

Prepájanie multimediálnych metadát

DP II.

Bc. Peter Gašpar

Vedúci: Ing. Jakub Šimko, PhD.

gasparpeter92@gmail.com

Úvod

- problémová doména:
TV program (a jeho entity)
- televízny program, ale aj
filmové databázy
- hlavné statické informácie



Cieľ

- obohatenie existujúcich báz metadát o nové druhy metadát
- využiť iný, dynamický spôsob **extrakcie metadát**
- a vytvoriť tak viac dynamické metadáta



Motivácia

- pomenované entity v príspevkoch sociálnych sietí
 - názvy relácií, postavy, herci, #hashtagy
- **metadát nie je nikdy dosť**
 - statické k dispozícii,
 - dynamické nedostupné,
 - príklad: zaujímavá udalosť z časti seriálu - NLP nad popisom seriálu - **nie vždy sa dá**).



Motivácia (2)

ČESKO SLOVENSKO MÁ TALENT
23 október, 17:25 · 🌐

Krásny víkend plný zábavy vám praje Junior a Marcel! 😊



VS.

ČESKO SLOVENSKO MÁ TALENT

Najväčšia jesenná show! Najlepšie talenty, šokujúce a dojímavé príbehy. Najúspešnejší program TV JOJ sa vracia v novej sérii. Porota: Lucie Bílá, Diana Mórová, Jakub Prachař, Jaro Slávik Slovensko-česká súťažno-zábavná show.MN 12
Rok vydania: 2015

Geniálne vystúpenie skupiny Fair Play Crew si včera vyslúžilo uznanie poroty a standing ovation!

Ako vznikajú ich nápadité choreografie? 😊



Tanečníci majú pri tvorbe nápaditých čísel zásadu: Čím je to primitívnejšie, tým lepšie!

Poliakom sa toto pravidlo vyplatilo.

CSMATALENT.JOJ.SK

👍 Páči sa mi to 💬 Komentár ➦ Zdieľať

357 ľuďom sa to páči.

Najrelevantnejšie ▾

21 zdieľaní



Napište komentár...



Katarina Krištofičová Boli super , určite lepši ako " roboti "

Páči sa mi to · Odpovedať · 👍 11 · 22 október, 15:43

➦ 5 odpovedí



Luboš Hucman Vobec to nebolo primitívne...nech su to niekto skusi dokazat co oni!boli sikovny a zabavny,nato ze Poland tak ich obdivujem a tesim sa na dalsie ich vystupenie:-)

Páči sa mi to · Odpovedať · 👍 1 · 22 október, 15:57

➦ 1 odpoveď

Zobraziť 13 ďalších komentárov

Navrhované riešenie

- využiť sociálne siete,
 - komentáre a správanie používateľov,
 - aktivita televízií.
- Ako?
- Postačujú? Nie!

63%

opakované
reakcie ¹⁾



1) percento statusov, pri ktorých aspoň 1 používateľ pridal aspoň 2 komentáre, facebook.com/csmatalent, máj - október 2015

Zdroj obrázku: Instagram TV Markíza (@tvmarkizaofficial)

Ako na to?

