

文書作成に役立つ 検索のTips

横野 光
明星大学

文書作成と検索

- 「書きたい内容はあるがどう書けばいいか……」
 - 文章ガイドブック的なものはあるが一般的なものがメイン
 - 自分が書きたい事柄についての作例がないこともある
 - 特に専門的，個別的な内容だとカバーされないことが多い
- 検索エンジンを使って似たような事例を探して比較してみる
- 一般的な単語の使い方か
 - 他にも適切な単語があるか

検索エンジンの使い方 (1/2)

- 書こうとしている文をそのまま入れてみる
 - 「””」(ダブルクォーテーション)で囲むとキーワードが分割されない
 - 全く同じ文を使っているページが表れる
- けど……
 - 検索キーワードが長すぎると全くヒットしないことがある
 - いくつかのまとまりに分割してみる
 - フレーズ単位とか
 - 具体的な表現は消してみる
 - 用言(動詞や形容詞)の場合, 途中までにしてみる

検索エンジンの使い方 (2/2)

- 検索結果の見方
 - まずはヒット件数を見る
 - ヒット件数が多い⇨一般的によく使われている
 - 検索結果を見る
 - 辞書サイトのページや文例集のようなページが多く出現していると、一般的にはあまり使われていない可能性がある
 - 各検索結果のスニペットでのキーワードの使われ方を見る
 - 検索キーワードをすべて含んでいないかもしれない

約 87,600 件 (0.32 秒)

<https://cancam.jp> > archives ▾

「職場で苦手な人」へのストレスから解放される方法 - CanCam

2020/01/09 — 「職場で苦手な人」へのストレスから解放される方法. 何かとストレスフルな「仕事」のシーン。中でも難しいのが「職場の人間関係」ですよね。

スニペットをもう少し見てみる

- 検索キーワードが含まれているスニペット
 - 自分が想定している文脈でキーワードが使われているか
- 検索キーワードが含まれていないスニペット
 - 自分が想定している文脈に近いものであれば、そこで出てくる単語を使ってみることも考えられる

単語の使い方 (1/2)

- 文脈に応じて適切な単語を選択したい
 - 意味としては似ている語であっても状況に応じて使い分けが必要な場合がある
 - 国語辞典の語義だけでは判断がつかない場合もある
 - 間違えやすい単語
 - 類義語
 - 同音異義語
 - 確立と確率
 - 誤字
 - 人工知能と人口知能
- 同じ内容をどう言うか
 - サ変名詞+する (e.g. 入力する)
 - 「入力してください」, 「ご入力ください」, 「入力をしてください」……

仮名漢字変換を使っているととりあえず最初のやつを選びがち……

単語の使い方 (2/2)

- 複数ある候補のそれぞれで検索
 - 単語だけでなくフレーズで検索し
 - 例: 「[ストレスから解放される](#)」
- 検索結果の数を比較する
 - 「[ストレスから解放される](#)」 611,000件
 - 「[ストレスから開放される](#)」 424,000件
 - 検索エンジンによっては「もしかして……」でより適切な検索キーワードを提示してくれる場合もある



使おうとしている言葉は一般的か (1/2)

- 専門用語，業界用語をうっかり使ってしまうことがある
 - 自分はよく知っているが，一般的にはあまり知られていない言葉
 - 例: 機械色，商材
- 読んでいる人が分かる言葉で説明する必要がある
 - けど，何が専門用語か，というのは判断が難しい
 - 用語辞典や用語集は存在するが，自社での用語までは当然フォローされない

使おうとしている言葉は一般的か (2/2)

- 検索エンジンでのヒット件数を見る
 - 件数が少ないと一般的ではないかも、と考えられる
 - どのくらいの数が「少ない」と言えるか
 - 感覚的には数千は少ない, 数万だとちょっと微妙(個人的な意見)
 - 元の文面中の別の単語との組み合わせで検索
- 検索結果に出てくるページも見てみる
 - 検索結果に「<検索キーワード>とは」みたいなサイトやスニペットがあるとヒット件数にかかわらず専門用語っぽいかも
 - その語についての説明が求められることが多い, と見ることができる

しっくりくる単語を探す

- 書きたい文はあるけど、なんとなくしっくりこない
 - もう少ししっくりくる単語があるかもしれない
- 元の文から単語を一つ抜いて検索する
 - 抜いた部分に“*” (アスタリスク) を入れて検索してみる
 - 「記事を*できる」で検索すると記事を投稿できる、記事を保存できる、などが検索結果に表れる
 - 関係ない単語を含めたくないときは“-” (マイナス) をその単語の前に入れる
 - 「記事を*できる -保存 -投稿」で検索すると、記事を読覧できる、記事を理解できる、などが出てくる

(発展)日本語を計算機で処理する

- 文を解析してその結果を見してみる
 - 解析結果がおかしいと、人にとっても解釈が難しい文の可能性
 - 注: 解析器はそれなりに高精度だが必ず正しい結果を出すというわけではない
- 単語や単語列の統計を取ることができる
 - 出現頻度などを求められる
- 日本語の解析器
 - 形態素解析器
 - MeCab (<https://taku910.github.io/mecab/>)
 - Juman (<https://nlp.ist.i.kyoto-u.ac.jp/?JUMAN>)
 - 係り受け解析器
 - CaboCha (<https://taku910.github.io/cabocha/>)
 - KNP (<https://nlp.ist.i.kyoto-u.ac.jp/?KNP>)

導入が少々面倒

(Linuxやプログラミングの知識があると良い)

```
太郎は花子の読んだ本を借りた
太郎は-----D
花子の-D |
読んだ-D |
本を-D
借りた

EOS
* 0 4D 0/1 -2.501312
太郎 名詞,固有名詞,人名,名,*,*,太郎,タロウ,タロー
は 助詞,係助詞,*,*,*,は,ハ,ワ
* 1 2D 0/1 0.352664
花子 名詞,固有名詞,人名,名,*,*,花子,ハナコ,ハナコ
の 助詞,連体化,*,*,*,の,ノ,ノ
* 2 3D 0/1 1.780279
読ん 動詞,自立,*,*,五段・マ行,連用タ接続,読む,ヨン,ヨン
だ 助動詞,*,*,*,特殊・タ,基本形,だ,ダ,ダ
* 3 4D 0/1 -2.501312
本 名詞,一般,*,*,*,本,ホン,ホン
を 助詞,格助詞,一般,*,*,*,を,ヲ,ヲ
* 4 -1D 0/1 0.000000
借り 動詞,自立,*,*,一段,連用形,借りる,カリ,カリ
た 助動詞,*,*,*,特殊・タ,基本形,た,タ,タ
EOS
```

おわりに

- 検索エンジンを使って文例を探す
 - 文の妥当性を調べることができる
 - 代替案を見つけることができる
 - 同じような文脈で使われている他の単語が見つかるかも
- 注意
 - ヒット件数が多いからといって正しいとは限らない
 - ヒットした検索結果のページでどのように使われているかを確認する