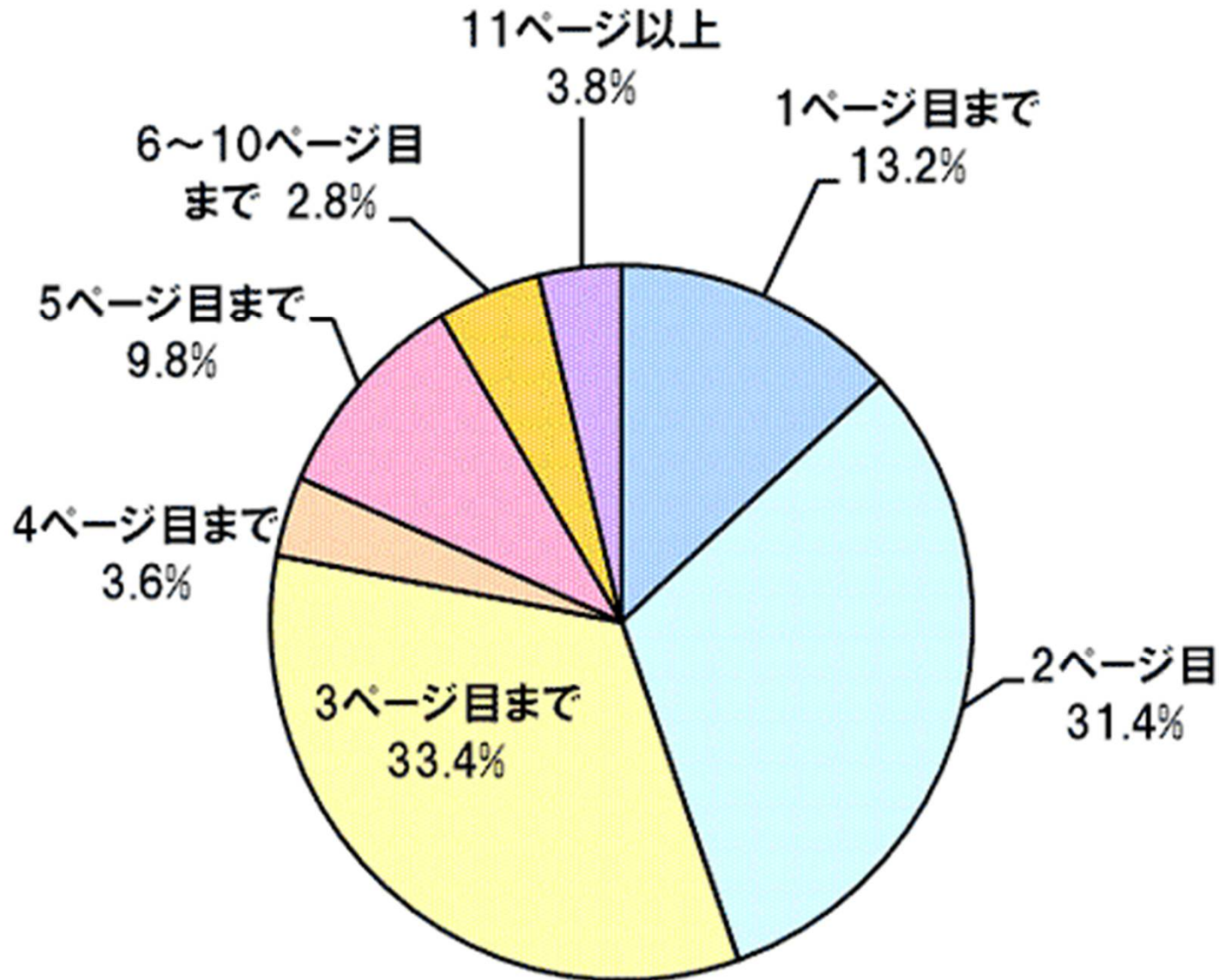


図7.12 検索順位とクリック数の関係 p87

SEMリサーチ:[調査] SEO - 検索順位とクリック数の関係 - 米AOLの検索行動データより, <http://www.sem-r.com/08h1/20080626172700.html>.

