

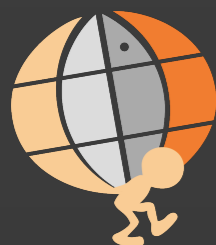
Prepájanie multimediálnych metadát

DP I.

Bc. Peter Gašpar

Vedúci: Ing. Jakub Šimko, PhD.

gasparpeter92@gmail.com



PeWe@FIIT
personalized web group

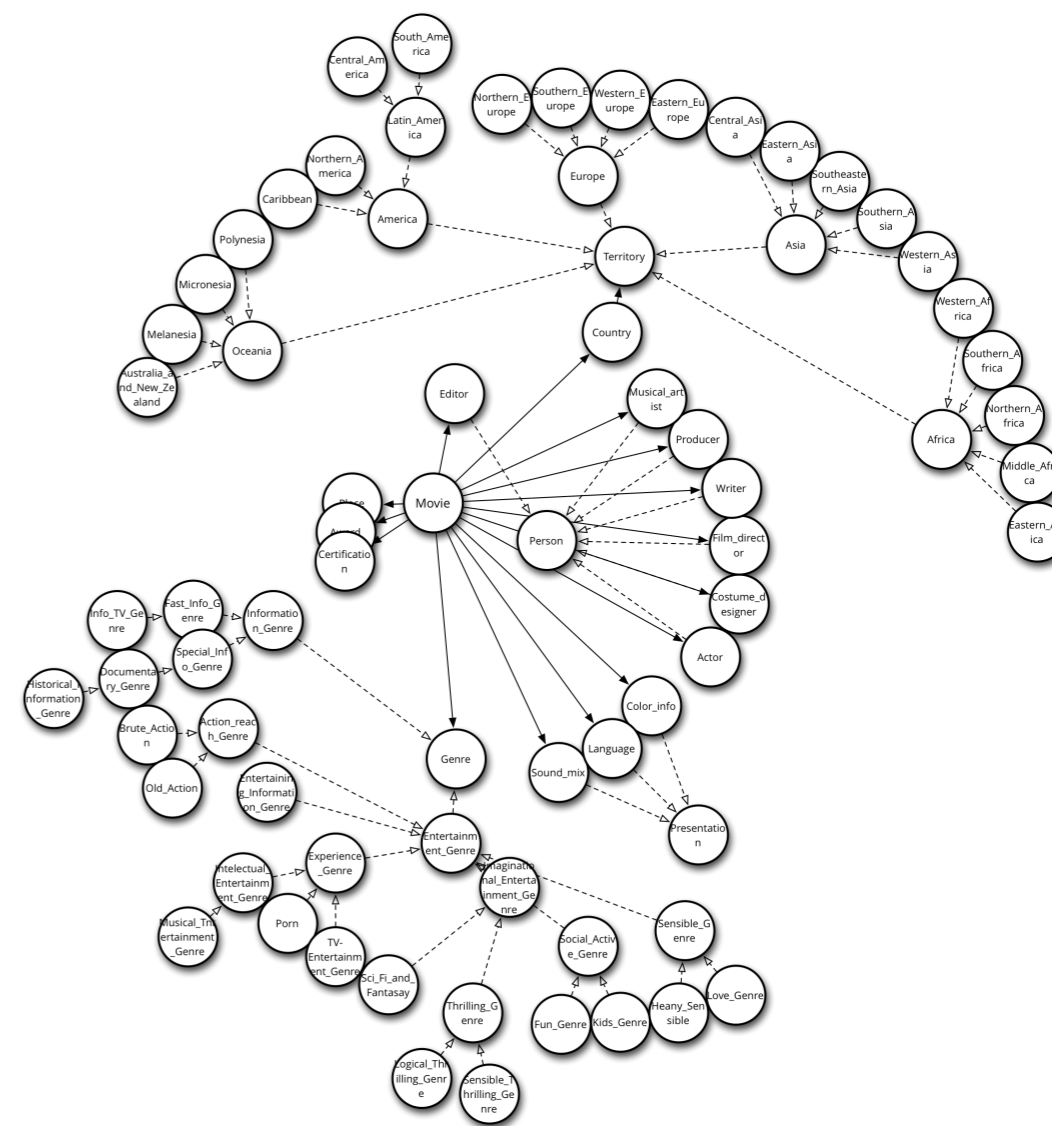
28.4.2015

- obohatenie existujúcich metód na **prepájanie multimedialných dát**
- získanie kvalitných a užitočných metadát
- zameranie na doménu **televízie**
 - metadáta k televíznemu obsahu (od relácií až po hercov)
- s cieľom využiť **niečo**, čo môže obohatiť existujúce databázy a služby

PREČO TO MÁ ZMYSEL?