1

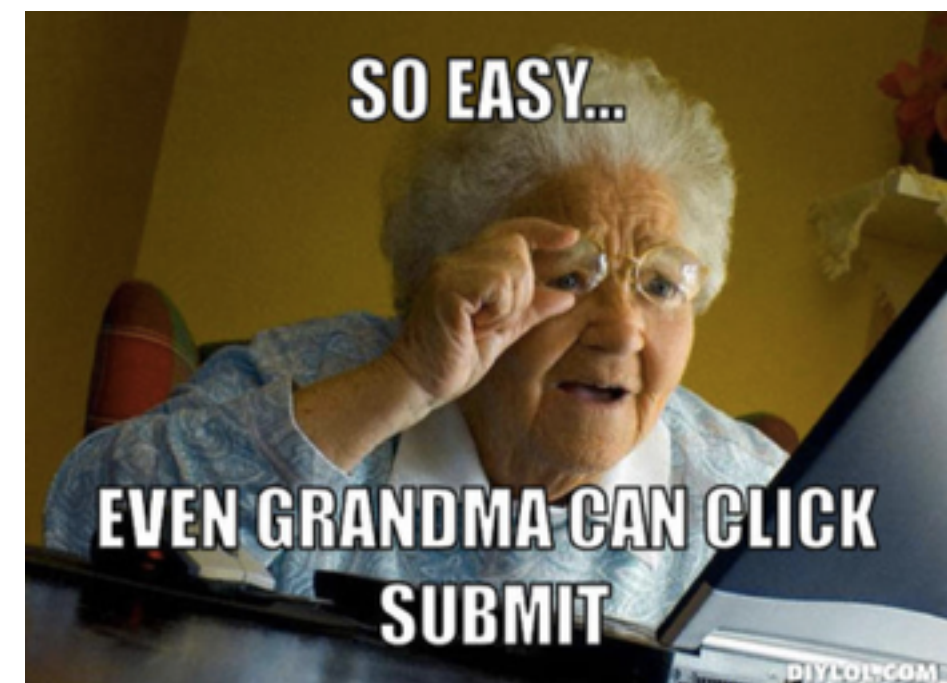
TV program a príspevky zo sociálnej siete

2

Vzájomné namapovanie

3

Extrakcia metadát



alebo ani nie?

Mapovanie TV programu

extrakcia relácií, hercov a postáv
(podľa možností)

