



Odhalování tématických sociálních sítí s pomocí WWW

Jiří Jelínek

Katedra managementu informací
Fakulta managementu J. Hradec
Vysoká škola ekonomická Praha

Obsah prezentace

Motivace

Dostupná řešení

Detekce tématických sociálních sítí

Prototyp a výsledky

Další postup

Motivace

Tématická sociální síť –

množina osob spojených s daným tématem,
oblastí či problematikou a jejich vzájemných
vazeb

Identifikace s využitím WWW prostoru jako zdroje
dat

Dostupná řešení – WWW vyhledávače

Google Book Search

Vyhledávat knihy

O této knize

Just for fun: the story of an accidental revolutionary

Autor: Linus



Autor: Linus Torvalds, David Diamond

Spolupracovník David Diamond

Vydáno 2001
[HarperBusiness](#)

Počet stran: 262

ISBN:0066620732

[Přidat do mé knihovny](#)

[Napsat recenzi](#)

Koupit tuto knihu

[Amazon.com](#)

[Barnes&Noble.com](#) - \$14,95

[Books-A-Million](#)

[BookSense.com](#)

[Google Product Search](#)

Vypůjčit si tuto knihu

[Nalézt tuto knihu v knihovně](#)

"I was a nerd. Geek. From fairly early on. I didn't duct-tape my glasses together, but I might as well have, because I had all the other traits. Good at math, good at physics, and with no social graces whatsoever. And this was before being a nerd was considered to be a good thing." In this witty and engrossing narrative, Linus Torvalds, the brilliant mastermind behind the latest Internet revolution, in collaboration with writer David Diamond, chronicles his transformation from a pale, skinny Helsinki college kid to an international folk hero. What began as a childhood hobby soon became the astonishing phenomenon known as the LINUX operating system. LINUX was created because Linus was curious to see if he could improve upon the operating systems already out there, such as UNIX. How could he create a system that could run all of his favorite software with fewer crashes and faster productivity? Linus posted his early versions of LINUX on the Internet and called upon the most brilliant minds in computer science to enhance his system. What happened next took the world by storm. Linus wasn't the first to use the power of the open source movement, but the excitement he generated when he offered his system to the world was unprecedented. What made LINUX revolutionary was its journey to become the marvel it is today -- and how the system grew from the meeting of thousands of minds around the world. Suddenly, Linus's

Zdroj: <http://books.google.com/>

Dostupná řešení – specializované zdroje

A Linux-based Real-Time Operating System

(1997) [\(Make Corrections\)](#) [\(21 citations\)](#)

Michael Barabanov

View or download:

nmt.edu/~rtlinux/papers/thesis.ps

Cached: [PS.gz](#) [PS](#) [PDF](#)

[Image](#) [Update](#) [Help](#)

From: nmt.edu/~rtlinux...homepage.shtml
(more)

Homepages: [M.Barabanov](#)

CiteSeer
Scientific Literature Digital Library

[Home/Search](#) [Context](#) [Related](#)

[\(Enter summary\)](#)

Rate this article: 1 2 3 4 5 (best)

[Comment on this article](#)

Abstract: This work describes the design, implementation, and possible applications of Real-Time Linux --- a hard real-time version of the Linux operating system. In this system, a standard time-sharing OS and a real-time executive run on the same computer. Interrupt controller emulation is used to guarantee a low maximum interrupt latency independently of the base system. The use of a one-shot timer makes it possible to achieve a low task release jitter without compromising throughput. Lock-free FIFO... [\(Update\)](#)

Cited by: [More](#)

Efficient Alternatives for Implementing - Fixed-Priority Schedulers Sergio [\(Correct\)](#)

Contiki - a Lightweight and Flexible Operating System.. - Dunkels, Grönvall, Voigt (2004) [\(Correct\)](#)

QoSOS: An Adaptable Architecture for QoS.. - Moreno, Neto.. (2003) [\(Correct\)](#)

