

1. このテーマ選定理由

今、インターネットの急激な普及により多くの情報をネットから得られるようになりました。しかしながら、それにより本当に必要な情報が掲載されているホームページを探し出すのが難しくなってきました。そのため検索エンジンの重要性が高まっています。なので、検索エンジンがどのようなプログラムで作られ、どのようにして動作しているかを理解し、小規模な検索エンジンを動かしてみたいと言う理由でこのテーマを選択しました。

2. 研究概要

今回は、これまでの検索エンジンに関することについて調べたものを簡単にまとめました。この研究の目的は検索エンジンの仕組みを理解し、自分のよくみるホームページや検索の対象となるホームページを自由に設定できるような自分専用の検索エンジンを自分のパソコン上で実際に運用してみることです。

現在、インターネット上では、さまざまな検索エンジンサービスが存在し、競い合っています。検索エンジンには大きく分けて2つの種類があります。1つはロボット型の検索エンジンで、この検索エンジンで最も有名なものは「Google」です。Googleはウェブロボットと呼ばれるプログラムを使ってホームページを巡回してインデックスを作成し、そのデータを元に検索を行います。ただしロボット検索エンジンの作業はすべて機械的に行われているため、利用者にとって有用な情報を含んだホームページとそうではないページが一緒に検索結果に出てしまいます。そしてもう1つはディレクトリ型検索エンジンです。ディレクトリ型検索エンジンは人によって、登録に値するかどうかの審査を受けています。ディレクトリ型の検索エンジンでもっとも有名なのは「Yahoo」です。ただし、人がチェックを行うので、登録されている件数はロボット型の検索エンジンに比べると圧倒的に少なくまた登録作業をしていないホームページは検索対象にならないという問題点があります。しかし、最近は2つの異なる検索エンジンの仕組みを組み合わせることで質の高い検索機能を提供しようとする混在型（ハイブリット型）が多い。

このため、定期的に更新されていて有利な情報が記載されたホームページのリンクを持っているホームページを検索対象として設定し、ロボット型の検索エンジンのように自動的に対象のホームページにアクセスし、データを回収することにより少しでも質の高い検索結果が得られるのではないかと考えています。

3．研究計画

表 1 9 月までの研究計画

時期		研究計画
8 月	下旬	開発環境の準備、特にウェブサーバのインストールや設定方法を調べる
9 月	上旬	参考文献 2 に基づいて小規模な検索エンジンを動かし、操作の確認を行う。牽引用テーブルの構造を理解する
	下旬	オブジェクト指向 Ruby 言語の基礎を学ぶ

4．その他

参考文献

- 1． 神崎 洋治 “体系的に学ぶ検索エンジンのしくみ”, 2008
- 2． “Ruby で作る検索エンジン”, 2009
- 3． 水野 貴明 “Web 検索エンジン Google の謎”, 2004
- 4． 用語解説辞典