

検索エンジンはどうのようにして有用なサイトと有用でないサイトを判別しているのか

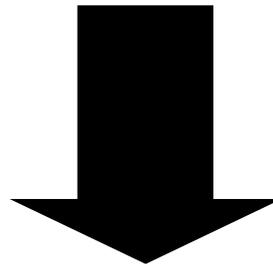
検索エンジンはどうのようにして
有用なサイトと有用でないサイトを判別しているのか

4A 電気電子情報学科 200502716 林 潤

なぜその研究が必要か？

インターネット上でサイトへのアクセス数を増やす(集客する)には検索エンジンの結果の上位表示が必要不可欠である。

Google™



YAHOO!
JAPAN®

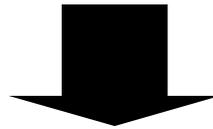
どのようにして検索エンジンが有用なサイトかどうかを判別しているのかが分かれば、自らのサイトを上位表示させることができ、集客が可能になる。

どのようにしてそれを解決するか？

GoogleやYahooでは正式な判別基準の情報は公開されていない。



日々検索エンジンのアルゴリズムの変化を追い、上位に表示されているサイトの特徴を研究し、検索エンジンが有用かどうかを判別するための要因をまとめる。



実際に自分でその要因に沿ったサイトを作ってyahooやgoogleの上位に表示させ、気付いたことや分かったことをまとめる。

検索エンジンの基礎知識

Crawlerと呼ばれるrobotが常にサイト内のリンクを伝って巡回しています。

Crawlerは、その情報をまとめてIndex(一時的に情報を保管する保管庫のようなもの)を作成します。

検索エンジンは、ユーザが検索したキーワードに対して、Algorithm(独自の方式)にしたがってIndexの情報の中から「関連性の高い」順に表示します。

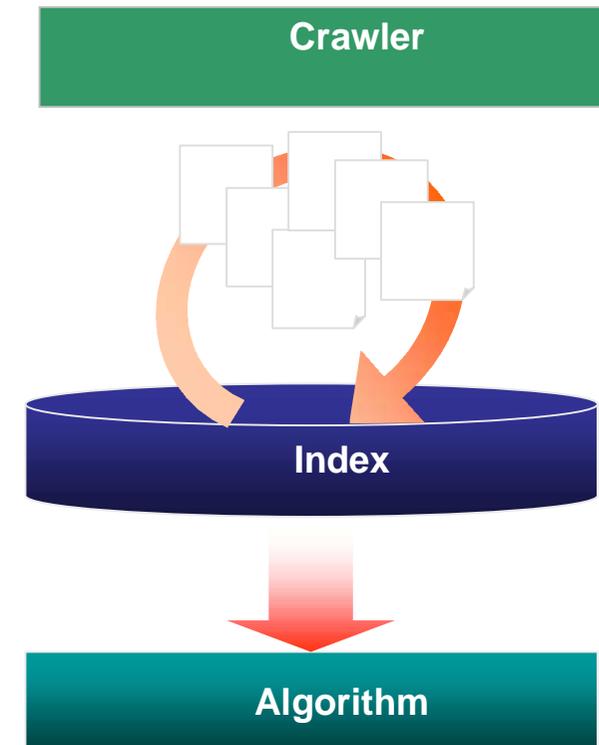


図1. 検索エンジンの概略

検索エンジンの検索順位はアクセス数順に表示されていない。

内部要因の内部ソース

```
1 .....<head>
2 .....<META http-equiv="Content-Type" content="text/html; charset=Shift_JIS">
3 .....<meta name="robots" content="all">
4 .....<title>タイトル</title>
5 .....<meta name="Keywords" content="キーワード 1 ,キーワード 2 ,キーワード 3 ,キー
6 .....ワード 4 ,キーワード 5 . . . ">
7 .....<meta name="Description" content="サイト説明">
8 .....</head>
```