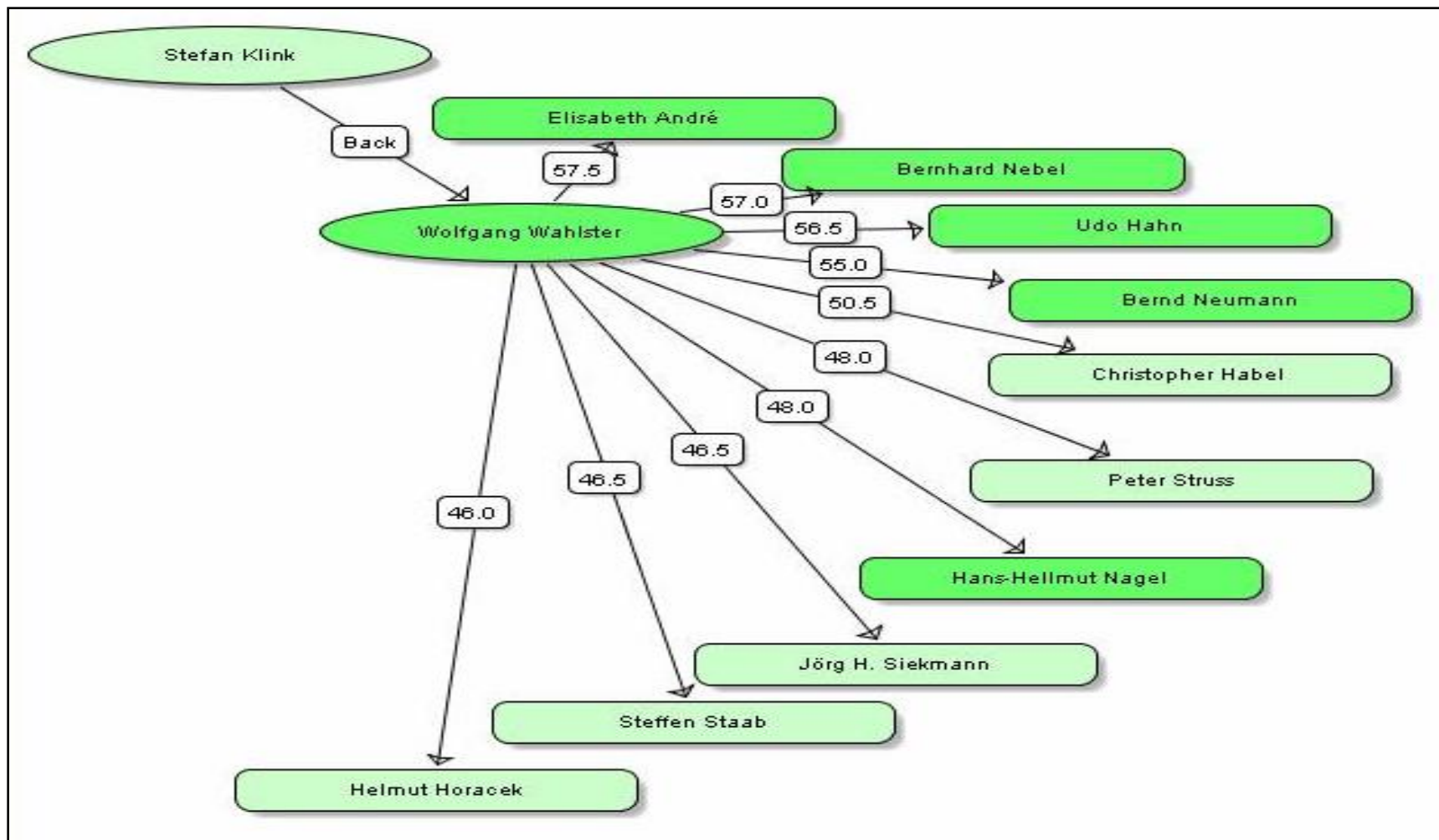
Active bibliography (related documents): [More](#) [All](#)

