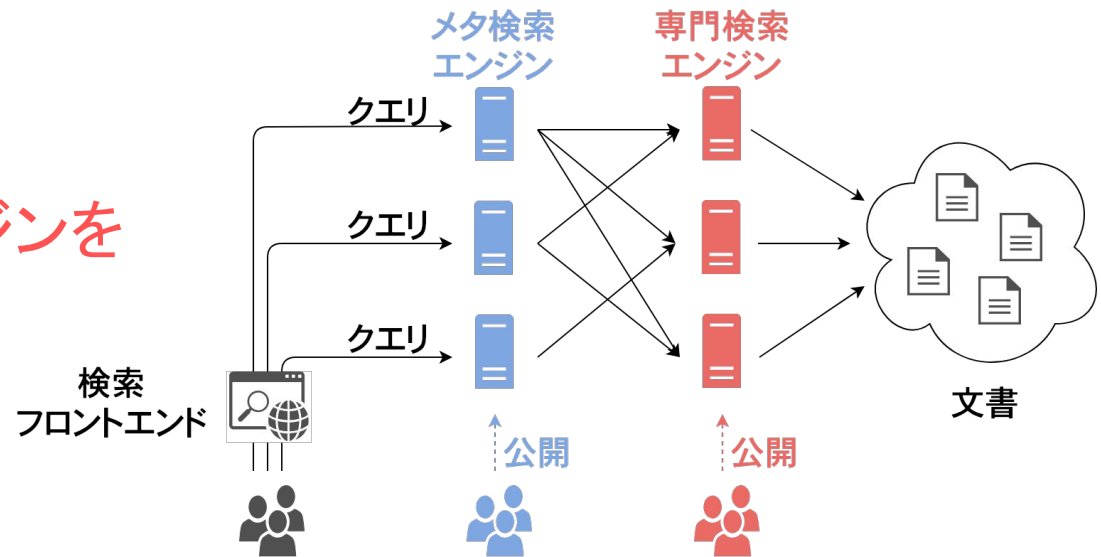


分野限定型検索エンジンを複数組み合わせた 分散型検索エンジン —自由な検索エンジンの開発—



kearch

分野を限定した検索エンジンを
複数つなぎ合わせて
自由な検索エンジン
を開発した。



世界中で利用されている検索エンジンのほとんどが不透明なものであるという現状を
打開すべく、省リソースで誰でも立ち上げ可能で、かつランキングアルゴリズムが公開で
ある、自由な検索エンジン「kearch(ケアチ)」を開発した。

アプローチ

… 2種類の検索エンジンをつなぎ合わせる

- ❖ **専門検索エンジン**: ある特定の分野に限定して、Web上の文書のクローリングを行い、検索機能を提供する。
- ❖ **メタ検索エンジン**: 複数の専門検索エンジンをつなぎ合わせることで、より広範囲の分野をカバーした検索機能を提供する。

従来の検索エンジンとの違い

- ❖ 省リソースで動作可能
既存の検索エンジンに比べて非常に少ないリソースでも動作する。
- ❖ 公開されたアルゴリズム
関連するソースコードはGPLライセンスにてGitHub上で公開されているため、誰でも閲覧することが可能。
<https://github.com/kearch>
- ❖ 段階的に構成可能
分野を限定した検索エンジンを段階的につないでいくことで、小規模に運用を開始できる。

検索例

… プログラミング言語 Haskell に関する検索エンジンを立ち上げた場合

Haskellのパッケージマネージャであるstackについて検索してみる。

実際のクエリは“stack tutorial”である。



kearch

stack tutorial

Specialist server name

select automatically

Haskellのstackに関連する文書が最上位にヒットする

stack/GUIDE.md at master · commercialhaskell/stack · GitHub

<https://github.com/commercialhaskell/stack/blob/master/doc/GUIDE.md>

Skip to content Why GitHub? Features → Code review Project management

1.9478343 from Haskell (163.43.31.234)

User guide - The Haskell Tool Stack

<https://docs.haskellstack.org/en/stable/GUIDE/>

The Haskell Tool Stack Home Changelog Tool documentation Install/upgrades

Hello World Example Inner Workings

1.9478343 from Haskell (163.43.31.234)

⋮

どの専門検索エンジンからの結果であるかが表示される