#motivácia

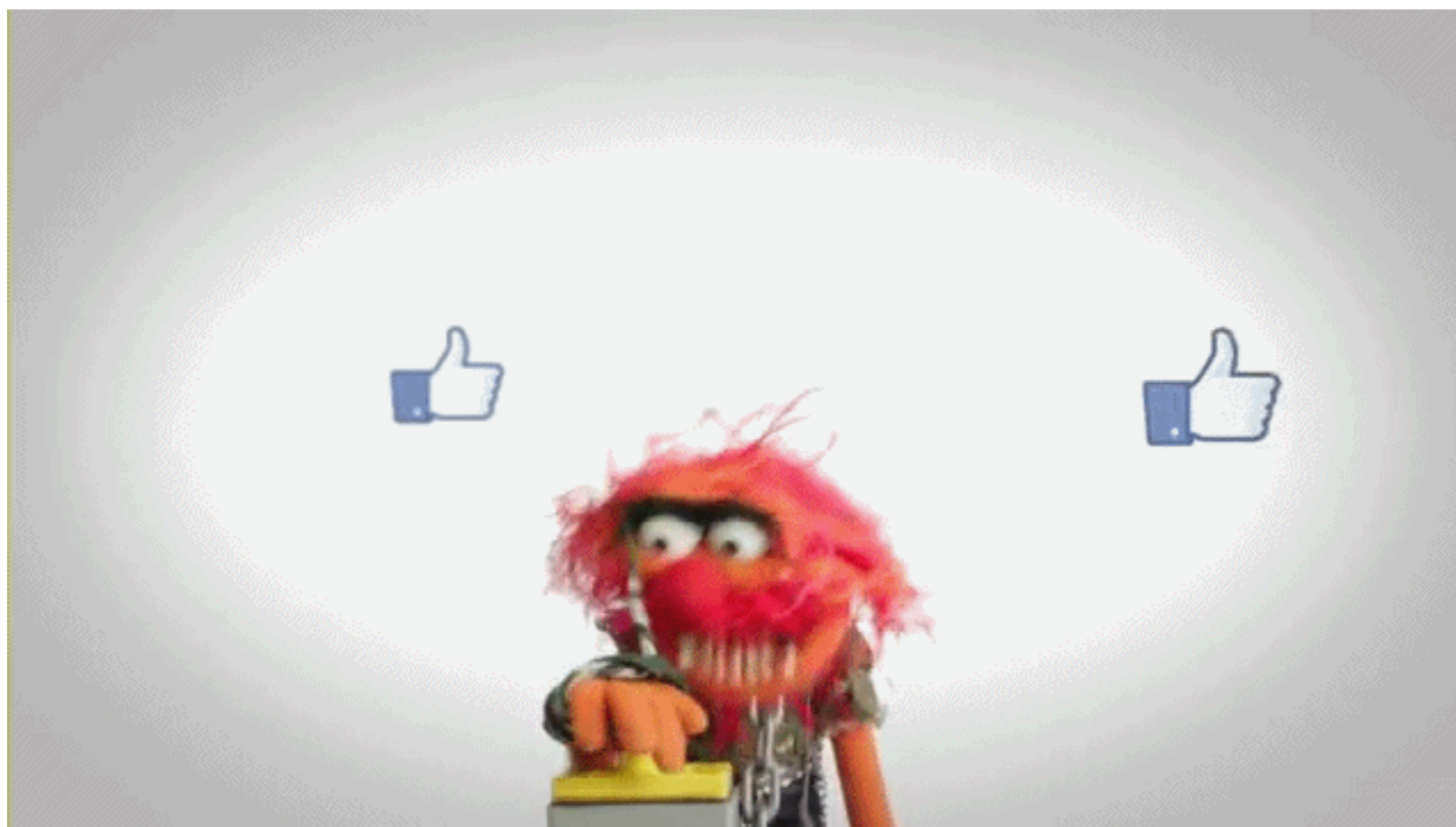
- dostupné metadáta
- často **statické** informácie o filmoch, seriáloch, reláciách
- vychádzajúce spravidla z oficiálnych materiálov alebo rôznych zákutí webu
- dynamika?



