

Závěrečná zpráva k semestrální práci  
YouTube reranking

Petr Kubín  
Herbert Waage  
FIT ČVUT

8. prosince 2015

## Obsah

<b>1</b>	<b>Popis projektu</b>	<b>3</b>
1.1	Zadání práce . . . . .	3
1.2	Vstup . . . . .	3
1.3	Výstup . . . . .	3
<b>2</b>	<b>Způsob řešení</b>	<b>3</b>
2.1	Komunikace s Youtube API . . . . .	3
2.2	Identifikace záznamů odpovídajících klíčovému slovu . . . . .	3
2.3	Získání metadat k (identifikovaným) záznamům . . . . .	3
2.4	Výpočty různých typů vzdáleností . . . . .	4
2.5	Třídění podle daných atributů . . . . .	4
2.6	Webový interface . . . . .	4
<b>3</b>	<b>Příklad výstupu</b>	<b>4</b>
<b>4</b>	<b>Experimentální sekce</b>	<b>5</b>
<b>5</b>	<b>Diskuze</b>	<b>5</b>
<b>6</b>	<b>Závěr</b>	<b>5</b>

# 1 Popis projektu

## 1.1 Zadání práce

Cílem projektu je vytvoření aplikace umožňující vylepšené vyhledávání ve fotografiích uložených na serveru Youtube. Youtube poskytuje aplikační rozhraní umožňující získat prakticky libovolná data o videích, která jsou přístupna i z oficiálního webového rozhraní. Cílem projektu je tedy aplikace, která umožní vyhledávání na Youtube založené na klíčových slovech, stejně jako je tomu na Youtube nyní, ale navíc bude možné zadat sekundární vyhledávání na libovolná metadata.

## 1.2 Vstup

Klíčové slovo, sada hodnot metadat pro setřídění a počet výsledků (omezení velikosti výstupu).

## 1.3 Výstup

Videa, jejichž popis odpovídá klíčovému slovu. V levém sloupci jsou videa v původním nepřerankovaném pořadí, v pravém sloupci jsou videa setříděna podle vzdálenosti ke zvoleným metadatům.

# 2 Způsob řešení

## 2.1 Komunikace s Youtube API

Začátek komunikace s Youtube API probíhá zadáním dotahu do formuláře, odkud nám Youtube vrátí veškerá videa odpovídající hledanému výrazu s omezením maximálního počtu videí. Zde vezmeme každé video a pomocí parseru zjistíme identifikátor videa, což je řetězec za watch?v=, dlouhý max 12 znaků a omezený buď koncem řetězce nebo znakem &. Po získání identifikátoru videa jsme přešli ke stažení metadat.

## 2.2 Identifikace záznamů odpovídajících klíčovému slovu

## 2.3 Získání metadat k (identifikovaným) záznamům

Pro získání jsme použili curl, neboť standartní HTTP požadavek není na školním serveru webdev, kde je naše aplikace umístěna, povolen. Pro začátek komunikace je potřeba mít vlastní API klíč, který umožní identifikovat náš program oproti YouTube API a tím dokončit požadavek. Pro samotné získání metadat jsme využili [https://www.googleapis.com/youtube/v3/videos?id=".\\$id."&key=".\\$API\\_KEY."&part=snippet,statistics,contentDetails,recordingDetails](https://www.googleapis.com/youtube/v3/videos?id=). Kde proměnná \$id je identifikátor videa získaný rozparsováním url adresy a API\_KEY je již zmiňovaný unikátní klíč.

## 2.4 Výpočty různých typů vzdáleností

V naší aplikaci se vyskytuje několik datových typů: řetězec, číslo, GPS, datun. Pro každý datový typ jsme museli použít správný výpočet vzdálenosti. Pro číselné datové typy jsme použili euklidovskou vzdálenost. GPSky jsme porovnali pomocí Great circle distance, která přesně určí nejmenší vzdálenost na naší zeměkouli.

