

TWD projekt 3

Raport Weekly 1

Skład zespołu: Adrian Stańdo, Patryk Tomaszewski, Kinga Ułasik

Główny cel:

Wykorzystać dane pobrane z [Google Takeout](#) oraz [YouTube Data API](#) aby przedstawić informacje o historii wyświetlanych filmów na YouTube.

Dostępne dane:

- Czas wyświetlenia filmu przez użytkownika
- Ilość wyświetleń filmu (łącznie, nie tylko danego użytkownika)
- Ilość lików i dislików
- Ilość komentarzy
- Długość trwania
- Kategoria wybrana przez twórcę filmu
- Język filmu (te dane mogą być niedokładne)
- Data wrzucenia filmu
- Automatycznie wygenerowane "tematy" filmu

Zarys systemu:

Całość zostanie podzielona na dwa główne moduły - moduł pre-procesujący dane i moduł odpowiadający za ich wyświetlenie jako dashboard. Pierwszy z modułów będzie pobierał plik html z historią oglądania filmów na YT bezpośrednio z pobranego [Google Takeout](#), konwertował w łatwo odczytywalny plik csv a następnie uzupełniał o dodatkowe dane na podstawie [YouTube Data API](#). Drugi z modułów będzie operować na wygenerowanym pliku i wyświetlać go w postaci interaktywnego dashboarda.

Główna idea dashboardu:

3-4 wykresy, u góry panel kontrolny, na którym się wybiera statystykę do wyświetlenia oraz filtry (np. z jakiego okresu czasu wyświetlać dane, maksymalny próg czasu trwania filmiku, minimalny próg ilości polubień).

Pomysły na wykresy:

- Heatmapa według dnia tygodnia i godziny - intensywność koloru zależy od wybranej statystyki np. ilość czasu spędzonego na YT czy ilość danych polubień
- Wykres liniowy przedstawiający zmianę danej statystyki w czasie
- Porównanie dwóch różnych statystyk między sobą za pomocą scatter plota
- Przedstawienie średniej z danej statystyki przedstawione za pomocą histogramu lub pojedynczej liczby, na przykład ile średnio polubień miały oglądane filmiki, ile czasu trwały
- Przedstawienie najczęściej oglądanych filmików
- Wykrywanie klasterów filmów o podobnych cechach