The screenshot shows a Google search result for '神奈川大学' (Kanagawa University). The search bar contains '神奈川大学' and the search button is labeled '検索'. Below the search bar, there are radio buttons for 'ウェブ全体から検索' (selected) and '日本語のページを検索'. The search results show '神奈川大学の検索結果 約 569,000 件中 1 - 10 件目 (0.16 秒)'. The first result is '神奈川大学ホームページ' (3回閲覧 - 2月12日), which is circled in red. A blue box highlights the description: '神奈川大学の公式ホームページです。大学案内、教育、研究活動、入試情報、公開講座情報などをご覧いただけます。' and the URL 'www.kanagawa-u.ac.jp/~14k - トップページ - 検索ページ'. Below the description, there are links for 'Web Station', '入試情報', 'アクセスガイド', '学部・大学院', '企業・一般の方', '卒業生の方', '神奈川の方', and '大学概要'. At the bottom, there is a link 'kanagawa-u.ac.jp からの検索結果 >'. Arrows from the HTML code above point to the 'タイトル' and 'サイト説明' elements in the search result.

内部要因のキーワード含有率

例1)「検索エンジンにはディレクトリ型とロボット型があります。ロボット型はGoogleのようなもので、ディレクトリ型はYahooのようなものです。」

例2)「検索エンジンにはディレクトリ型検索エンジンとロボット型検索エンジンがあります。ロボット型検索エンジンはGoogleのようにロボットが自動で集めてきたデータの中から検索する検索エンジンです。ディレクトリ型検索エンジンはYahooのようにエディタによって手動で登録される検索エンジンです。」

検索エンジンはどのようにして有用なサイトと有用でないサイトを判別しているのか

例1) ♪

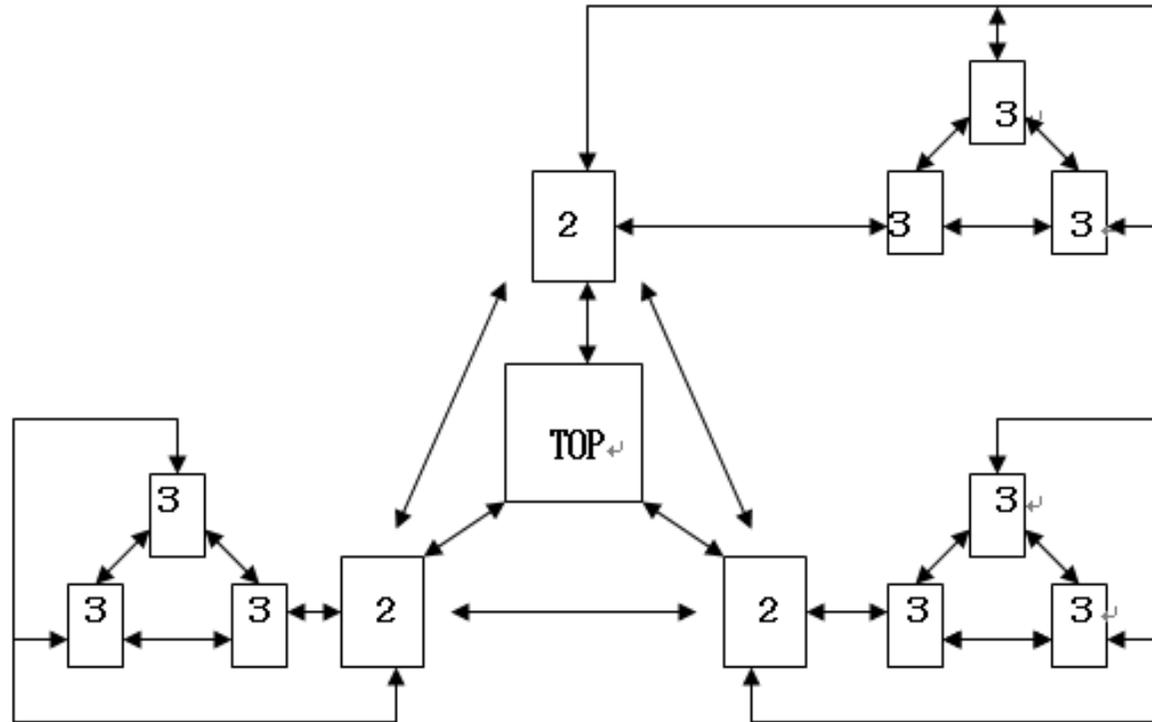
検索 | エンジン | に | は | ディレクトリ | 型 | と | ロボット | 型 | が | あります | ロボット | 型 | は | google | の | よう | な | もの | で | ディレクトリ | 型 | は | yahoo | の | よう | な | も | の | です。 ♪