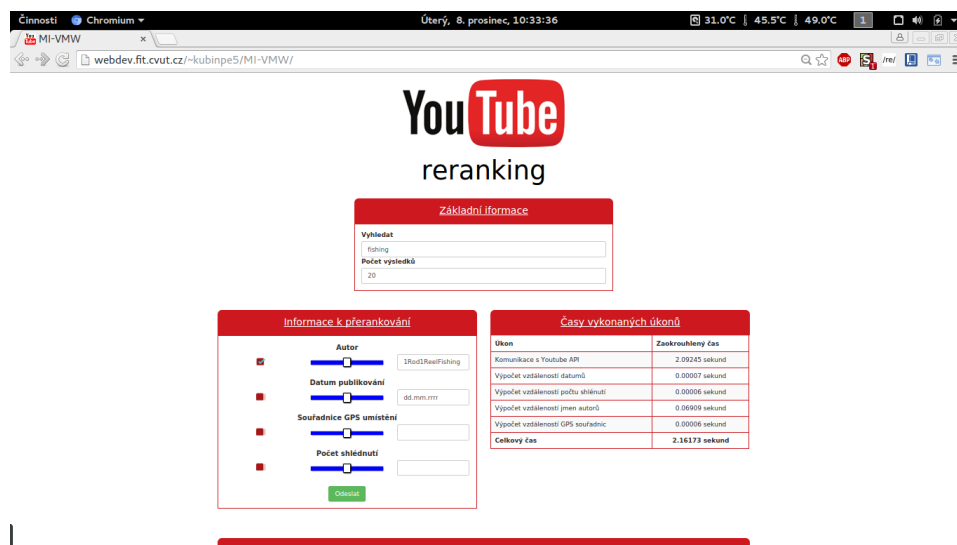
Pro typ řetězec, který je použit u jména Autora videa, jsme naprogramovali Levenshteinovu(editační) vzdálenost, která reprezentuje počet operací nutných k převedení jednoho řetězce na druhý.

## 2.5 Třídění podle daných atributů

// TODO

## 2.6 Webový interface

Naše aplikace je nasazena na serveru webdev. Pro styly jsme použili framework bootstrap, který umožňuje efektivní formátování webových elementů.

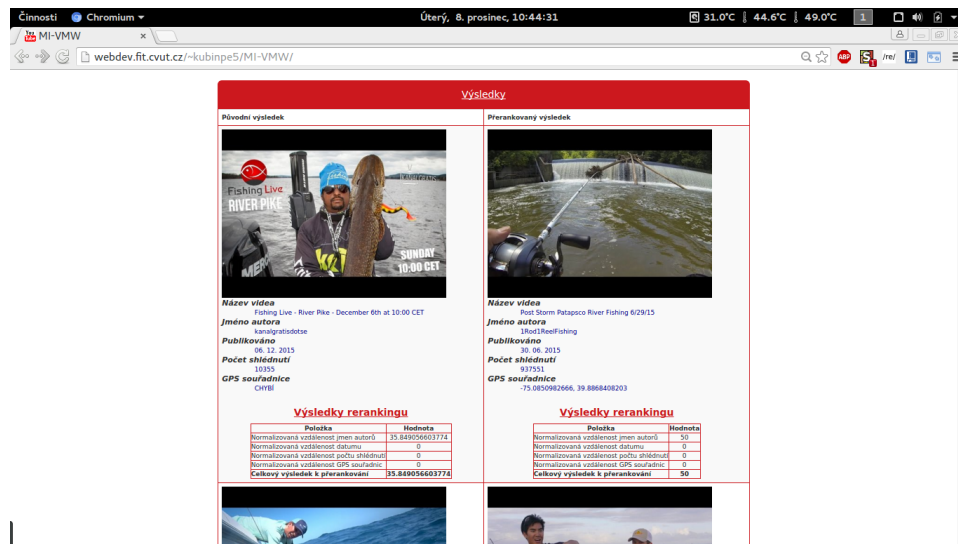


Obrázek 1: Náhled aplikace

## 3 Příklad výstupu

Výstup se skládá ze dvou sloupců. Levý sloupec reprezentuje staré uspořádání dat a pravý sloupec nové pořadí dle uživatelského vstupu v sekci informace

k přerankování. Celkový počet videí je omezen znovu uživatelským vstupem. V následující kapitole bude zmíněno testování rychlosti.



Obrázek 2: Náhled výsledku

## 4 Experimentální sekce

//TODO

## 5 Diskuze

## 6 Závěr

## Citace

- [1] Getting Started with the YouTube Data API: *Google Developers* [online]. [cit. 13. 11. 2014]. Dostupné z WWW: <<https://developers.google.com/youtube/v3/getting-started>>.