

メタ検索エンジン

- 複数の検索エンジンを同時に並行的に検索
- 検索デスク SearchDesk
 - <http://www.searchdesk.com/>
 - 渡り検索型メタ検索エンジン
 - キーワード入力はそのままで次々と異なった検索エンジンへ
- Ceek.jp
 - <http://www.ceek.jp/>
 - 構断検索型メタ検索エンジン
 - 複数の検索エンジンを同時検索
 - 検索結果を統合表示

19

第5回演習課題：好きなテーマで 2つの検索エンジンを検索して比較

- 検索エンジンによって検索結果が大きく変わることを実際に検索してみても実感してみよう！
- Yahoo! JapanとGoogleを両方を同じ検索語で検索
 - 10件以上ヒットする検索語を選ぶ
 - 2つの検索結果を見比べてみる

20

アーカイブ・サイト(1)

- 定期的にWWW上の情報資源を保存・提供
- 過去の情報が見れる
 - 検索エンジンでは、現在のWebページしかみれない(「404 Not Found Error」)
- 検索可能(URLやキーワードなど)

第7回資料

21

アーカイブ・サイト(2)

- WayBackMachine:
<http://archive.org/web/web.php>
 - 巨大なアーカイブ・サイト
 - 1996年以降から収集開始
 - 世界の多くのサイトを収集
- WARP: <http://warp.ndl.go.jp/>
 - 国立国会図書館のウェブ・アーカイブ事業
 - 日本の中央省庁、立法機関、都道府県、各種法人、国立大学のホームページなど

第7回資料

22

第6回演習課題

- 日本首相官邸の過去のWebページをアーカイブ・サイトで見てみよう。
 - 首相官邸ホームページ:
<http://www.kantei.go.jp/>
- 0. その前に現在の首相官邸ページをみてみよう
 - 1. IEを起動
 - 2. <http://www.kantei.go.jp/> を入力し「Enter」

第7回資料

23

演習

— 1. WayBackMachine—

1. WayBackMachineのサイトに行く
 - <http://archive.org/web/web.php>
2. 首相官邸のホームページのURLを入力
 - <http://www.kantei.go.jp/>
3. 「Take Me Back」ボタンをクリック
4. 日付のリンク(例:Nov 08,1996)をクリックして過去のWebページをみる
 - 1つだけではなく、いろいろみるみる！
 - 例:1997年の首相は誰だったかな？
 - 例:2000年は？

第7回資料

24

演習

— 2. WARP—

1. WARPのサイトに行く
 - <http://warp.ndl.go.jp/>
2. キーワード:「首相官邸」を入力
3. 「検索」ボタンをクリック
4. 「首相官邸／内閣官房」のリンクをクリックして過去のWebページをみる
5. 「2004年11月19日」収集 のリンクをクリック

第7回資料

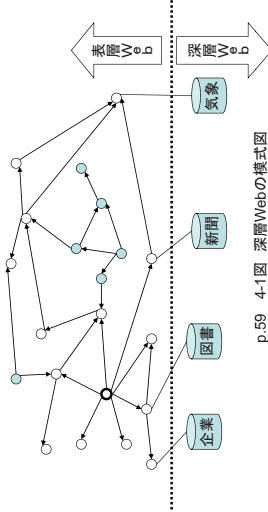
25

ハイブリット型検索エンジン

- デイレクト型とロボット型を組み合わせたもの
 - 最近ではほとんどがこのタイプ
 - 大抵どちらからのタイプが優先的に
 - Yahoo! はデイレクト型優先のハイブリット型
 - ロボットで収集したサイトも検索できる。
 - Googleはロボット型優先のハイブリット型
 - 「デイレクト」で、カテゴリでブラウジングできる。
 - 「デイレクト」のリンクをたどってみよう

10

表層Webと深層Web



p.59 4-1図 深層Webの模式図

13

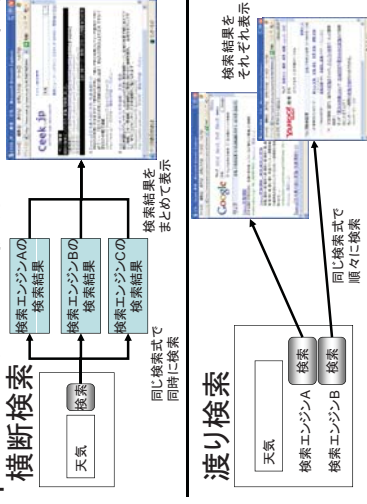
Google(2)

- 検索オプション(詳細検索)
 - **語の区切りには半角**を使うこと！
 - 論理演算を使った検索
 - ページの記述言語の指定
 - 情報のファイルタイプの指定: doc, 画像ファイル
 - 情報の更新日付
 - ドメイン指定
 - 検索式で書けるものもある

