

DATA A ZNALOSTI 2017

Sborník konference

Editori

Josef Steinberger
Martin Zíma
Dalibor Fiala
Martin Dostal
Michal Nykl



Plzeň, hotel Angelo
Česká republika
5. – 6. října 2017
www.dataznalosti.cz



Data a znalosti 2017 organizují

**Katedra informatiky
a výpočetní techniky**
Západočeská univerzita v Plzni
Technická 8
306 14 Plzeň



Fakulta aplikovaných věd
Západočeská univerzita v Plzni
Technická 8
306 14 Plzeň



**Nové technologie
pro informační společnost**
Západočeská univerzita v Plzni
Technická 8
306 14 Plzeň



Partneři vydání

IntraWorlds s.r.o.
Bělohorská 476/7
301 00 Plzeň



KadeL Data service, spol. s r.o.
Pod Vinicemi 931/2
301 00 Plzeň



RTsoft s.r.o.
Pod Všemi svatými 40
301 00 Plzeň



DATA A ZNALOSTI 2017

Sborník konference

Editori

Josef Steinberger
Martin Zíma
Dalibor Fiala
Martin Dostal
Michal Nykl

Plzeň, hotel Angelo
Česká republika
5. – 6. října 2017
www.dataznalosti.cz

Vydáno: Západočeskou univerzitou v Plzni

Data a znalosti 2017

Editoři

Josef Steinberger, Dalibor Fiala, Michal Nykl
Nové technologie pro informační společnost
Fakulta aplikovaných věd
Západočeská univerzita v Plzni
Technická 8
306 14 Plzeň

Martin Zíma, Martin Dostal
Katedra informatiky a výpočetní techniky
Fakulta aplikovaných věd
Západočeská univerzita v Plzni
Technická 8
306 14 Plzeň

Partneři vydání

IntraWorlds s.r.o.
KadeL Data service, spol. s r.o.
RTsoft s.r.o.

Autoři příspěvků jsou uvedeni v obsahu. Každý příspěvek byl recenzován, recenzenti jsou členy programových výborů konferencí.

Tato publikace neprošla redakční ani jazykovou úpravou.

Vydala

Západočeská univerzita v Plzni
P.O.Box 314, Univerzitní 8, 306 14 Plzeň

Elektronická verze sborníku konference

1. vydání, 248 stran

ISBN 978-80-261-0720-0

© Západočeská univerzita v Plzni, 2017

Předmluva

Data a informační technologie ovlivňují naši společnost na každém kroku stále více. Přirozeně tak datové, informační a znalostní inženýrství tvoří významnou složku náplně činnosti mnoha výzkumných skupin působících v Česku a na Slovensku. Ve dnech 5. a 6. října 2017 se v Plzni uskutečnilo již 37. setkání odborníků zabývajících se datovými a znalostními technologiemi, které bylo prostorem pro výměnu poznatků a zkušeností mezi výzkumníky z českých a slovenských univerzit, výzkumných ústavů a také vývojových týmů předních nadnárodních firem. Pohodlné zázemí našlo více než 60 vědců v příjemném prostředí hotelu Angelo přímo naproti proslulému plzeňskému pivovaru.

Konference *Data a znalosti* navazuje na dlouholetou tradici dvou prestižních konferencí *DATAKON* a *Znalosti*. *DATAKON* existoval od roku 2001, kdy navázal na konferenci *DATASEM* s tehdy dvacetiletou tradicí. Konference *Znalosti* se pořádala rovněž od roku 2001. Tyto dvě konference se v roce 2015 sloučily v jednu a daly tak vzniknout současné podobě největšího setkání výzkumníků v oboru datových a znalostních technologií z oblasti celého bývalého Československa.

Hlavními tematickými okruhy konference v roce 2017 byly:

- získávání, ukládání a zpracování Big Data,
- dolování dat,
- vizualizace velkých dat,
- analýza nestrukturovaných dat,
- strojové učení, klasifikační a prediktivní systémy,
- tvorba, publikování a využívání otevřených a propojených dat,
- indexování a vyhledávání textových a multimediálních dat,
- modelování uživatelů, adaptivní a personalizované systémy,
- pokročilá uživatelská rozhraní softwarových a informačních systémů,
- systémy pro správu znalostí v organizacích,
- expertní, inteligentní a agentní systémy, počítačová inteligence,
- počítačová lingvistika,
- ontologické a konceptuální modely a
- automatické odvozování a plánování.