www.movieontology.org

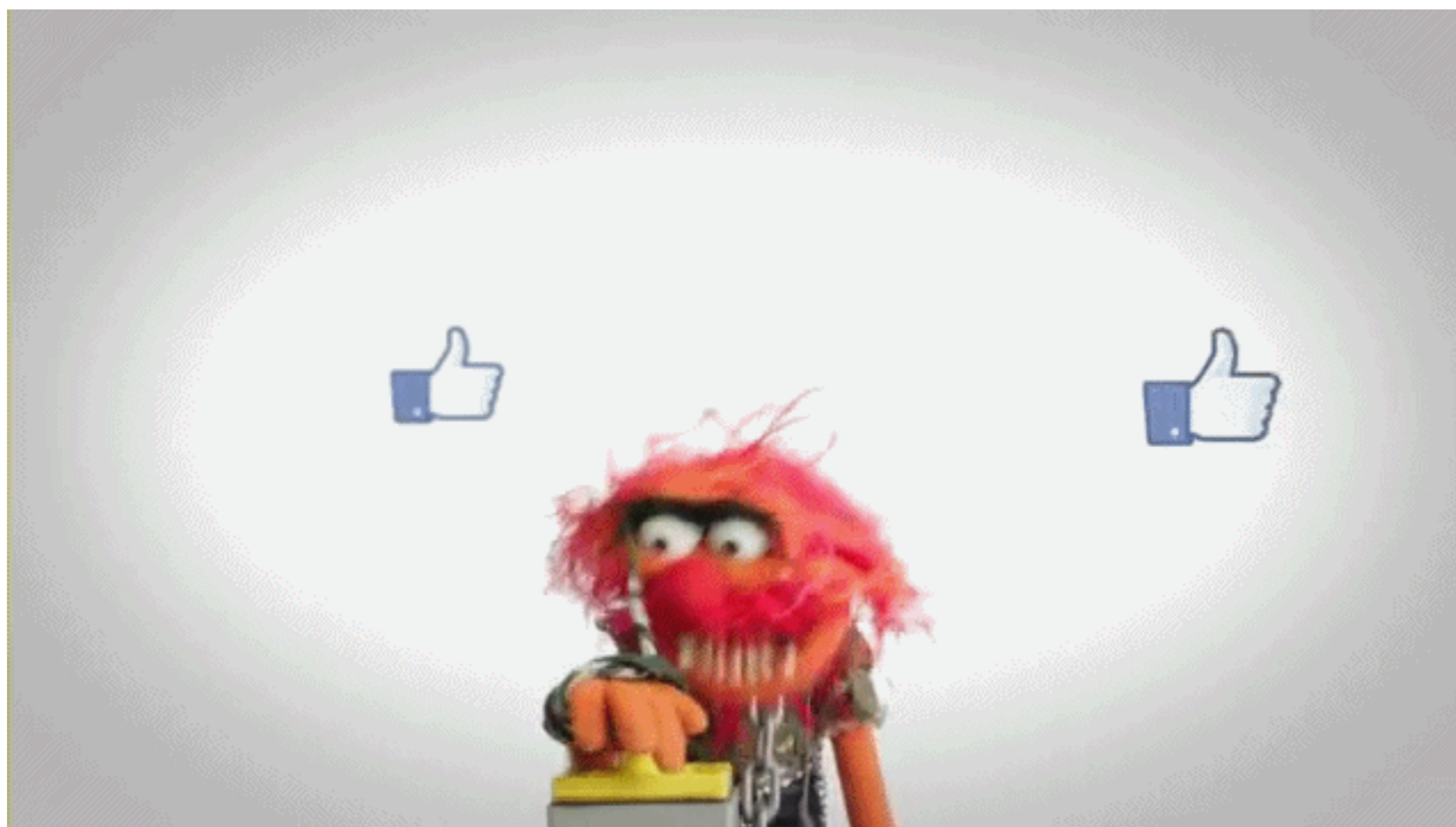
ČO NIE JE STATICKÉ?

#facebook #twitter #instagram



ČO NIE JE STATICKÉ?

#facebook #twitter #instagram

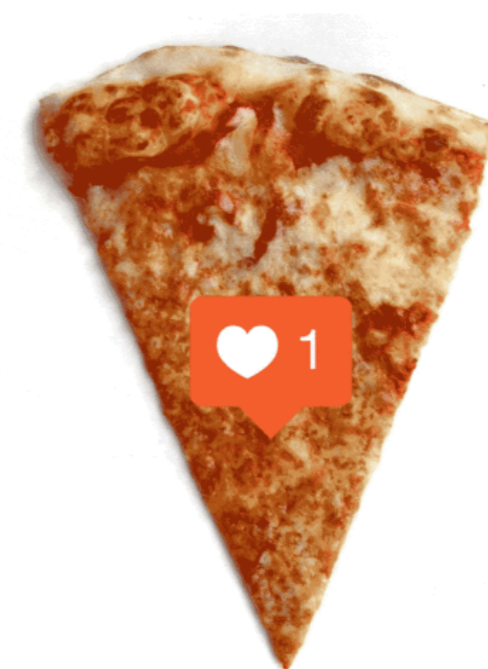


sociálne siete

- sociálne siete ako **second screen**
- množstvo diskutujúcich a interagujúcich používateľov:
 - vyjadrujú svoje názory a pocity
 - zdieľajú zaujímavé odkazy, obrázky, videá
 - (i)relevantný obsah
- priveľa obsahu, ale máme kategorizáciu
 - Facebook stránky, značky (#hashtagy)



- sociálne siete ako **second screen**
- množstvo diskutujúcich a interagujúcich používateľov:
 - vyjadrujú svoje názory a pocity
 - zdieľajú zaujímavé odkazy, obrázky, videá
 - (i)relevantný obsah
- priveľa obsahu, ale máme kategorizáciu
 - Facebook stránky, značky (#hashtagy)