自然検索での検索順位とクリック数の関係 (AOL, 2006)

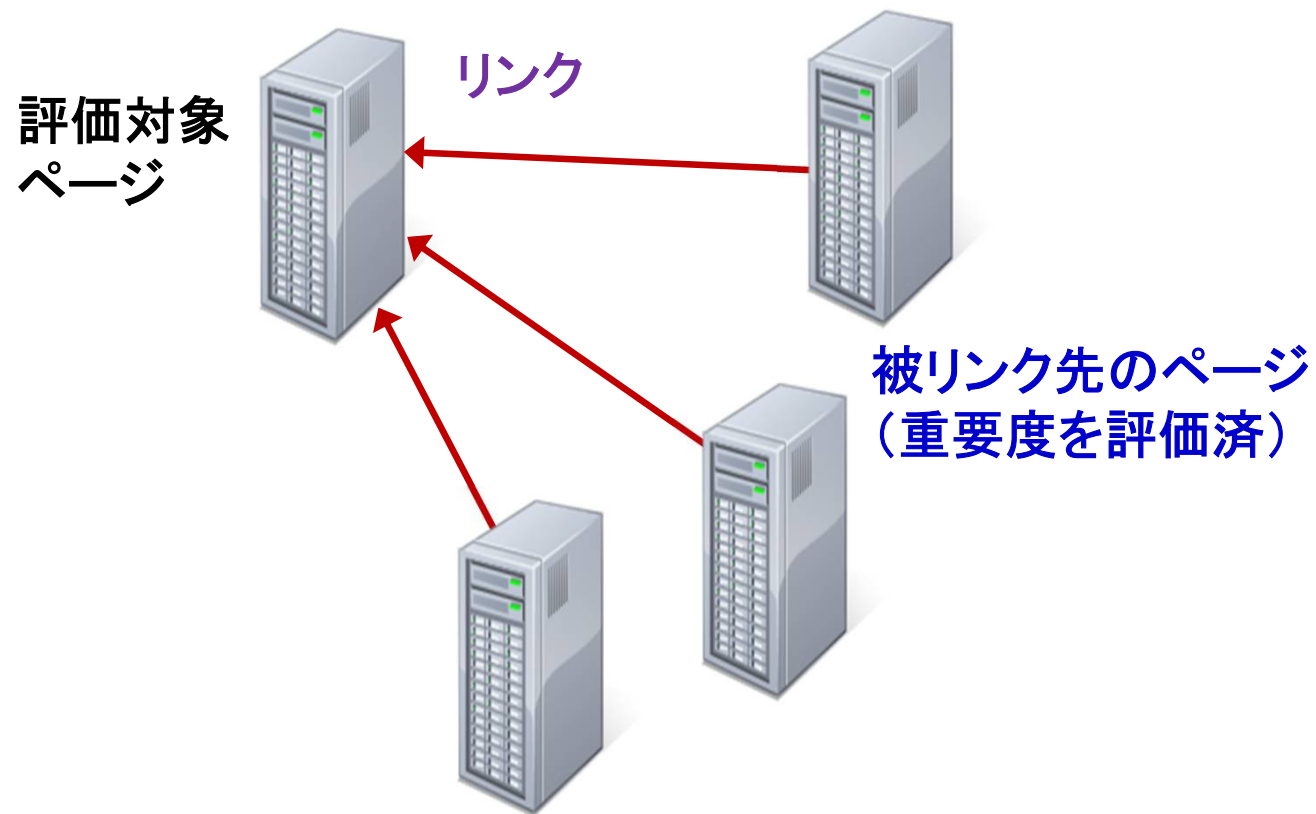
全検索数 9,038,794
全クリック数 4,926,623

順位	クリック数	クリック率	クリックシェア	1位との差(倍)
1	2,075,765	22.97%	42.13%	
2	586,100	6.48%	11.90%	3.5
3	418,643	4.63%	8.50%	5.0
4	298,532	3.30%	6.06%	7.0
5	242,169	2.68%	4.92%	8.6
6	199,541	2.21%	4.05%	10.4
7	168,080	1.86%	3.41%	12.3
8	148,489	1.64%	3.01%	14.0
9	140,356	1.55%	2.85%	14.8