0.6: A Real-Time Linux - Yodaiken, Barabanov [\(Correct\)](#)

0.5: Concurrent Processing of JBIG Decompression and Skew Estimation - Kanai, Bagdanov

Zdroj: <http://citeseer.ist.psu.edu/>

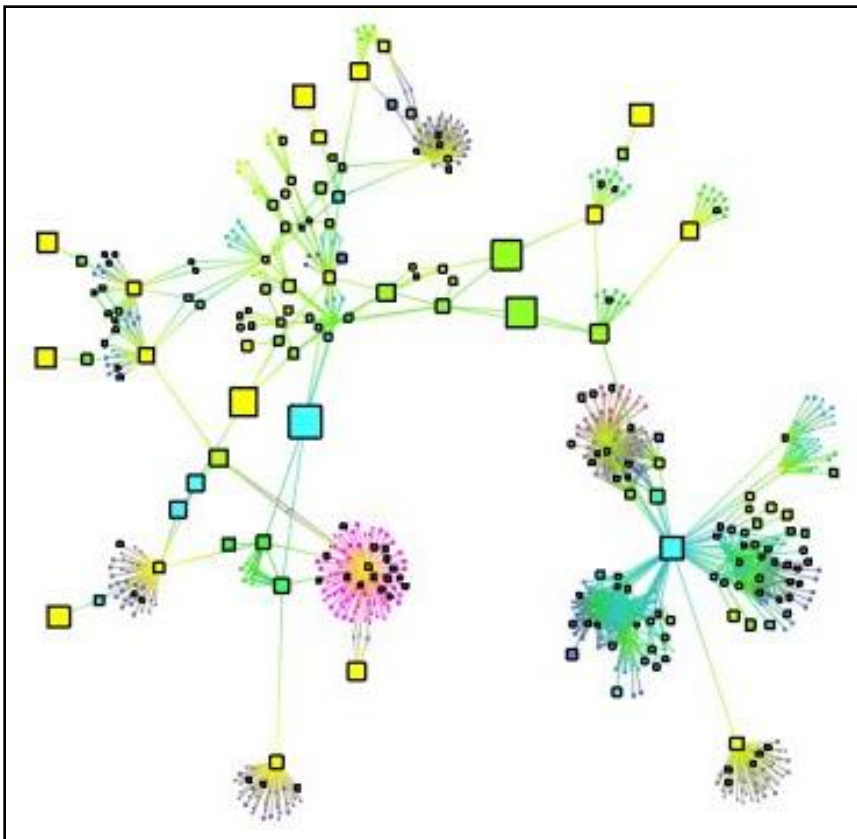
Dostupná řešení – práce se vztahy



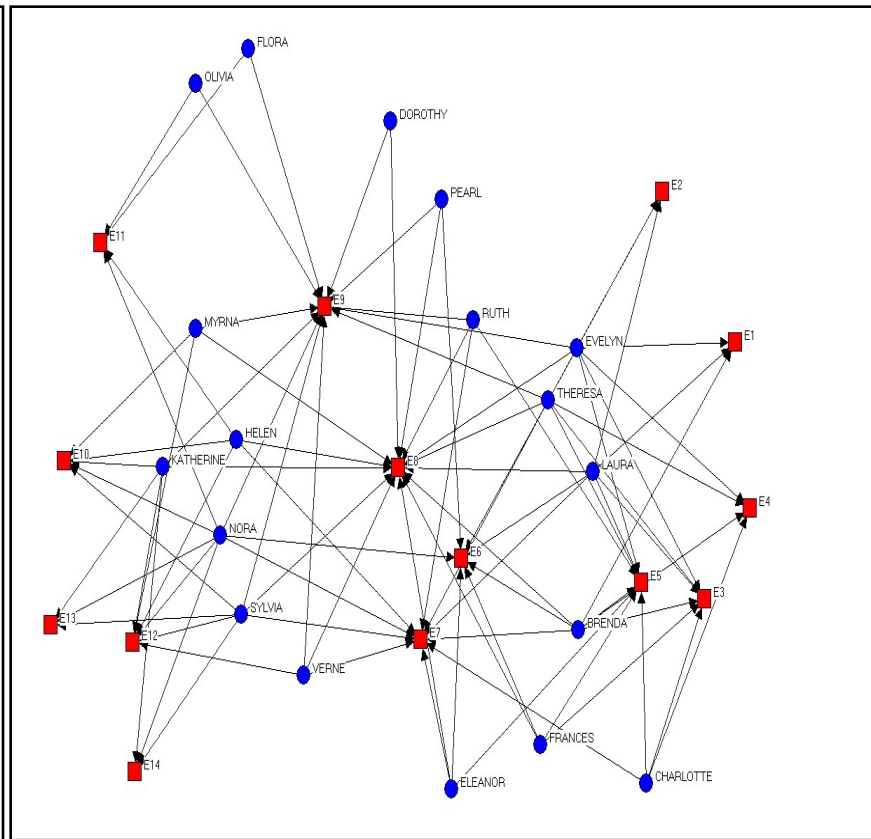
Zdroj: <http://dbis.uni-trier.de/DBL-Browser/ScreenShots/RelAuth.jpg>