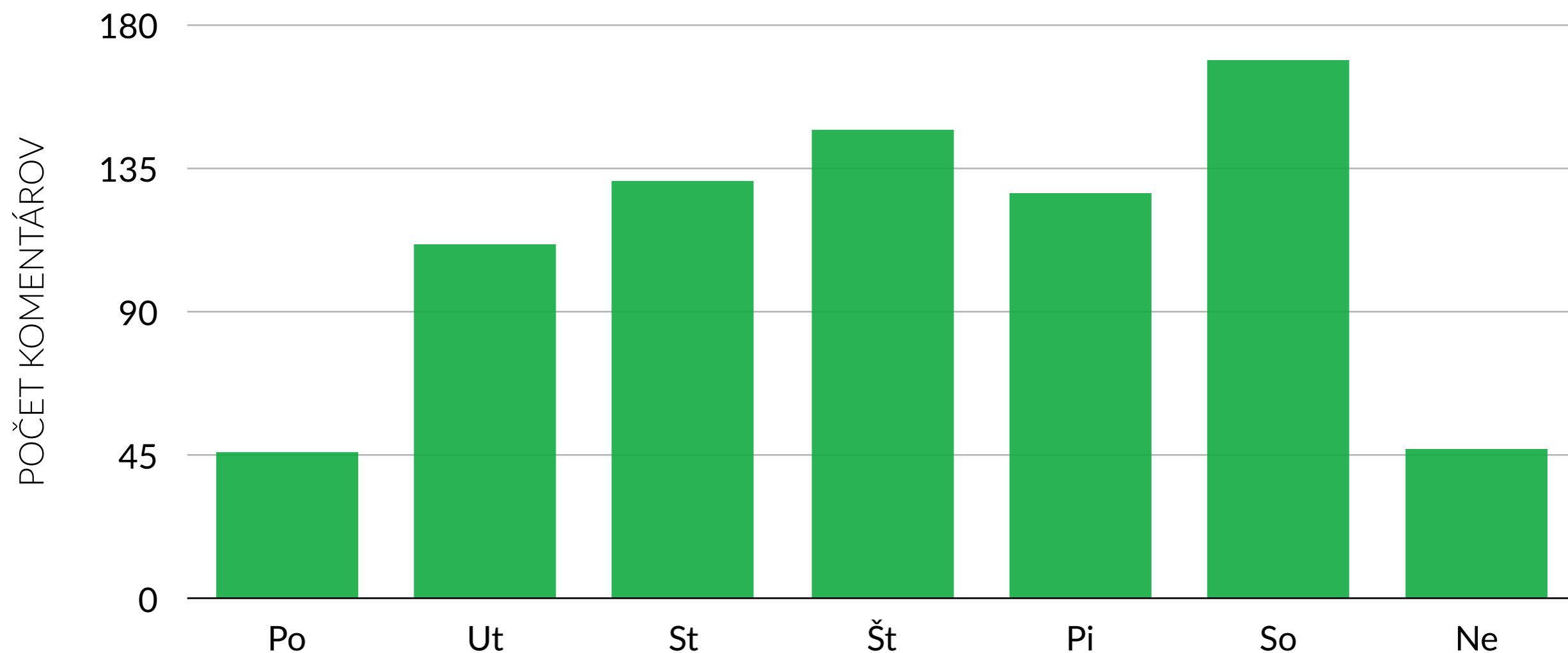
- sentiment - odhalenie postoja alebo pocitu používateľa
 - subjektivita -> hodnotenie
- rôzne prístupy (lexikálne slovníky, smajlíky, ...)
- ale nie vždy je to jednoduché...
 - “happy days of training going to end in a few days #sad #happy”

Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2010. Enhanced sentiment learning using Twitter hashtags and smileys. In Proceedings of the 23rd International Conference on Computational Linguistics: Posters (COLING '10). Association for Computational Linguistics, Stroudsburg, PA, USA, 241-249.

- Facebook stránky TV Markíza a TV JOJ
 - a ich podstránky (pre jednotlivé programy)
- pokrytie programu:
 - gro tvoria najsledovanejšie relácie
 - tiež večerné programy (po 18:00), vlastné relácie a filmy
 - vybrané programy (najmä živé prenosy) bohaté na počet príspevkov