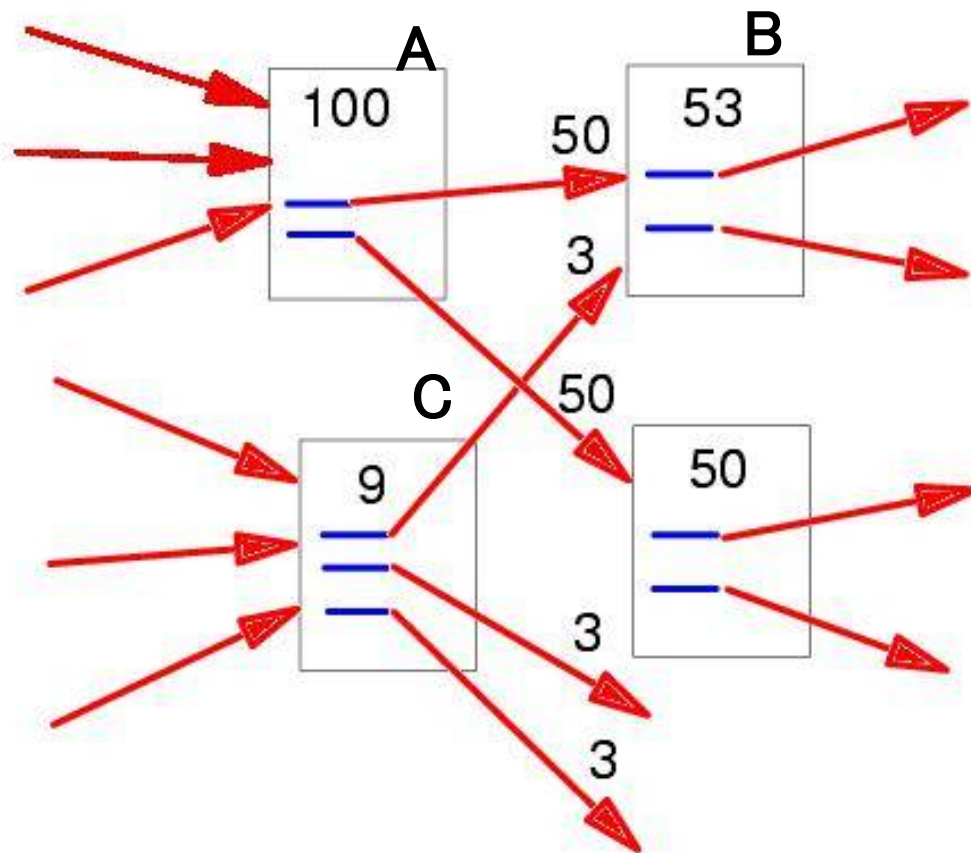
PageRankテクノロジー(Googleの技術)

- PageRankとはページ重要度を評価したもの
- 多数の重要度の高いページから参照されるページは重要度が高いと評価
- 評価対象ページに対する他のページからのリンクを投票とみなす
- 評価は被リンク先のページ数とそのページの重要度によって評価
- 重要度の低いページから多数のリンクを受けても重要度は増加せず
- 重要度の高いページの表示は優先している。

図7.10 PageRankテクノロジーの仕組み p86



PageRank計算方式模式 番外



Larry Page, 'PageRank: Bringing Order to the Web', <http://hci.stanford.edu/~page/papers/pagerank/>

7.4 SEO (Search Engine Optimization)

Webサイトに手を入れ検索エンジンで
上位に表示させるよう適切化すること

SEOとは

- SEM(Search Engine Marketing)で検索サービスから自社サイトへ誘導
 - ◆コンバージョンを狙う
- SEOは自社のページが上位に表示されるように、自社サイトを最適化すること
- ユーザが検索するキーワードをHPに組み込む
 - ◆タイトルや見出しにキーワードを入れたり、キーワードを意識した本文を制作すること
- 自分でHPを作成した時どうする