TV program

NLP, generovanie hashtagov

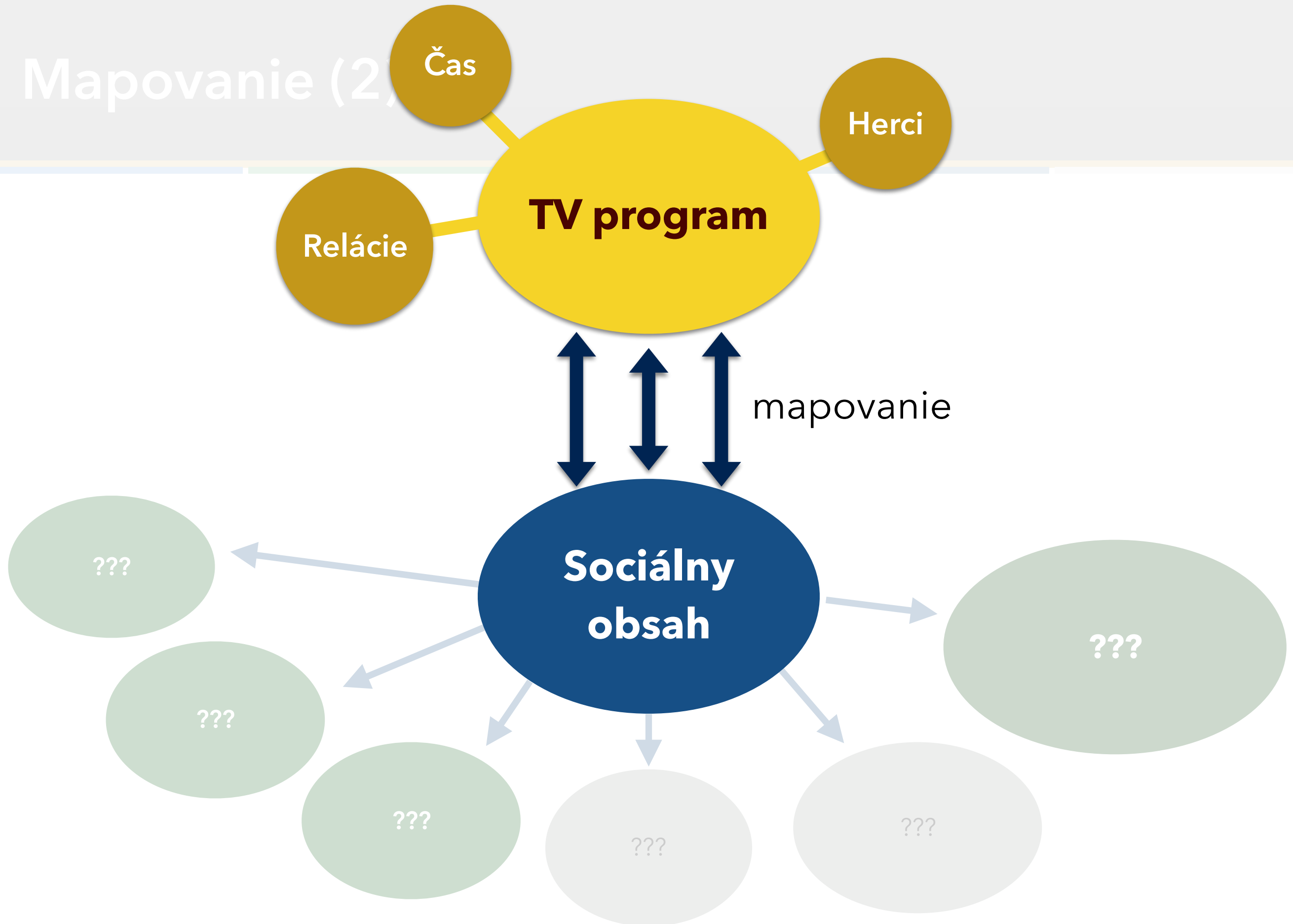
Obsah zo soc. sietí

NLP, extrakcia hashtagov

problém: slovenčina
súčasný prístup: pre články

Mapovanie

- využitie existujúcich prístupov (n-gramy)
- zohľadnenie vysielacích časov



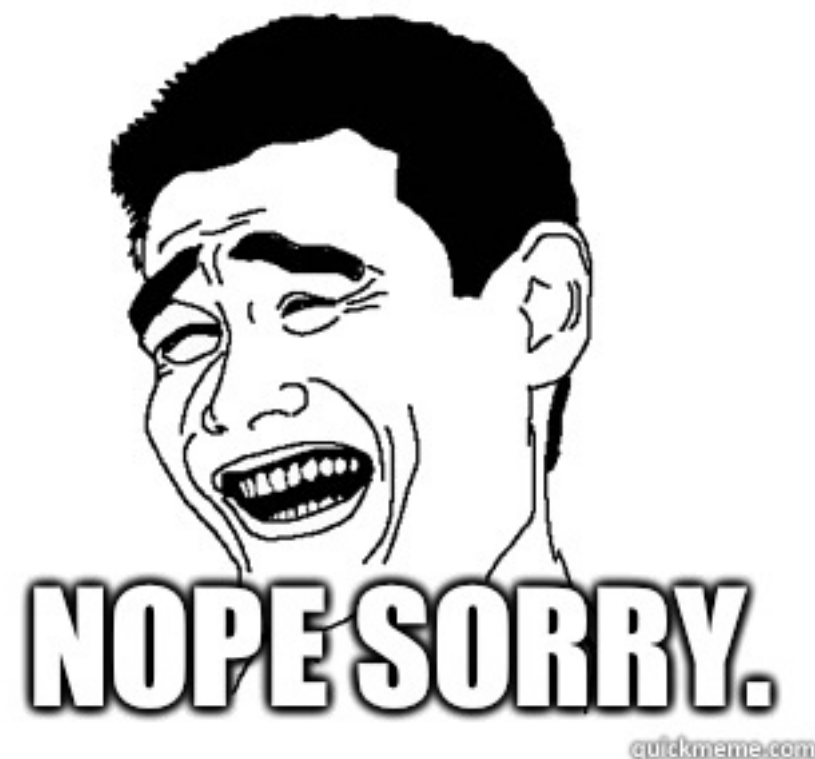
Problémy s mapovaním

- niekedy obsah statusu nestačí,
- žiadne hashtagy ani pomenované entity,
- treba sa však vynájsť :)