16

メタ検索エンジン

p. 40 3-2図 メタ検索エンジンの種類



11

Yahoo! Japan: <http://www.yahoo.co.jp/>

- 日本でよく利用されている
- アナリストは200人といわれている
- 日本人に有用と考えられるページを収集
- 簡単な解説をサイト毎に付与
- Yahoo!カテゴリー
 - リンクをたどって「教育図書館」← p.44 検索例-1
- 登録サイト検索
 - 「ユネスコ」を検索 ← p.45 検索例-2
- Web検索
 - キーワード検索: 「島津製作所 ノーベル賞」 ← p.45 検索例-3
 - 検索オプション(詳細検索)

14

Google(3)

- p.48 検索例-5
 - 久留米出身の洋画家で坂本繁二郎を除く青木繁、古賀春江のサイトを検索する
- P.49 検索例-6
 - 図書館と著作権に関する、日本語のページで、日本の省庁(go.jp)のページを検索したい。

17

検索エンジン利用の一般的な注意

1. 表層Webしか検索できない
2. 検索結果の質や信頼性を確認する
3. ヘルプを活用する
4. 検索エンジンによって結果が大きく異なる
5. キーワードはWebページに表記されたものが対象となる
6. 検索式の入力には英数字、記号は半角を使う

12

Google(1): <http://www.google.co.jp/>

- Google: 検索式
 - AND検索: 「半角スペース」
 - OR検索: 「OR」
 - Not検索: 「-」
 - フレーズ検索: 「"」でくくる
- 「I'm Feeling Lucky」ボタン
 - 最高位の検索結果を表示
 - 演習: 「気象庁」 p.47 検索例-4

15

Google(4)

- イメージ検索
 - 画像が検索できる
 - 演習: 「空」 p.49 検索例-7

18

情報検索サービス Webページ、Webサイト

聖徳大学 書補講習
江草田佳
国立教育政策研究所
yuka@nier.go.jp

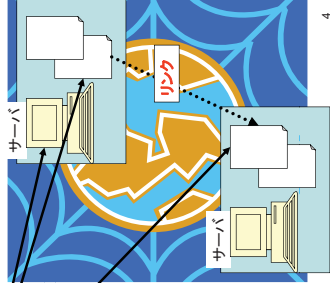
1



WWWクライアント:
WWWブラウザ: Internet Explorer



Web (2)



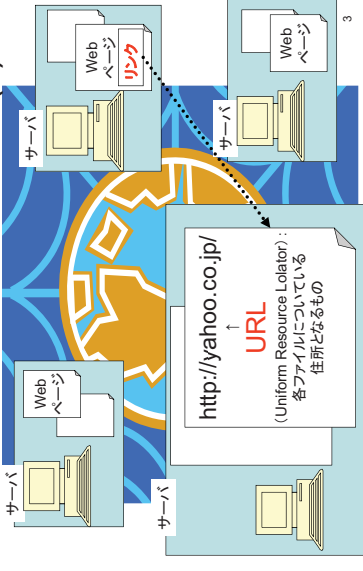
4

ロボット型検索エンジン

- 代表的なもの: Google, Goo
- ロボット (収集プログラム) でページを収集
 - 登録しているサイトは膨大
 - 検索結果の精度が低くなりがち
- ランキング (スコア)
 - 関連の高いものから順に表示する

7

WWW: World Wide Web(1)



2

本日のお品書き

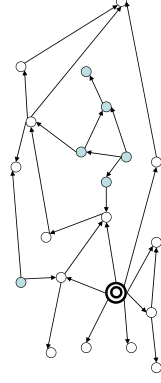
- WWWとは
- 検索エンジン
 - 検索エンジン (Yahoo! Japan, Google)
 - メタ検索エンジン (検索デスク、Ceek.jp)
 - 第5回演習課題
- アーカイブサイト
 - WayBackMachine
 - WARP

Web(3)

- Webページ
- Webサイト
- ホームページ、トップページ
- 検索エンジン
- アーカイブサイト
- サブジェクトゲートウェイ、リンク集
- サイト内検索、サイトマップ
- ポータルサイト

5

p.38 3-1図 Webの模式図



8

検索エンジン

- サーチエンジンとも言う
- 種類
 - デイレクトリ型 (カテゴリ型)
 - ロボット型
 - ハイブリット型
 - メタ検索エンジン
- ポータルサイトの核となる一機能

6

デイレクトリ型検索エンジン

- カテゴリ型検索エンジンともいう
- 代表的なもの: Yahoo! カテゴリ
- 人間がページを収集
 - 一定の収集基準がある
 - 有用なサイトが得られる確率が高い
 - ロボット型に対して登録されているサイトが少ない
- 人手でページの説明情報を作成
- 人手で適切な階層 (カテゴリ) に位置づけ
 - 求める主題やトピックをたどってブラウジングができる
- 検索システムは階層や説明情報も検索

9