「サプリメント 通販」検索 第1位

ロサンゼルスから直送！高品質と低価格を両立できるアメリカのサプリメント専門店

サブリンクスと

サブリンクス 本店
SUPLINX

最短5日でお届け、送料**390円** 今だけ
サプリメントアドバイザー常駐・信頼と実績で17周年

ログインしてお買い物するとポイントがたまり
▶アマゾンログイン&ペイメント **amazon payments**

商品数 **3,372** 点

成分や商品名・商品番号を入力 **検索** **マイページ** **お気に入り** **注文履歴**

カテゴリから探す

- ビタミン・ミネラル
- ヘルスケア (お悩み)
- スポーツ
- ダイエット
- ビューティー
- フード&ドリンク
- 癒し&グッズ

新着 ▼ SALE ▼ 定期購入 ▼ お役立ち特集 ▼

サブリンクス本店 **You Tube** 限定 **200円OFF** **クーポン配布中!**
チャンネル登録してね!
動画を見たあなただけに! 1円以上のご購入で使える
有効期間 6/30(木) 14:59まで

サブリンクス本店 リニューアル!
全国どこでも 送料390円!
お友だち紹介で ポイントGET
動画を見た方にお得なクーポン!
サブリンクスの オリジナル商品

200位



妊娠・授乳ママに選ばれる赤ちゃん和妈妈のための葉酸サプリメントとは
ライフプラス葉酸サプリメント通販
葉酸配合総合栄養バランスサプリメント「デイリーバイオベーシックス」紹介サイト

商品および注文に関するお問い合わせは
【通話料無料】平日5:30~13:00
当サイトに関するお問い合わせは、

トップページ TOP PAGE	ライフプラスのこだわり CONCEPT	ご注文方法 ORDER	愛用者の声 VOICE	よくある質問 FAQ
--------------------	------------------------	----------------	----------------	---------------



-  はじめてご購入ガイド
-  なぜ世界中で人気のライフプラス
-  お得に購入できる定期購読
-  妊娠 & 授乳ママの愛用者

～ 赤ちゃんが欲しいとおもったら～

妊娠・出産・産後に役立つ プレママとママのための葉酸サプリメント情報サイトです♪
健康な赤ちゃんを産み育てるには、まずお母さんの体内環境を整えることが大切！先輩ママの体験談とともにその重要性も紹介してい



サイト運営者情報

こんにちは♪当サイトを運営しているのは、ママのゆきです♪

妊娠中は、無事に育って

サプリメントで検索 200位ソース

ファイル(F) 編集(E) 書式(O)

```
1 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
  "http://www.w3.org/TR/html4/loose.dtd">
2 <html lang="ja">
3 <head>
4 <meta http-equiv="Content-Type" content="text/html; charset=Shift_JIS">
5 <meta http-equiv="Content-Style-Type" content="text/css">
6 <meta http-equiv="Content-Script-Type" content="text/javascript">
7 <meta name="GENERATOR" content="JustSystems Homepage Builder Version 16.0.1.0 for Windows">
8 <meta name="Keywords" content="ライフプラス,葉酸サプリメント 通販">
9 <meta name="Description" content="ベビ待ちママから妊娠・授乳ママ・子育てママに選ばれる赤ちゃんと
  ママのための葉酸サプリメントとは？ライフプラスの葉酸配合総合栄養バランスサプリメント「デイリーバ
  イオベシックス」 通販サイト">
10 <title>ライフプラス葉酸サプリメント 通販</title>
11 <link rel="stylesheet" href="hpbparts.css" type="text/css" id="hpbparts">
12 <link rel="stylesheet" href="container_5A_2c_top.css" type="text/css" id="hpbcontainer">
13 <link rel="stylesheet" href="main_5A_2c.css" type="text/css" id="hpbmain">
14 <link rel="stylesheet" href="user1.css" type="text/css" id="hpbuser">
15 </head>
16 <body id="hpb-template-05-01-01" class="hpb-layoutset-01">
17 <div id="hpb-skip"><a href="#hpb-title">本文へスキップ</a></div>
18 <!-- container -->
19 <div id="hpb-container">
20 <!-- header -->
21 <div id="hpb-header">
22 <div id="hpb-headerMain">
23 <h1>妊娠・授乳ママに選ばれる赤ちゃん和妈妈のための葉酸サプリメントとは</h1>
```

検索エンジンのアルゴリズム

- テキストマッチング
 - ◆ 形態素解析 単語の切り出し
- キーワード
 - ◆ 出現頻度/近接度/出現位置
- タグごとの重み付け <head> <body>
- リンクポピュラリティ PageRankの考え方
- アンカーテキスト(クリックできる文字列)の重要視