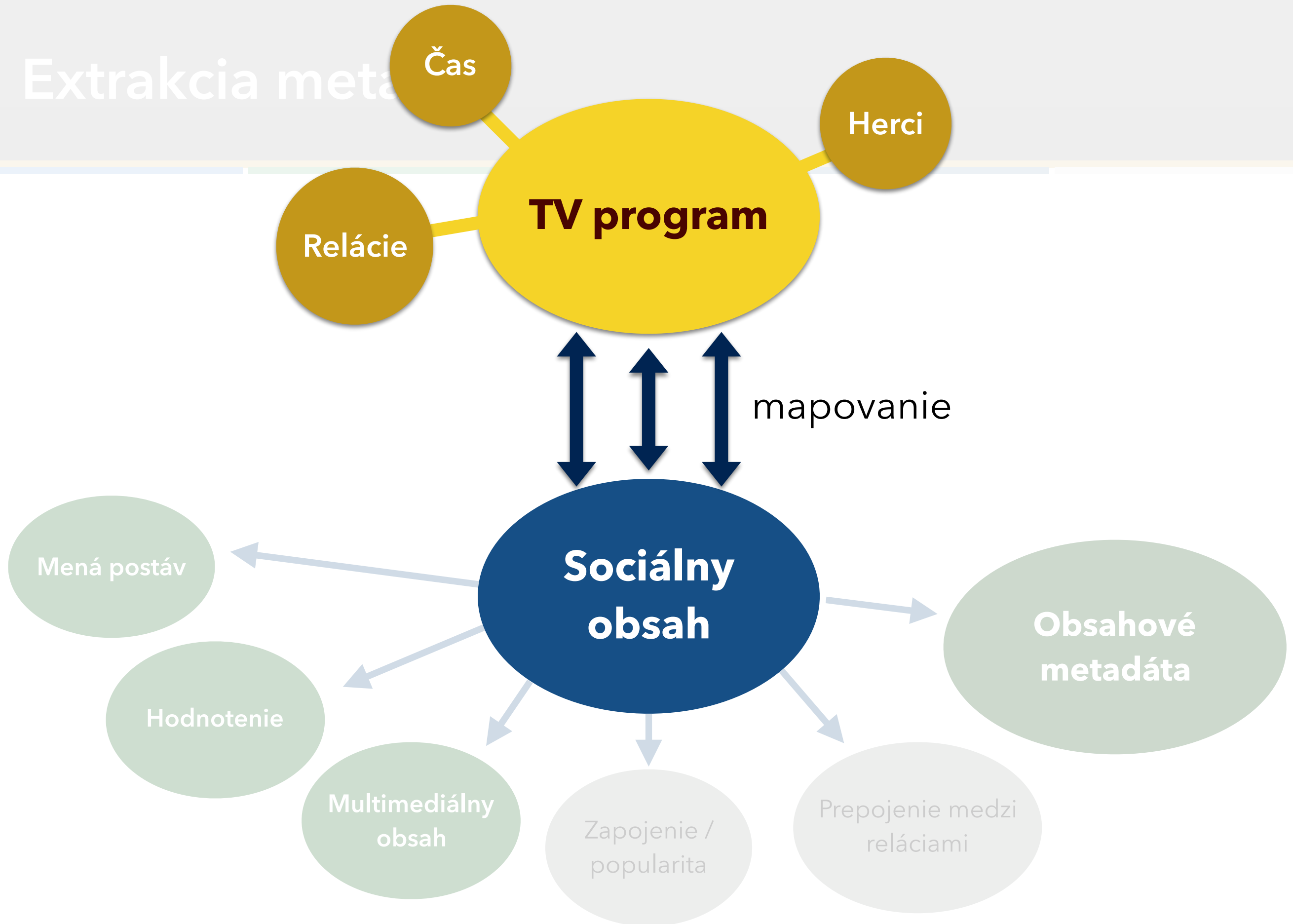
Mapovanie TV programu (3)