Dostupná řešení – vizualizace výstupů

Odhalování tématických sociálních sítí s pomocí WWW

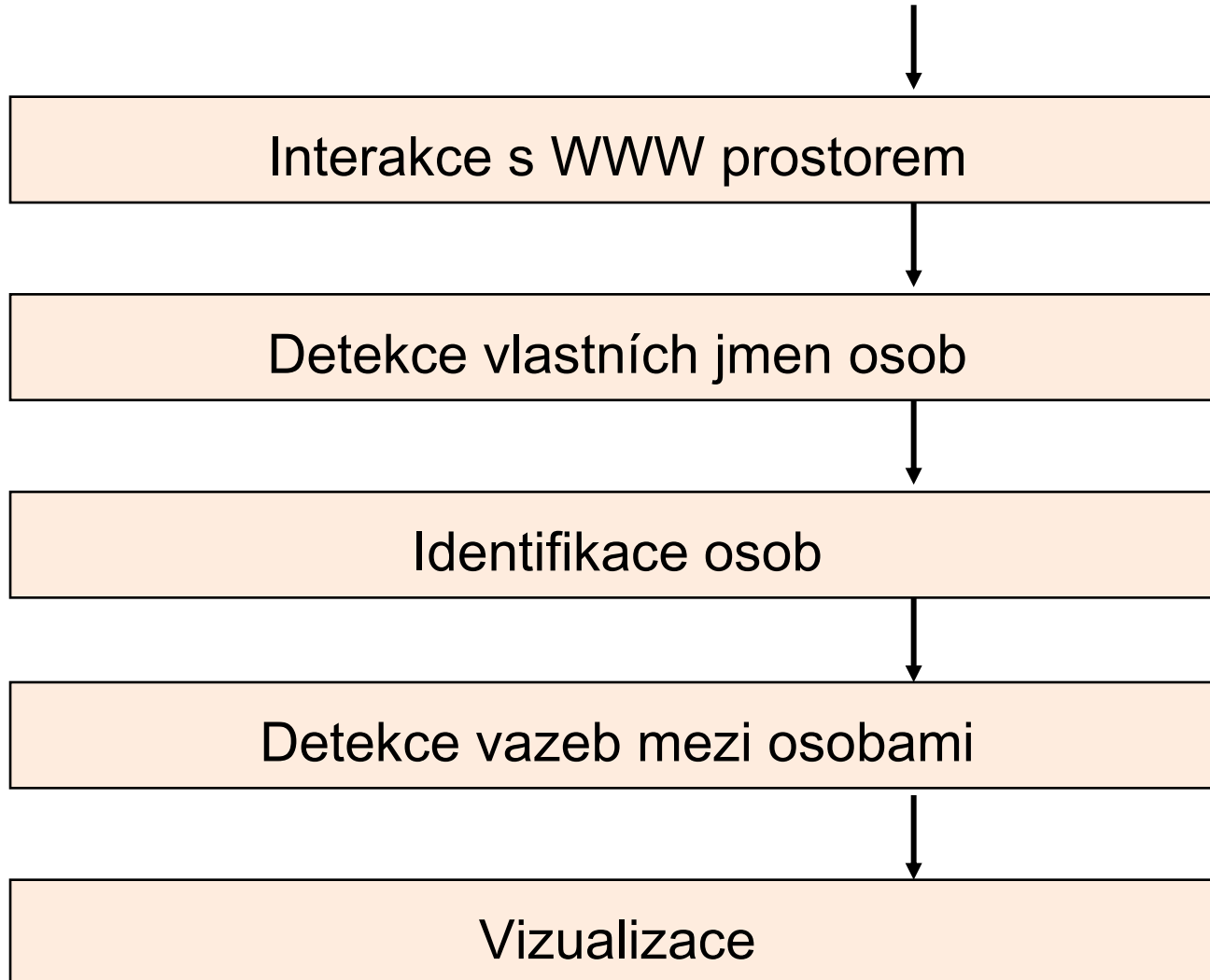


Zdroj: <http://tulip.labri.fr/samples/screenshot02.jpg>

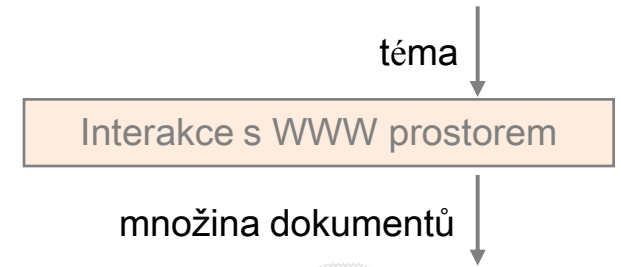


Zdroj: <http://www.analytictech.com/Netdraw/netdrawsamples.htm>

Identifikace tématických sociálních sítí



Interakce s WWW prostorem

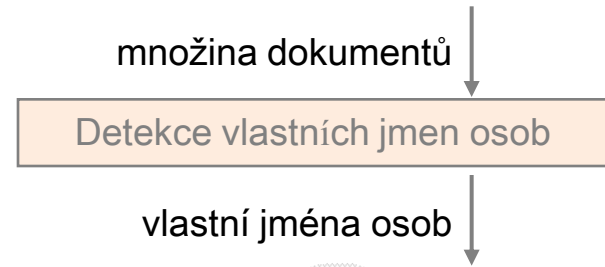


Použití obecných WWW vyhledávačů a dalších WWW zdrojů