♪

例2) ♪

検索 | エンジン | に | は | ディレクトリ | 型 | 検索 | エンジン | と | ロボット | 型 | 検索 | エンジン | が | あります | ロボット | 型 | 検索 | エンジン | は | google | の | よう | に | ロボット | が | 自動 | で | 集めて | きた | データ | の | 中 | から | 検索 | する | 検索 | エンジン | です | ディレクトリ | 型 | 検索 | エンジン | は | yahoo | の | よう | に | エディタ | によって | 手動 | で | 登録 | される | 検索 | エンジン | です ♪

内部要因の内部リンク



- 1 階層目 . . . <http://1.com> (TOP)
- 2 階層目 . . . <http://2.1.com> または、 <http://1.com/2/>
- 3 階層目 . . . <http://2.1.com/3/> または、 <http://1.com/2/3/>

内部要因のURL

Google 神奈川大学 検索 検索オプション
表示設定

ウェブ全体から検索 日本語のページを検索

ウェブ 神奈川大学 の検索結果 約 573,000 件中 1 - 10 件目 (0.22 秒)

他のキーワード: [神奈川大学 偏差値](#) [神奈川大学 図書館](#) [神奈川大学 サッカー部](#) [神奈川大学生協](#) [神奈川大学 硫化水素](#)

[神奈川大学ホームページ](#) - 2 回閲覧 - 15:46

神奈川大学の公式ホームページです。大学案内、教育、研究活動、入試情報、公開講座情報などをご覧ください。

www.kanagawa-u.ac.jp/ - 14k - [キャッシュ](#) - [関連ページ](#)

Web St@tion	企業・一般の方
入試情報	卒業生の方
アクセスガイド	神大生の方
学部・大学院	大学概要

[kanagawa-u.ac.jp からの検索結果](#) »

[神奈川大学ホームページ](#)

事務局のある1号館や入試センターやイベントの多い16号館へは「**神奈川大学入口**」で、またシンポジウムなどが多く行われる20・23各号館や**神奈川大学生協**へ行く場合は「六角橋西町」で下車すると便利です。...

www.kanagawa-u.ac.jp/02/accessmap/ - 17k - [キャッシュ](#) - [関連ページ](#)

[神奈川大学 - Wikipedia](#)

神奈川大学(かながわたいがく、英語: Kanagawa University)は、神奈川県横浜市神奈川区六角橋3-27-11に本部を置く日本の私立大学である。1949年に設置された。大学の略称は**神大**(じんたい)を大学が公式に使用しており、登録商標(第4372633号および ...

ja.wikipedia.org/wiki/神奈川大学 91k - [キャッシュ](#) - [関連ページ](#)

外部要因の歴史

新しいドメイン (Url) を取得してすぐにGoogleやYahooにサイト登録しても、一定期間は上位に掲載されない現象をエイジングフィルタというように、ドメインには歴史が関係している。
これは単純に、古いドメインほど上位に表示されやすいことを意味している。

外部要因の更新頻度

サイトは「更新頻度が高いほど上位表示される」といわれている。
具体的には、文章やデザインやページ数など、コンテンツの追加の更新をしていくうちに有用なサイトだと判別される可能性は上がる。