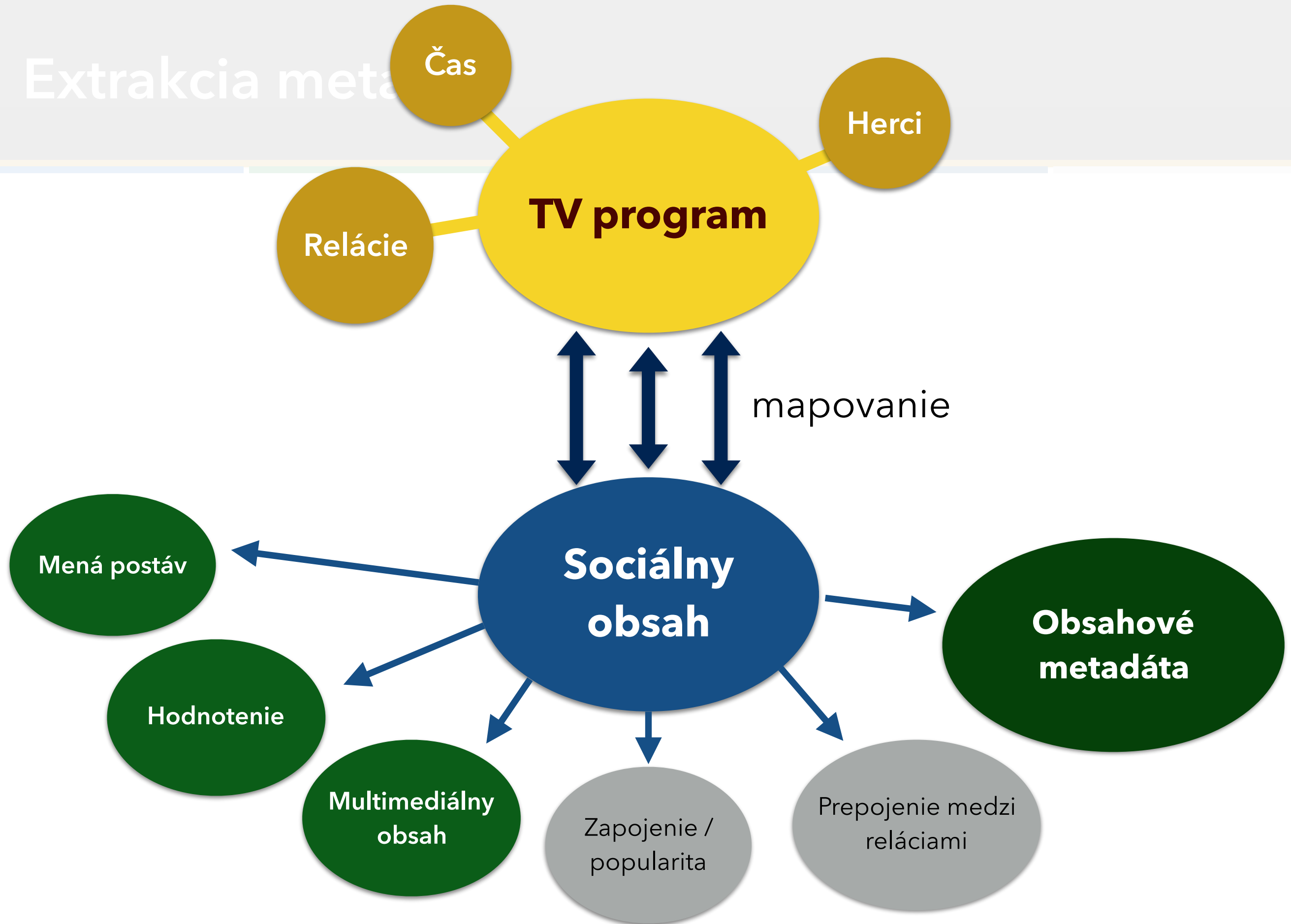
Statusy a komentáre sa kamarátia s TV programom. Máme metadáta!



