

AFCEA - seminář PS07 Inteligence

OD DAT KE ZNALOSTEM



Program semináře 17. 1. 2019

Čas	Přednášející	Téma
8:30		Registrace
9:00 - 9:10	Josef Strelec, předseda pracovní skupiny AFCEA PS07 Inteligence	Zahájení semináře, úvodní slovo
9:10 - 9:40	Jan Černocký, <i>Fakulta informačních technologií, VÚT v Brně</i>	Dolování informací z řeči pro obranu a bezpečnost <i>Speech recognition, gender recognition, speaker recognition.</i>
9:40 – 10:10	Tomáš Bia, <i>Phonexia</i>	Revoluce v rychlosti a přesnosti hlasové biometrie. <i>Nová generace technologie pro hlasovou biometrii přináší nové možnosti jejího využití v oblasti bezpečnosti a obrany. Technologii založenou výhradně na neuronových sítích uvedla na trh jako první na světě česká společnost Phonexia. Ukážeme vám, jak hlasová biometrie výrazně zefektivňuje práci při zpracování nahrávek, zvyšuje přesnost výsledků a dokáže výrazně ušetřit čas analytiků i vyšetřovatelů.</i>
15 min		Přestávka na kávu
10:25 – 10:55	Richard Salač, <i>Datera</i>	Jak se řeší tvarosloví se zaměřením na češtinu <i>Příspěvek pokrývá problematiku podpory tvarosloví ve fulltextovém vyhledávání a analytických řešeních pracujících s přirozeným jazykem. Budou představeny v současnosti nejpoužívanější techniky - lemmatizace a stemming - a různé způsoby jejich realizace. Na reálných příkladech budou ukázány rozdíly mezi jednotlivými přístupy, jejich výhody i nevýhody. Dále budou ilustrována specifika podpory tvarosloví pro extrakci entit.</i>
10:55 – 11:25	Ondřej Bojar, <i>MFF UK, Ústav formální a aplikované lingvistiky</i>	Strojový překlad pro investigativní účely <i>V přednášce představím technologii strojového překladu, její předpoklady a omezení, to jest zejména závislost na dostupnosti přeložených textů a obecně spolehlivost. Dále popíši zkušenosti s naším projektem, kdy překládáme SMS z vietnamštiny do češtiny pro účely Národní protidrogové centrály Policie České republiky.</i>



AFCEA - seminář PS07 Inteligence

OD DAT KE ZNALOSTEM

60 min		Oběd
12:25 – 12:55	Pavel Pecina, <i>MF</i> UK, Ústav formální a aplikované lingvistiky	Metody vyhledávání, elastic search <i>Průspěvek představí oblast vyhledávání informací a zkušenosti Ústavu formální a aplikované lingvistiky s vyhledáváním v multimodálních (audio, video, text) a multilingválních datech (dokumenty v různých jazycích).</i>
12:55 – 13:25	Miroslav Nečas, TOVEK	Techniky skládání střípků informací <i>Zpravodajství je založené na sběru informací z různých zdrojů, jejich ověřování a skládání do uceleného obrazu. Jednou z hlavních výzev je v různých datech rozpoznat relevantní informace o osobách, organizacích, věcech, místech či událostech a jejich vzájemných vztazích. Průspěvek uvádí praktické zkušenosti s využitím různých technik (<u>extrakce entit, ontologie, biometrie</u> a další) pro získání a propojení informací ze strukturovaných i nestruturovaných dat.</i>
15 min		Přestávka na kávu
13:40 – 14:10	Michal Gürtner,, IBM	Entity Matching <i>Schopnost rozpoznávat a párovat (ztotožňovat) entity různých typů od jmen přes adresy po jiné údaje typu názvy produktů je klíčová pro mnoho organizací a jejich aktivity. Společnost IBM disponuje metodikami, propracovanými technologiemi a softwarovými nástroji, které s vysokou mírou pravděpodobnosti dokáží data správně párovat. Ve svém portfoliu má i unikátní, celosvětovou znalostní databázi jmen a variant jejich zápisu v různých kulturách. Tato znalostní databáze je vhodná pro párování a analýzu jmen osob a organizací bez ohledu na to v jakých znakových sadách jsou zapsána. V rámci prezentace bude představena metodika, nástroje i zmíněná znalostní databáze včetně ukázky vybraných lingvistických parametrů, které se při párování využívají.</i>
14:10 – 14:40	Luděk Michálek, Policejní akademie ČR	Možnosti využití analýzy nestruturovaných dat v praxi bezpečnostních sborů <i>V současnosti dostupné obrovské množství nestruturovaných dat je pro praxi bezpečnostních sborů nejen velkou příležitostí, ale především velkou výzvou. Průspěvek se zabývá typologií nestruturovaných dat využitelných v praxi bezpečnostních sborů, jejich významem a potřebností využití a nastiňuje, jak je možné těchto možností využít a kudy by se v této oblasti mohl ubírat další vývoj.</i>
10 min	Josef Strelec	Ukončení semináře