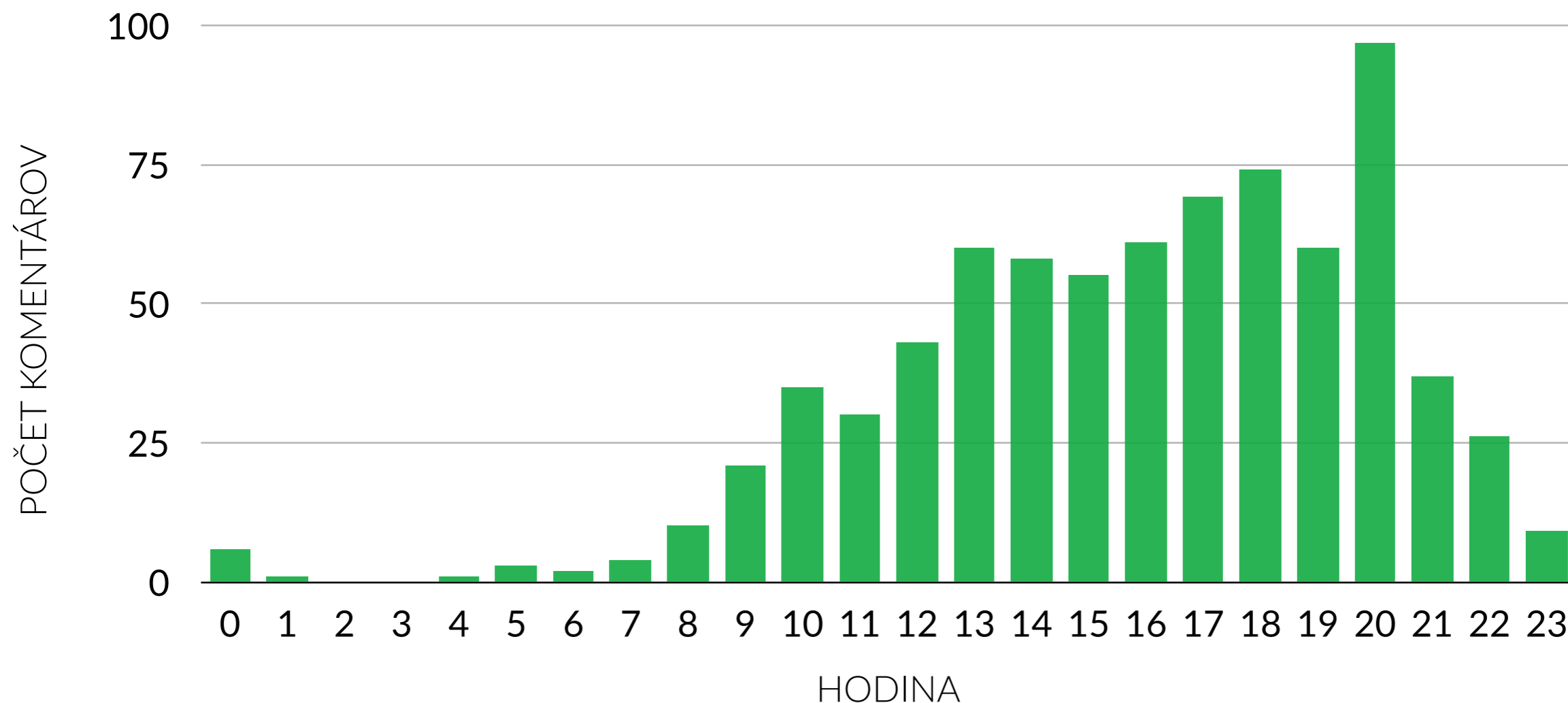
- príspevky televízií:
 - zákulisné články, fotografie, videá, upútavky
 - rôzne súťaže (a ľudia, ktorí na nich “participujú” v komentároch)
- príspevky fanúšikov:
 - komentáre (skrývajú v sebe sentiment a referencie na iné entity)
 - ďalší zaujímavý obsah (fotografie, články, ...)

Komentáre k večernému seriálu - TV Markíza
(2.2. - 22. 2. 2015)



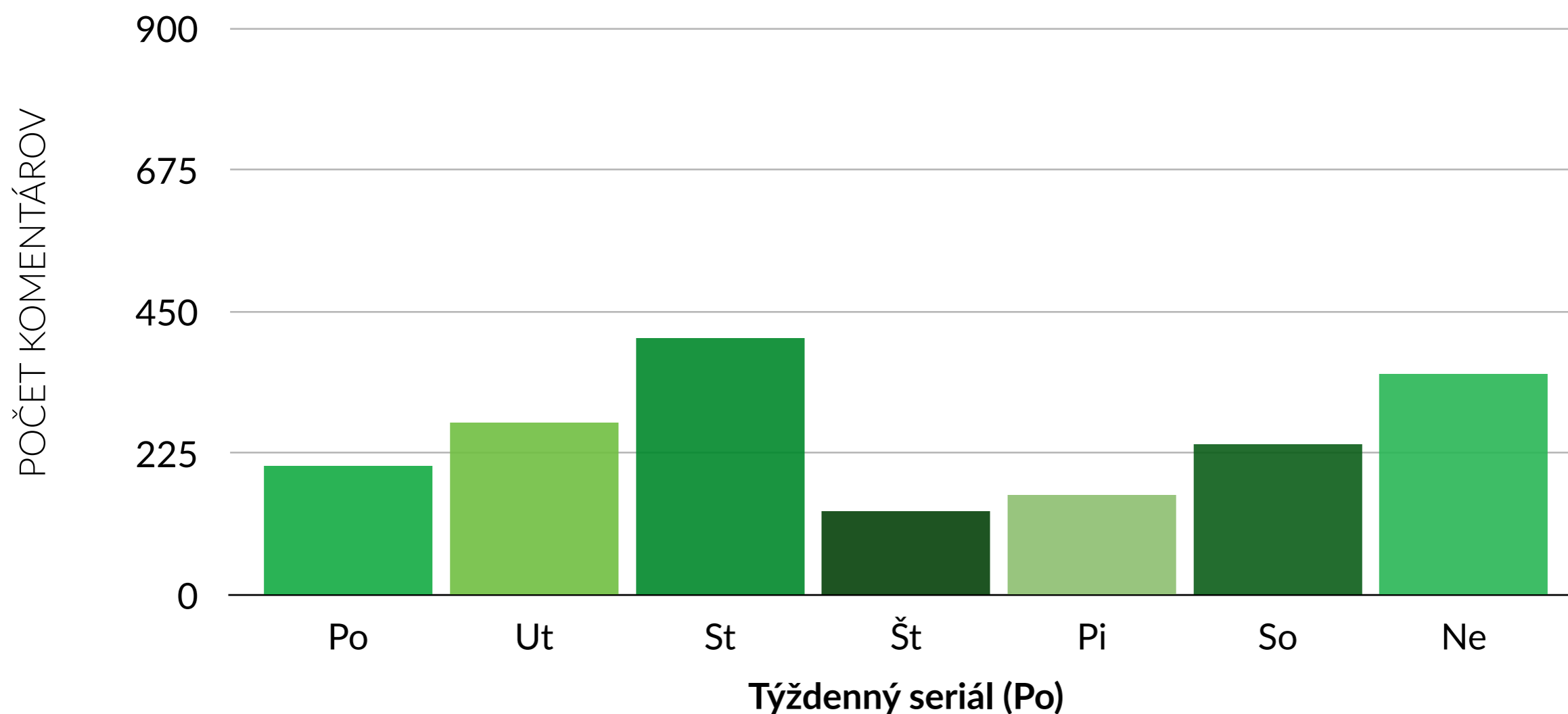
Vysielací čas: Ut - Št 20:30

Komentáre k večernému seriálu - TV Markíza
(2.2. - 22. 2. 2015)