Extrakcia metadát

- veľa možností
- implementačne napr. rozpoznávanie (a klasifikácia) pomenovaných entít
 - využitie existujúcich riešení a ich adaptácia
- súčasné prístupy:
 - zamerané na Twitter

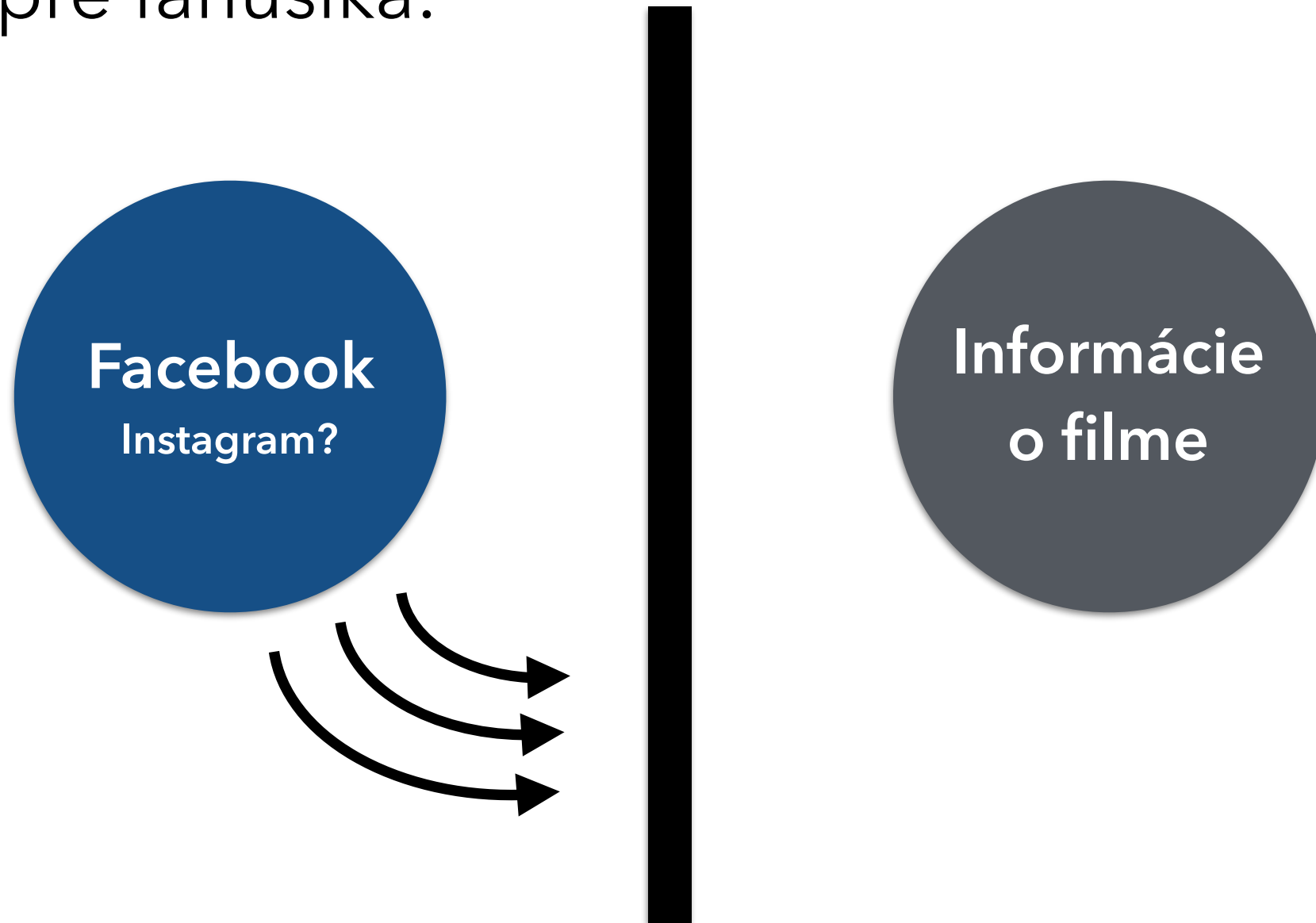




- mapovanie:
 - manuálna anotácia príspevkov zo sociálnej siete,
- extrakcia:
 - zložitejšie,
 - bude vstupovať aj (externý) ľudský faktor.

Aký to má zmysel?

- napr. pre fanúšika:



Aký to má zmysel? (2)

- napr. pre fanúšika:

