

Source code for the LitsearchR keyword generation

```
library(litsearchr)
curdir <- dirname(rstudioapi::getSourceEditorContext()$path)
setwd(curdir)

PP_import <- import_results(
  directory = "./initial_selection/"
)
PP_results <- remove_duplicates(
  PP_import, field = "title", method = "exact")

raked_words <- extract_terms(
  text = paste(PP_results$title, PP_results$abstract),
  method = "fakerake",
  min_freq = 2,
  ngrams = TRUE,
  min_n = 2,
  max_n = 5
)
real_words <- extract_terms(
  text = NULL,
  keywords = PP_results$keywords,
  ngrams = TRUE,
  min_n = 2,
  method = "tagged"
)

all_keywords <- unique(append(real_words, raked_words))

PP_dfm <-
  create_dfm(elements = paste(PP_results$title, PP_results$abstract),
            features = all_keywords
  )
groupingGraph <-
  create_network(
    search_dfm = PP_dfm,
    min_studies = 3,
    min_occ = 3
  )

cutoff <-
  find_cutoff(
    groupingGraph,
    method = "cumulative",
    percent = .80,
    imp_method = "strength"
  )

reducedgraph <-
  reduce_graph(groupingGraph, cutoff_strength = cutoff[1])
```

```

searchterms <- get_keywords(reducedgraph)
write.csv(searchterms, "./search_terms.csv")

grouped_terms <- read.csv("./ngrp_search_terms.csv")
plant_terms <- as.character(grouped_terms$word[grepl("plant", grouped_terms$group)])
pest_terms <- as.character(grouped_terms$word[grepl("pest", grouped_terms$group)])
method_terms <- as.character(grouped_terms$word[grepl("method", grouped_terms$group)])
danger_terms <- as.character(grouped_terms$word[grepl("danger", grouped_terms$group)])
goal_terms <- as.character(grouped_terms$word[grepl("goal", grouped_terms$group)])
chemical_terms <- as.character(grouped_terms$word[grepl("chemical", grouped_terms$group)])
farming_terms <- as.character(grouped_terms$word[grepl("farm", grouped_terms$group)])

plants <- unique(append(c("maize", "desmodium", "sorghum"), plant_terms))
pests <- unique(append(c("stemborer", "striga", "grasshopp"), pest_terms))
farms <- c("intercrop", "agricultu", "farm", "cereal")
farming <- unique(append(c("fix* nitrogen"), farming_terms))
method <- unique(append(c("behavio*ral response", "push pull", "push*pull", "semiochemi",
"stimulo deterrent"), method_terms))
goals <- unique(append(c("pest control"), goal_terms))
dangers <- unique(append(c("yield loss", "crop loss", "famine", "plant damage"), danger_terms))
dango <- unique(append(c(), dangers))
dango <- unique(append(dango, goals))

mysearchterms <- list(
  farms, farming, pests, dango, method, plants
)
my_search <-
write_search(
  groupdata = mysearchterms,
  languages = "English",
  stemming = TRUE,
  closure = "none",
  exactphrase = TRUE,
  writesearch = FALSE,
  verbose = TRUE
)
write(my_search, "./search.md")

musthave <- as.character(read.csv("./must_have.csv", header = FALSE)$V1)
q4_import <- import_results(
  directory = "./NewApproach/query4/"
)
q1_import <- import_results(
  directory = "./NewApproach/query1/"
)
#extract_import <- rbind(q2_import, q1_import)
q4_results <- remove_duplicates(
  q4_import, field = "title", method = "exact")
q1_results <- remove_duplicates(
  q1_import, field = "title", method = "exact")
articles_found <- check_recall(true_hits = musthave,

```

retrieved = q4_results\$title)

articles_found