Téma reprezentováno klíčovými slovy

Výstupem plné texty

Detekce vlastních jmen osob



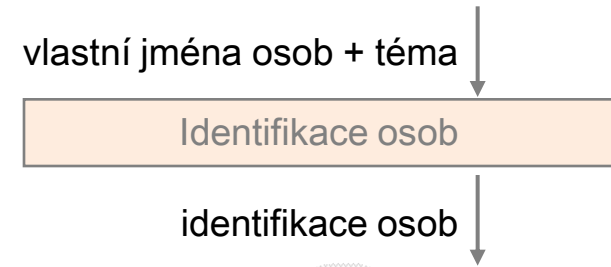
Metody NLP – analýza přirozeného jazyka

Statistický přístup – učení z předchozích rozhodnutí

Slovníky vlastních jmen osob – náročné na vytvoření

Využití kontextu – okolí možných vlastních jmen

Identifikace osob



Odlišit osoby se stejným vlastním jménem

užití tématu jako doplňkové informace

zpětný dotaz

Identifikace osob s různou formou zápisu vlastního jména

John Smith – J. W. Smith – John William Smith

z možných zápisů vybrán ten nejúplnější

Detekce vazeb mezi osobami

identifikace osob + množina dokumentů

Detekce vazeb mezi osobami

identifikace vazeb

$$h_{ijS} = k(w_{iS} + w_{jS}) + (1 - k)p_{ijS}$$

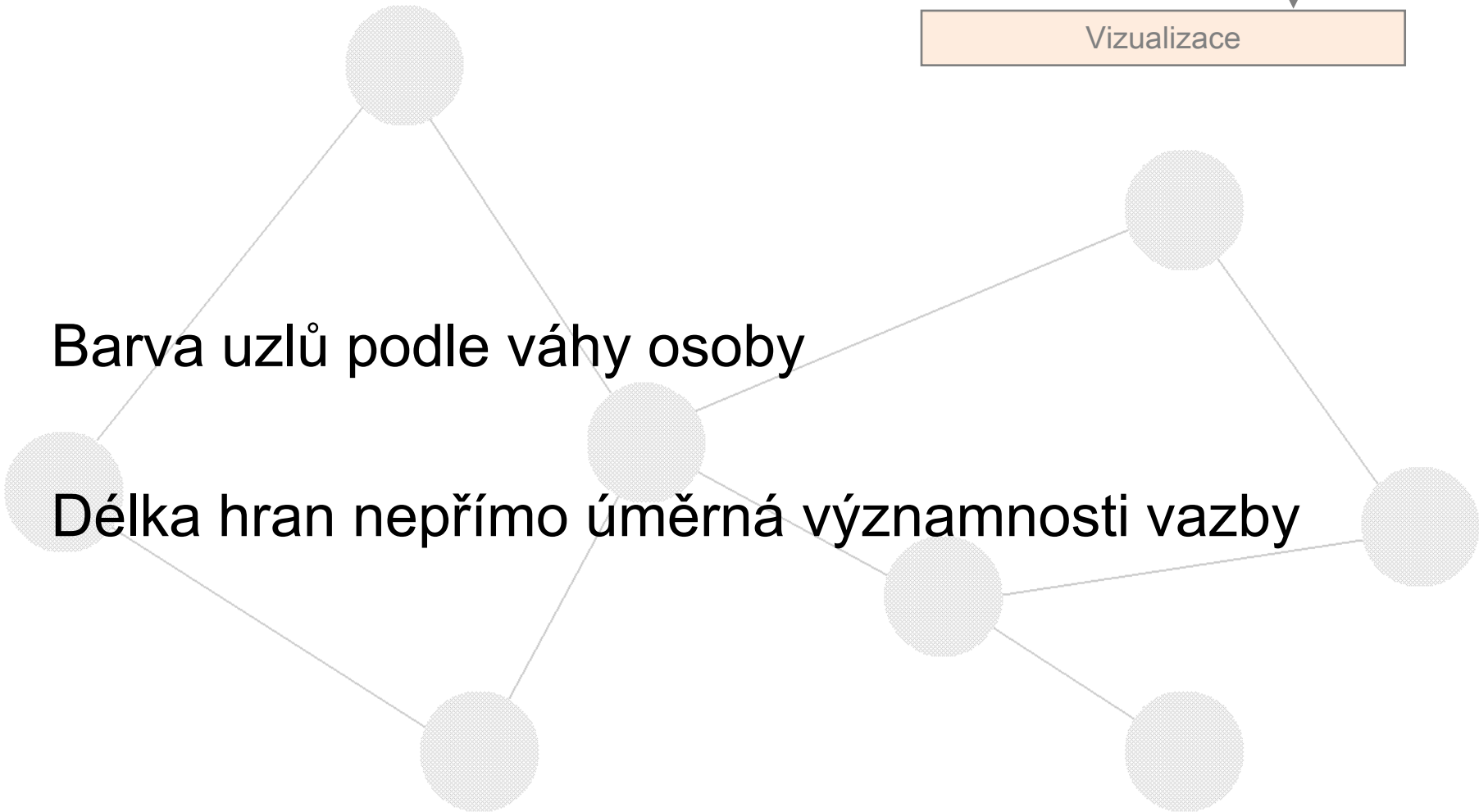
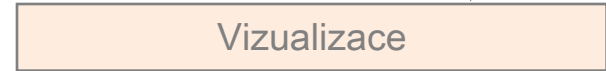
Váha osoby vzhledem k tématu S

