

tagger

<https://github.com/trinker/tagger>

- ・ openNLP の品詞タグ付与

必要なパッケージをインストール

タグ一覧

`penn_tags()`

Penn Tree Bank 式のタグではなく、一般的な品詞記号にまとめることもできる：
`as_universal()`

品詞を明記することもできる：`as_basic()`

コマンド

- ・ NICER より NS502 のデータの本文部分だけを取り出したものを例に

タグ付与：`tag_pos()`

タグ頻度 `count_tags()`

タグ頻度のプロット：`plot()`

タグ付きのテキストの出力：`as_word_tag()`

品詞付きデータで、品詞を指定して検索

- ・ 形容詞 + 名詞 を例に
- ・ `grep` だと、該当する文字列がある一文全体が表示される

stringr パッケージの利用

stringr のインストール

stringr パッケージの機能を使って該当文字列だけの検索

<https://sugiura-ken.org/wiki/wiki.cgi/exp?page=grepExtract>

- ・ これを修正