Wybrane Zagadnienia Współczesnej Informatyki

Tytuł projektu:

Aplikacja webowa do akwizycji i analizy danych z kanału RSS portalu <u>gofin.pl</u> (prawnik radzi).

Link do repozytorium:

https://github.com/piotranon/Shiny

Wykonał:

Piotr Długosz, 3 rok Informatyki, Inżynierskie, grupa lab.1

Spis treści:

- 1. Setup / pierwsze uruchomienie.
- 2. Opis projektu / funkcjonalności.
- 3. Lista wykorzystanych bibliotek.
- 4. Podgląd aplikacji.

Setup / pierwsze uruchomienie

- 1. Instalacja SOLR.
- 2. Utworzenie rdzenia w SOLR (./solr create -c projekt).
- 3. Zmiana portu w pliku (solr.R linia 8, zgodnie z dokumentacja port domyślny to 8983).
- 4. Uruchomienie pliku test.R jako Aplikacji przez R Studio.

Opis projektu / funkcjonalności

Dane na każdej zakładce są generowane automatycznie. Należy chwilę poczekać aby dane się załadowały.

Projekt posiada 3 pliki:

- test.R Główny plik aplikacji
- solr.R funkcje dotyczące biblioteki Solr, odczytu zapisu
- functions.R funkcje dotyczące analizy danych

Lista funkcjonalności

- 1. Projekt pozwala na pobieranie danych z kanału RSS i zapis danych w bazie <u>SOLR</u>.
 - W zakładce "dane" mamy możliwość pobrania najnowszych danych, podglądu danych z bazy. Dane zapisywane w bazie są unikalne.
- 2. Analizę częstości występowania słów
 - W zakładce "czestość wystepowania", mamy możliwość generowania wykresu czestości wystepowania słów.
- 3. Wizualizacji Chmury słów
 - W zakładce "chmura slow" mamy możliwość generowania chmury z danych zapisanych w bazie.
- 4. Klasteryzacje danych.
 - W zakładce "klasteryzacja" mamy możliwość klasteryzacji danych.

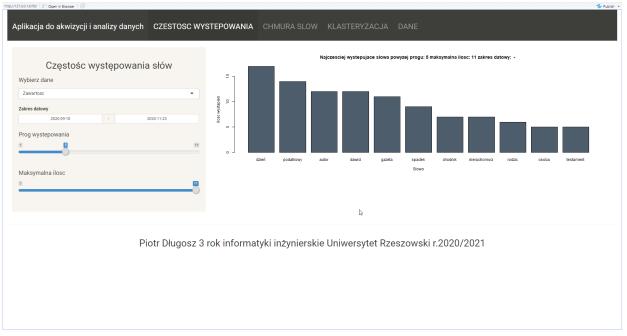
Lista wykorzystanych bibliotek z funkcjonalnościami

1. Tm Czyszczenie tekstu. 2. Shiny Biblioteka do aplikacji webowej. 3. Shinythemes Motyw bootstrap do aplikacji webowej. 4. Wordcloud2 Generowanie chmury słów 5. DT Wyświetlanie danych w tableach 6. tidyRSS Pobieranie danych z systemu RSS. 7. Solrium Połączenie z silnikiem Solr. 8. Gsubfin Do wyciagania zawartości z tekstu. 9. rList Do manipulacji danych. 10. stringi Do modyfikacji tekstu. 11. httr Do zapytań http. 12. XML Do analizy plików xml. 13. cluster Do klasteryzacji

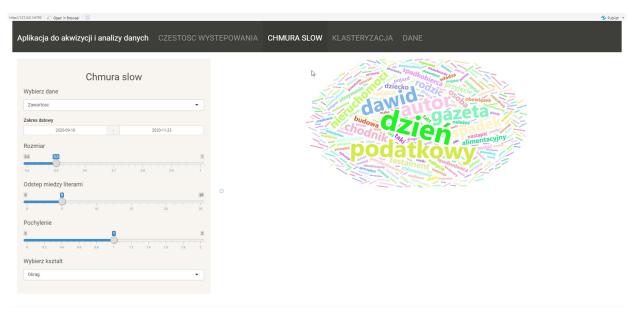
Wybrane Zagadnienia Współczesnej Informatyki – Aplikacja webowa do akwizycji i analizy danych z kanału RSS portalu <u>gofin.pl</u> (prawnik radzi). Długosz Piotr – 3 rok Informatyki inżynierskie – grupa lab 1

Podgląd aplikacji

Podgląd zakładki "czestosc wystepowania"

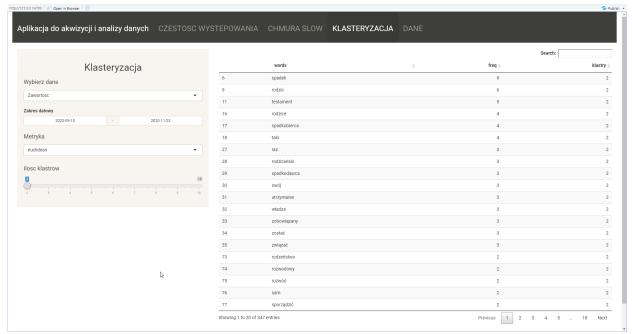


• Podgląd zakładki "chmura slow"



Piotr Długosz 3 rok informatyki inżynierskie Uniwersytet Rzeszowski r.2020/2021

Wybrane Zagadnienia Współczesnej Informatyki – Aplikacja webowa do akwizycji i analizy danych z kanału RSS portalu <u>gofin.pl</u> (prawnik radzi). Długosz Piotr – 3 rok Informatyki inżynierskie – grupa lab 1 • Podgląd zakładki "klasteryzacja"



Podgląd zakładki "dane"