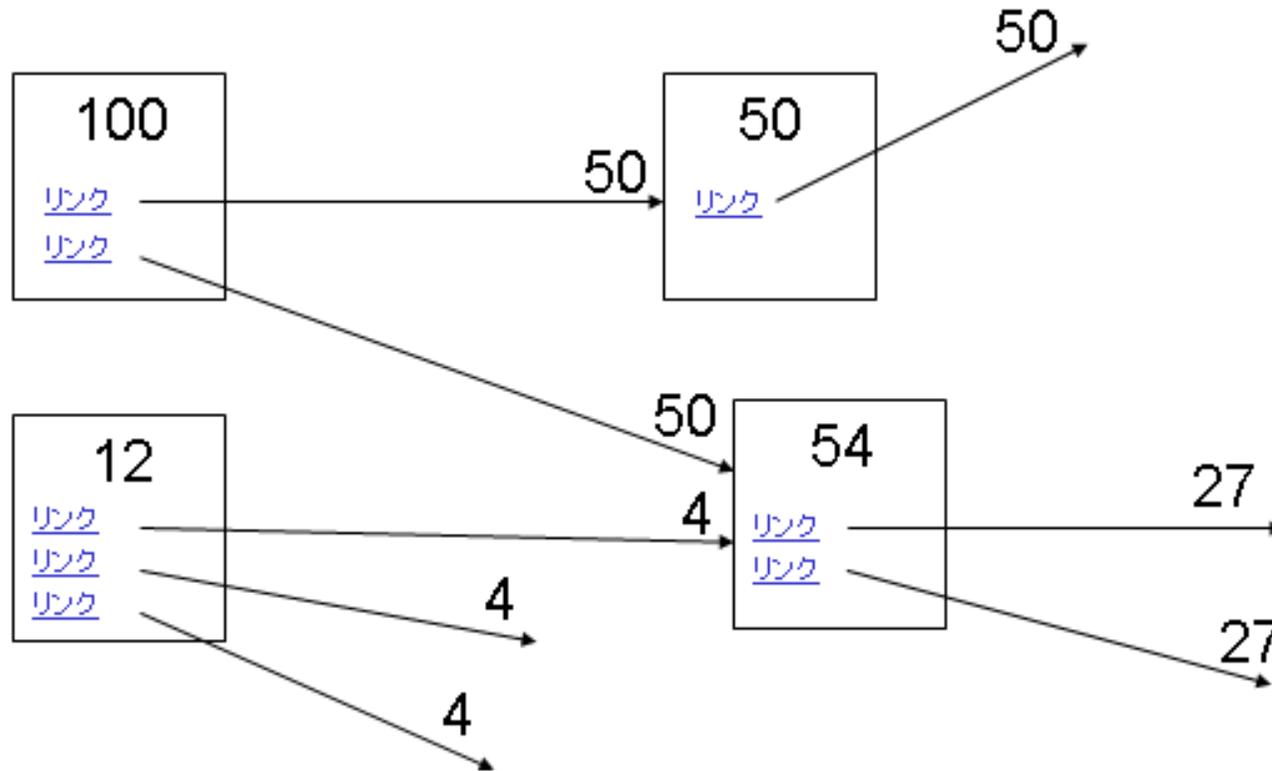
外部要因のAnchor Text

Anchor Textとは、リンクが設定されている部分に書かれた文字のことである。

YahooやGoogleで「出口」と検索すると、YahooやGoogle自身が1位に表示されるが、これは、このAnchor Textが原因で起こる現象である。



外部要因の外部リンク

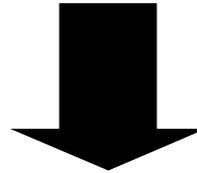


Googleの基本概念

「多くの良質なページからリンクされているページは、やはり良質なページである」

実験と考察

あるキーワードで上位表示を狙ってサイトを作成(2008/5/21)した結果、2008/1/25時点で、「Google 9位、Yahoo 33位、Googleモバイル1位、Yahooモバイル2位」に表示された。



本実験では、内部要因対策は十分に行えたが、外部要因対策の内、歴史などは期間的に短く不十分であった。このことから、PCの検索エンジンは外部要因を重視するのに対し、モバイルの検索エンジンは内部要因を重視することが順位から推測できる。

GoogleとYahooで順位の違いが生じたのも、Googleは比較的内部要因を重視するのに対し、Yahooは比較的外部要因を重視することが順位から推測できる。

まとめ

検索エンジンは内部要因と外部要因の2つの要因から判別されている。今後検索エンジンは、

- ・個人によって異なる検索結果を表示するパーソナライズ検索。
- ・検索した場所によって異なる検索結果を表示するGPS連動検索。
- ・動画や地図や画像などさまざまなコンテンツを表示するユニバーサル検索。

なども用いて有用なサイトかどうかを判別していくと予想されているのでそれらについても今後検討していく必要がある。