番外 検索エンジン市場

■ 世界

◆ Google 69%

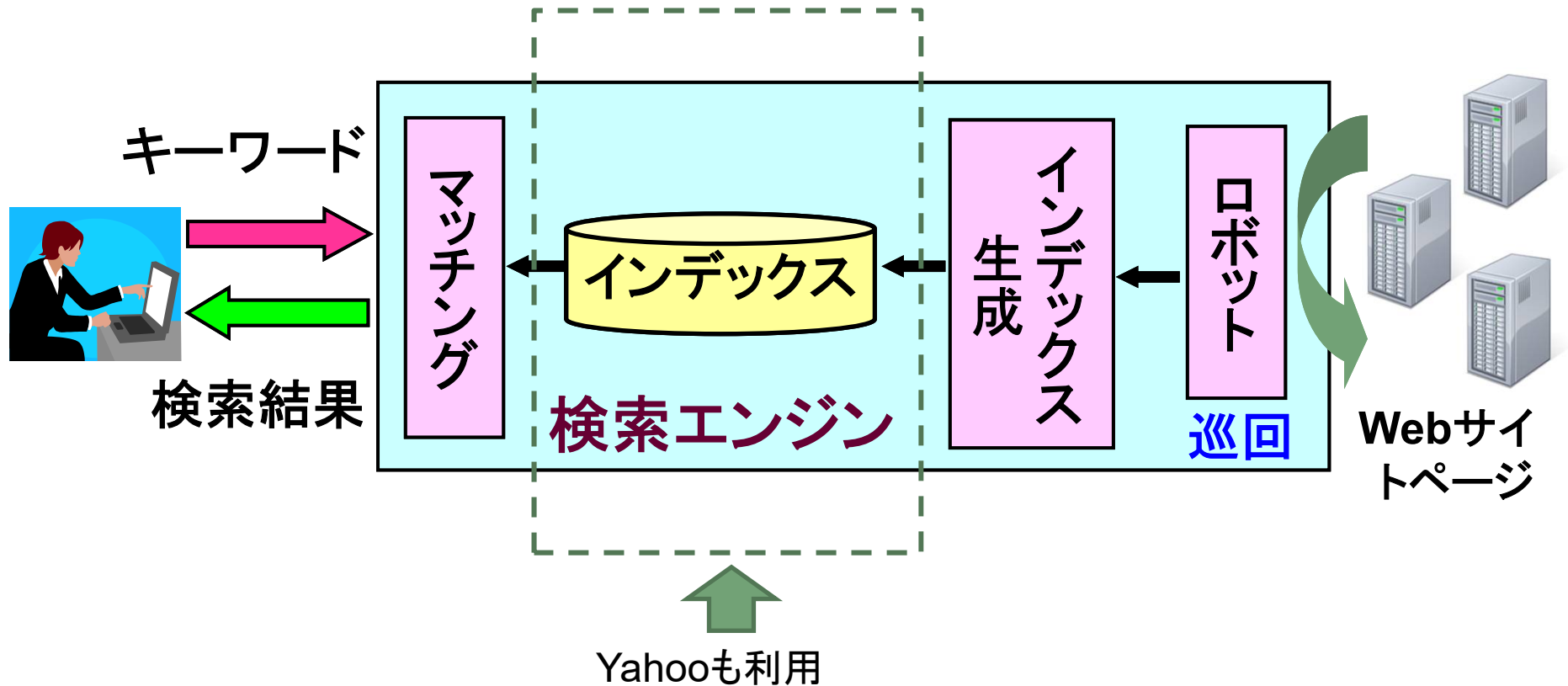
◆ マイクロソフト 10%

◆ Baidu (百度) 9%

□ <https://www.baidu.com/>

(注)YahooはGoogleの検索エンジンを利用

Google検索エンジン



まとめ

- **ディレクトリ型検索エンジン**
 - ◆ Open Directory プロジェクトがカテゴリ制作
- **ロボット型エンジン**
 - ◆ 多数のロボットが世界中のサイト検索しインデックスを作成
 - ◆ 技術開発とシステム運用に膨大な費用が必要
- **SEO**
 - ◆ 自社のページが**上位に表示**を目指し、自社サイトを最適化すること、各社が必死になっている

FIN