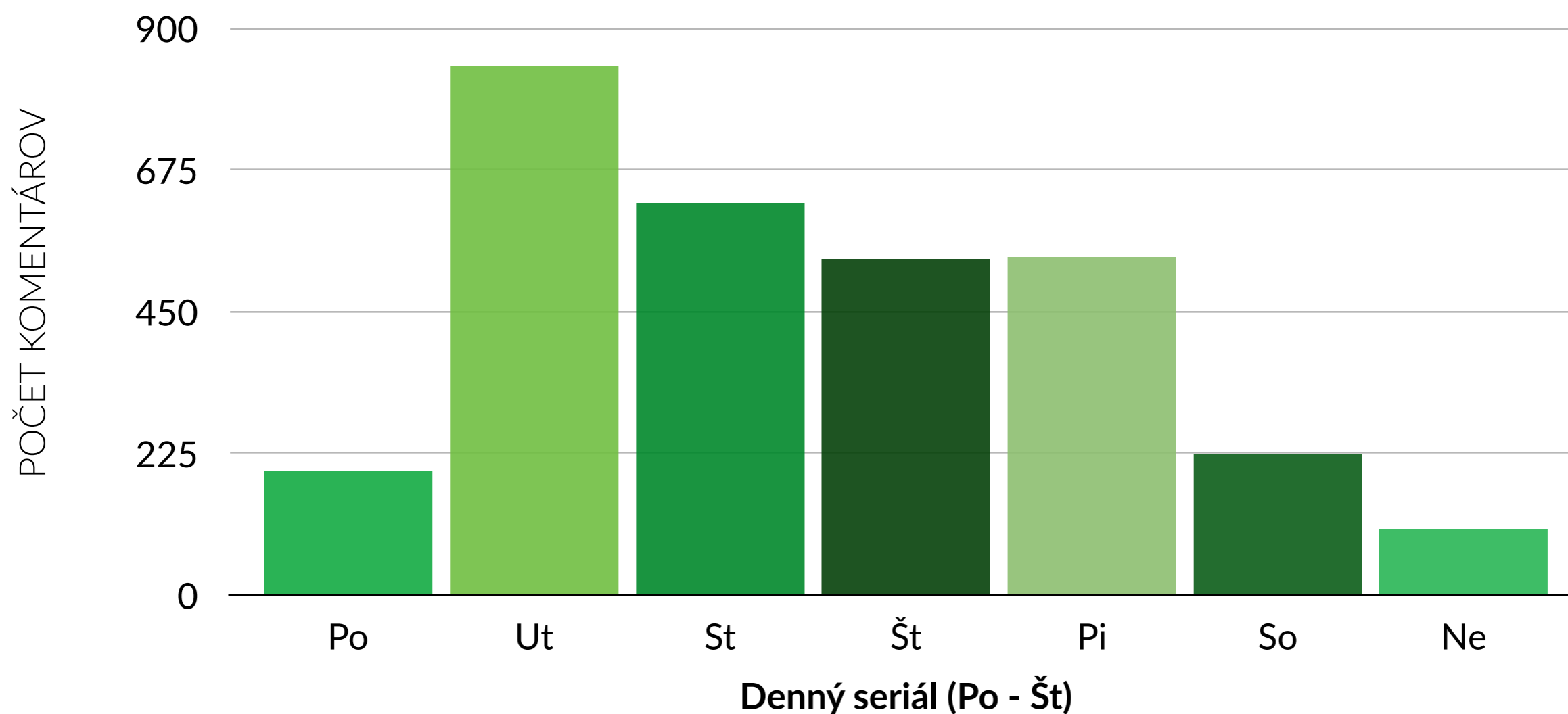


Vysielací čas: Ut - Št 20:30

Počet komentárov týždenný vs. denný seriál
(2.2. - 22. 2. 2015)



Počet komentárov týždenný vs. denný seriál
(2.2. - 22. 2. 2015)



...a teraz zopár príkladov...



...a teraz zopár príkladov...



TO NAJLEPŠIE Z “NAJLEPŠIEHO”

Gott žúroval so svetovou topmodelkou! Kde bola Ivana?

Uplynul približne mesiac odvtedy, čo Jana Kirschner porodila druhú dcéru Yolu. My sme kompletnú rodinku stretli minulú sobotu v Prahe. Pozrite si exkluzívne zábery vydarených dcérok Jany Kirchner.

SMOTANKA.MARKIZA.SK


Páči sa mi to · Komentovať · Zdieľať ·  64  6  1

 64 ľuďom sa to páči.

Najlepšie komentáre ▾

 1 zdieľanie

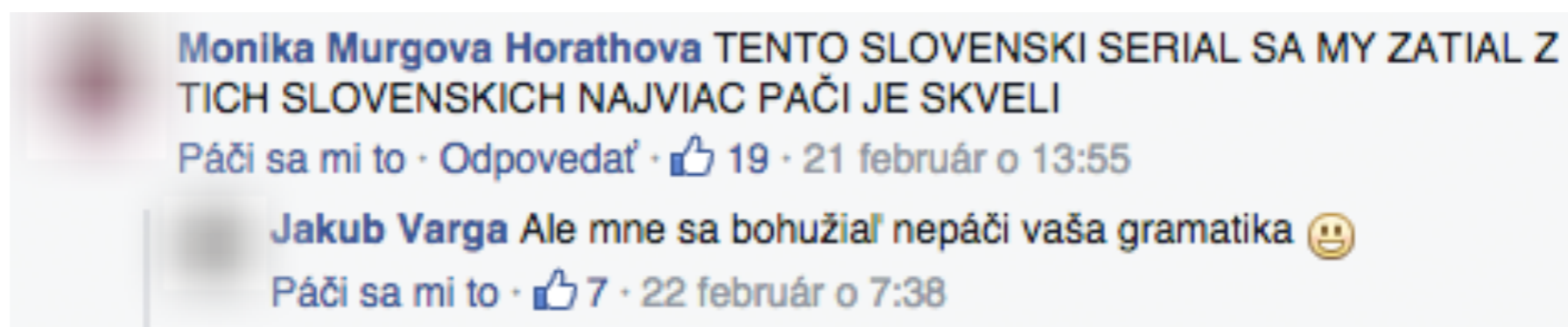
-  **Marika Červeňáková** Ivana bola v Levoči
Páči sa mi to · Odpovedať · 22 február o 8:26
-  **Milan Sutech Cibula** aj tak je to iba ďalšia štetka v smotanke 😞
Páči sa mi to · Odpovedať · 22 február o 7:22
-  **Ivona Horánska** Moje nervy robíte z toho vedu !!!
Páči sa mi to · Odpovedať · 22 február o 3:23
-  **Ildii Rozpávkárka** Ivana bola inde
Páči sa mi to · Odpovedať · 22 február o 13:30
-  **Gabo Hrebíček** pod stolom
Páči sa mi to · Odpovedať · Včera o 10:01

 kto sa dneska pogrca LIKE :d

Páči sa mi to · Odpovedať ·  1 · 21 február o 16:49


TO NAJLEPŠIE Z “NAJLEPŠIEHO” (2)

- problém sú aj preklepy a gramatika




TO NAJLEPŠIE Z “NAJLEPŠIEHO” (3)

- fanúšikovia reagujú na zmeny vysielacieho času a rušenie programov

Моника Чайёва čo strašíte už som chcela mať navretú žilu na krku


Linda Oros No to kdo povedal že nepôjde Panelak??' toto odpovezdte??? Až som sa zlakla???!!!
Páči sa mi to · Odpovedať · 27 február o 21:20

Veronika WeWe Vektorova To uz byva o pol 9???
Páči sa mi to · Odpovedať · Včera o 14:07

Anna Partilová a ty si spadla z oblakov? burlive vino o 20:30 už chodí takmer 2 roky
Páči sa mi to ·  3 · 23 hod.

