

Title	6.情報検索の実際：検索エンジンで検索できないこと
Author(s)	安形, 輝
Citation	聖学院大学図書館情報学研究, 第6号 寄附講座「インターネット時代の情報資源活用」特集号, 2011.3 : 55-57
URL	http://serve.seigakuin-univ.ac.jp/reps/modules/xoonips/detail.php?item_id=3362
Rights	

SERVE

聖学院学術情報発信システム：SERVE

SEigakuin Repository for academic archiVE

6 情報検索の実際

－検索エンジンで検索できないこと－

安 形 輝

概要

人々が何がしかの行動を行うさいに、まず検索エンジンを調べてから、ということが多くなってきている。このことは、検索エンジンの利用数にも顕れている。例えば、米国の人口2億人強による一ヶ月あたりの総検索数は延べ100億回を超えている(2008年5月時点)。検索エンジンは幅広く利用され、様々な行動に影響を与えるため、公益性が高まっていると考えられる。しかし、検索エンジンのシェアは私企業であるGoogleの独占あるいはGoogleとYahoo!の寡占という状況になっている。講演では検索エンジンが私企業によって独占あるいは寡占であることの問題を二つの点から紹介した。

一つ目は、検索エンジンがインターネット全体をどの程度カバーしているかである。機関リポジトリを用いて代表的な検索エンジンがカバーしている割合を推計した。調査からは公開されたウェブのうち、Google単独で検索できるのは5割、Yahoo!、Bingと合わせて検索できるのは7割程度となった。

二つ目は、検索エンジンが意図的に検索結果を排除した事例を紹介した。一般的に「初音ミク」事件と言われ、2007年10月頃にGoogle、Yahoo!の画像検索において初音ミク(音声合成ソフト)に関連した画像が検索されなくなった事件である。同時期にWikipediaでも関連項目が消去されている。両検索エンジンの公式発表では、意図的な削除はなかったとされている。しかし、競争下にある検索エンジンのどちらによっても同時に検索できなくなったこと、同時にWikipediaの項目も削除されたことに関する合理的な解釈は難しく、不自然な現象と言わざるを得ない。つまり、何かしらの操作があったことを示唆している。

以上のことから、重要なことについて検索を行う場合には、1) できるだけ複数の検索エンジンを用いてできるだけ網羅的に検索を行ったほうがよい、2) 検索結果は必ずしも信頼できるとは限らないことを留意すべきである、ということになる。

検索エンジンで検索できないこと

安形輝(亜細亜大学)

1 検索エンジンの公益性

1.1 検索エンジンの利用とシェア

検索エンジンの利用は多い

Google による独占あるいは Google と Yahoo!による寡占

1.2 検索エンジンの収益源

ターゲット広告 : Google アドワーズ、Google アドセンス

2 検索エンジンのカバー率

仕組み上、検索エンジンが検索できないこと

2.1 深層ウェブの調査

深層ウェブとは何か

検索エンジンによるページ収集方法

機関リポジトリを用いたカバー率の調査方法

2.2 検索エンジンのカバー率と重複率

公開されたウェブのうち、Google単独で検索できるのは5割、Yahoo!、

Bingと合わせて検索できるのは7割程度

3 検索エンジンからの排除の事例

検索エンジンが意図的に排除した事例の紹介

3.1 Google 八分

3.2 初音ミク事件

Google、Yahoo!などの画像検索から、一時期、ある種の画像がいつせいに排除された事件

[おまけ]

- ・インターネットへの情報流出とその対策