Detekce vazeb mezi podle společného výskytu jmen ve vstupních dokumentech

Významnost vazby h_{ijS} mezi termy
vazby významných osob nebo pevné týmy – koef. k

Vizualizace

osoby + vazby + témata



Barva uzlů podle váhy osoby

Délka hran nepřímo úměrná významnosti vazby

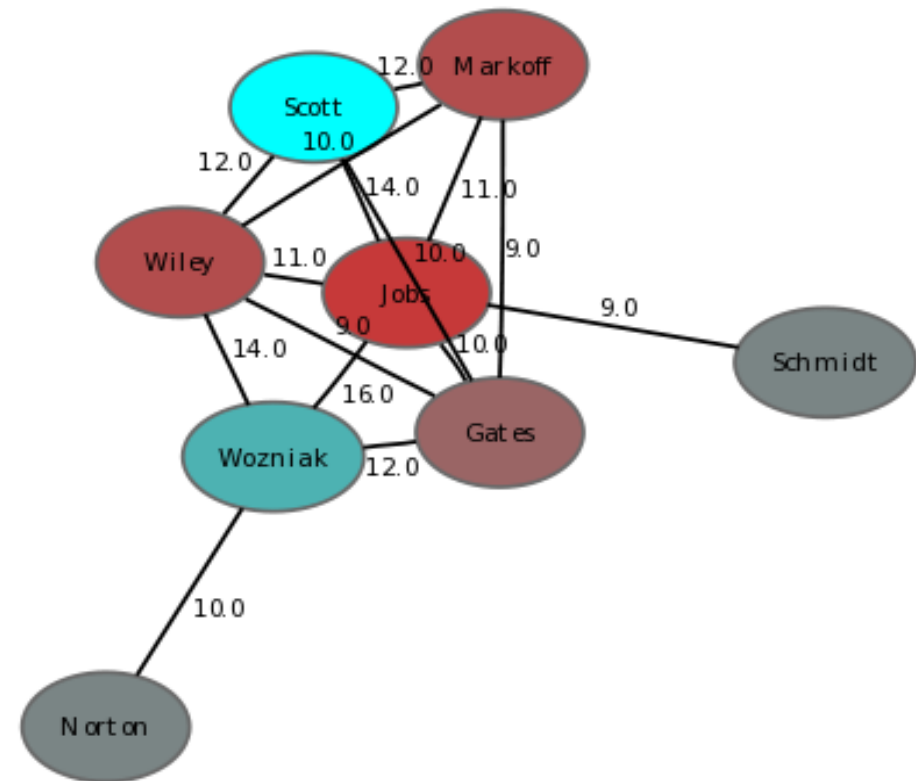
Prototyp a dosažené výsledky

WWW aplikace

35 témat

1462 WWW stránek

53778 klasifikovaných termů, z toho 14041 pozitivně hodnocených jako vlastní jména osob



Další postup

Rozšíření záběru - vstupní importní filtry pro citační servery a další specializované zdroje

Zpřesnění dosahovaných výsledků

- automatické stanovení hodnot uživatelsky nastavitelných koeficientů - genetické algoritmy
- zpřesnění identifikace osob s využitím WWW zdrojů

Vizualizace výstupů - 3D zobrazení

Závěr

Příspěvek do oblasti detekce potenciálních sociálních sítí

Výstupy použitelné všude, kde je potřeba identifikovat tématicky definované sociální sítě výzkum, finančnictví, ekonomika, atd.



Děkuji za pozornost