...ALE AJ PRÍSPEVKY PO 22:00



Problém: veľa príspevkov (a komentárov) mimo TV tému

NÁČRT RIEŠENIA

1. Získanie informácií:

- s využitím API sociálnych sietí.

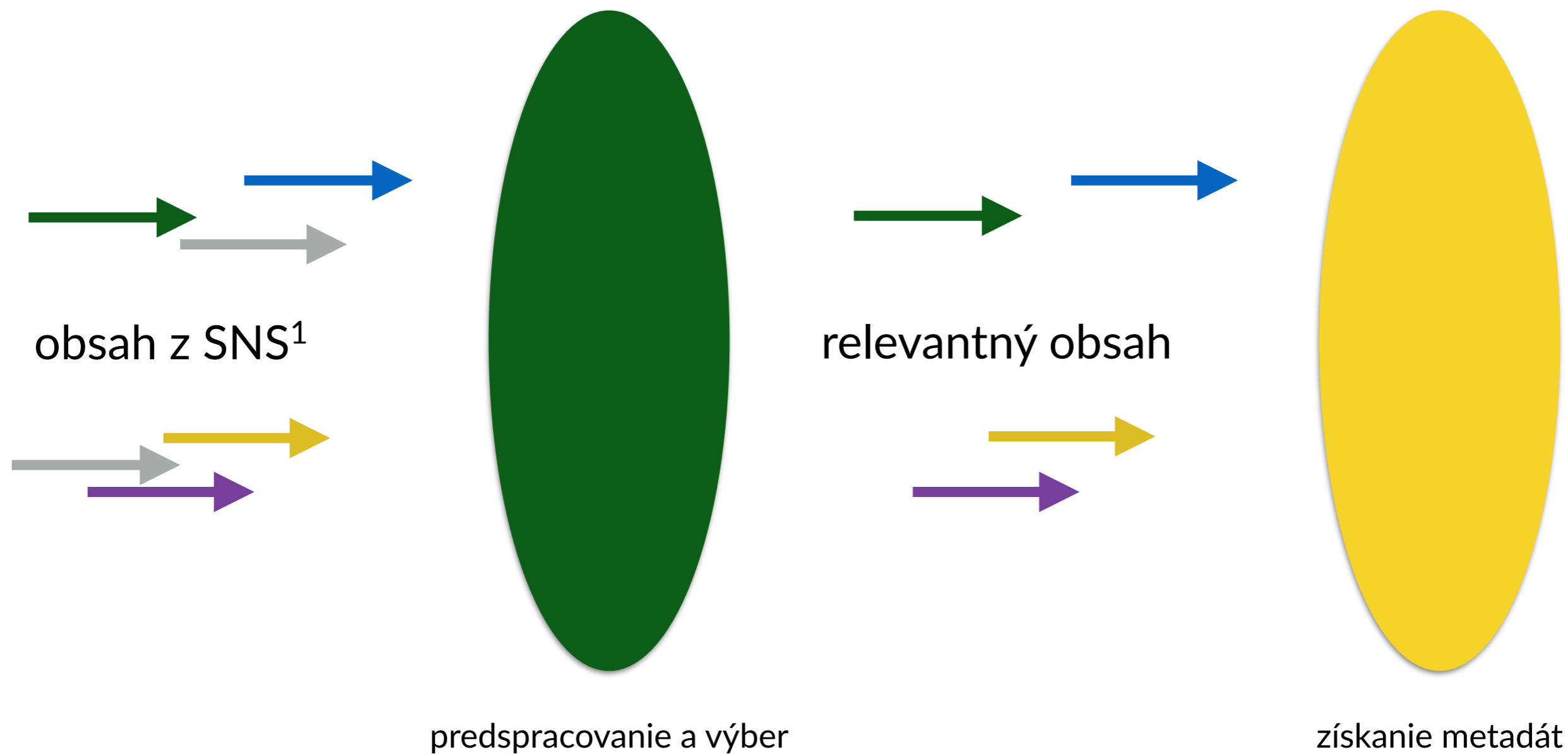
2. Detekcia televízneho obsahu:

- predspracovanie prirodzeného textu, hľadanie pomenovaných entít,
- mapovanie na TV program.

3. Spracovanie získaného obsahu.

4. Poskytnutie metadát.

NÁČRT RIEŠENIA (2)



1) social networking service

- získavanie dát z Facebook API
 - slovenské televízie (Markíza, JOJ)
 - najsledovanejšie nespravodajské programy (slovenské seriály a reality show)
 - svetové dokumentárne televízie (History, Discovery, Travel, Animal Planet)
- prvé pokusy s mapovaním (lematizácia, n-gramy, detekcia pomenovaných entít)

- **inteligentná televízia**
- obohatenie TV programu o nové informácie
 - hodnotenie, zaujímavosti, články, obrázky, ...
- odporúčanie
- lineárna televízia vs. archív



Getty