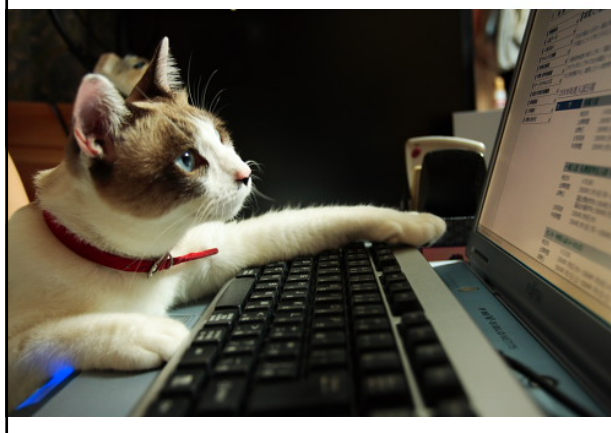


専門データベースと 検索エンジン (情報検索演習)

2015
時実 象一
tokizane@aichi-u.ac.jp



検索エンジンと専門情報データベース



3

検索エンジンと専門データベース



- 検索エンジン (サーチエンジン)
 - インターネットのホームページ (Web ページ) の検索ツール
 - Google, Yahoo!, bing, Goo など
- 専門データベース
 - 特定の情報を収集・選択・加工して提供する検索ツール
 - OPAC, 新聞記事、雑誌記事、百科事典

4

検索エンジン

- ロボット (Crawler と呼ばれる) が収集
 - Googlebot, Yahooobot, など



専門データベース

- 専門家が収集・選択・加工
 - 新聞記事データベースの例



専門データベースの例

- OPAC
 - 図書館の蔵書
- NDL-OPAC 雑誌記事索引
 - 雑誌の記事
- Webcat (CiNii Books)
 - 全国の大学の蔵書
- 新聞記事データベース
 - 新聞記事
- 特許庁電子図書館
 - 日本の特許文書



7

検索エンジンと専門データベースの違い

- 含まれている情報
 - 検索エンジン
 - ホームページに記載されている情報でロボットが収集
 - 玉石混交
 - 非常に情報量が多いが網羅性はない
 - 専門データベース
 - 特定の分野で特定の目的で作成
 - 人手で集められ、選択され、作成された情報
 - 評価済みで信頼性が高い
 - その分野では網羅性がある



8

検索エンジンと専門データベースの違い

- 検索項目
 - 検索エンジン
 - ページ内のすべての語句 (全文)
 - フィールド (検索項目) の区別はない
 - 書誌情報がない
 - カマカセの検索
 - 専門データベース
 - 選択したデータまたは作成したデータ
 - フィールドに分かれている
 - 書誌情報がある
 - データの形式 (たとえば年表示) が整えられている
 - 統制語 (あらかじめ定義された語) がある場合がある
 - きめ細かい検索が可能



9

検索エンジンと専門データベースの違い

- 情報の遡及と記録
 - 検索エンジン
 - 過去の情報が少ない (5-15 年)
 - インターネットがそもそも新しい (1995 年頃から)
 - 情報が更新されると前の情報は消される
 - 古いサイトや情報は失われる
 - 専門データベース
 - 過去の情報が充実 (10-100 年)
 - 記録として保存される



10

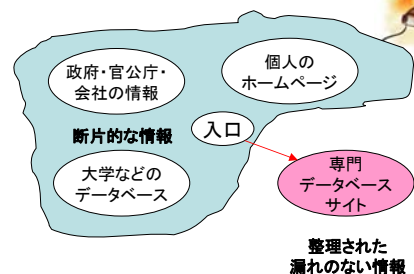
検索エンジンの限界

- ロボット索引の限界
 - 索引を禁止されているページには入れない
- 専門情報データベースの中身 (deep web) には入れない
 - つまり
 - OPAC の入口までは索引するが、
 - OPAC の中身は索引しない
 - **検索エンジンは専門データベースの中身は検索できない!**



11

インターネットの情報



12

検索エンジンと専門データベースの違い



- 利用形態
 - 検索エンジン
 - 無料で使える (広告料でまかなっている)
 - 専門データベース
 - しばしば有料
 - 図書館で購読していれば使える
 - 新聞記事データベースなど
 - 公的機関が作成しているものは無料
 - OPAC、特許データベース

13

検索エンジンと専門データベースの違い



- 検索エンジンの強い分野
 - 調べ方がわからないトピック
 - ことばの意味など
 - 特殊なトピック
 - 官公庁文書、公開文書
 - 生活・趣味などの分野
- 専門データベースの強い分野
 - ビジネスに必要な情報
 - 研究開発に必要な情報
 - 新聞記事・雑誌記事
 - 公的機関が作成するデータベース (判例、特許)
 - その他専門的な情報
- レポートや卒論ではなるべく専門データベースを使う