Autoři posílali příspěvky ve formě rozšířených abstraktů v českém, slovenském nebo anglickém jazyce do následujících kategorií:

- výzkumný příspěvek,
- příspěvek o probíhajícím výzkumu,
- aplikační příspěvek,
- vizionářský příspěvek,
- projektový příspěvek a
- doktorandské symposium.

Předmluva

Celkově bylo obdrženo 50 příspěvků, s největším zastoupením příspěvků o probíhající výzkumu. Každý příspěvek posoudili minimálně dva členové programového výboru. Výsledkem recenzování bylo rozhodnutí o přijetí 42 příspěvků. Autoři prezentovali své dílo buď formou kratší prezentace (6 příspěvků), nebo živou diskuzí při posterech (25 příspěvků), které předcházela minutová upoutávka na každou práci. Jeden příspěvek byl vybrán, kvůli jeho zaměření a velmi pozitivnímu hodnocení, pro delší prezentaci a diskuzi na páteční ráno. 10 příspěvků zaslaných do doktorandské sekce bylo představeno formou prezentace. Mladší kolegové tak měli šanci sbírat zkušenosti a také cenné rady od těch zkušenějších. Na programu bylo také pět vyzvaných přednášek na zajímavá a aktuální témata od expertů z firem zvučných jmen.

Zájem o konferenci *Data a znalosti* byl tento rok velký a téměř se vyrovnal loňskému rekordnímu a velmi podařenému setkání na Smolenickém zámku na Slovensku. Jsme velmi potěšeni, že do Plzně přijeli zástupci všech skupin působících v předmětné oblasti v Česku a na Slovensku.

Děkujeme členům programového výboru, kteří se ochotně podíleli na posuzování příspěvků. Zároveň děkujeme Martinu Zimovi za kvalitní práci při přípravě tohoto sborníku a také všem členům organizačního výboru, kteří vynaložili nemalé úsilí o to, aby se Plzeň, „Evropské hlavní město kultury“ roku 2015, stala poprvé dějištěm vědeckých diskuzí konference *Data a znalosti*.

Plzeň, říjen 2017

Josef Steinberger a Dalibor Fiala

Obsah

Zvané přednášky (abstrakty)

Sémantická analýza ve forenzním vyšetřování	2
<i>Kateřina Veselovská</i>	
Data Science a umělá inteligence v O2.....	3
<i>Jan Romportl</i>	
Bot framework a Cognitive services aneb sestavte si vlastního inteligentního bota	4
<i>Lukáš Kohut</i>	
Jak se pečou data v Socialbakers	5
<i>Milan Lepík</i>	
Moderní přístupy ke strojovému generování textů	6
<i>Jiří Materna</i>	

Výzkumné příspěvky

Relační a NoSQL databáze: dvě strany téže mince?	8
<i>Jaroslav Pokorný</i>	
Interaktívna vizualizácia výsledkov vyhľadávania informácií pomocou konceptových zväzov.....	15
<i>Veronika Novotná, Peter Butka, Miroslav Smatana</i>	
Včasná identifikácia trendov v správaní používateľov elektronického zľavového portálu	20
<i>Ondrej Kaššák, Mária Bieliková</i>	
Towards User-friendly and High-performance Analytics with Big Data Historian	27
<i>Martin Possolt, Václav Jirkovský, Marek Obitko</i>	
Ontology Learning for Facilitating Ontology Matching in Automotive	31
<i>Ondrej Šebek, Václav Jirkovský, Nestor Rychtycký, Petr Kadera</i>	

Aplikační příspěvky

Minimal Transportation Disruptions Model and Ontologies for Modelling of Disruptive Events	36
<i>Josef Petrák</i>	
Porovnanie algoritmov na analýzu sekvencií pohľadu	41
<i>Róbert Móro, Michal Melúch, Martin Mokry, Mária Bieliková</i>	

Obsah

OLAP Recommender: Supporting Navigation in OLAP Cubes Using Association Rule Mining	46
<i>Bohuslav Koukal, David Chudán, Vojtěch Svátek</i>	
Recommending News Articles using Rule-based Classifier.....	51
<i>Christián Golian, Jaroslav Kuchař</i>	
Využití EasyMiner API v projektu OpenBudgets.eu.....	56
<i>Stanislav Vojíř, Václav Zeman, Jaroslav Kuchař, Tomáš Kliegr</i>	
Hodnocení (ne)zajímavosti asociačních pravidel za využití báze znalostí	61
<i>Přemysl Václav Duben, Stanislav Vojíř</i>	
Efektivní analýza velkých dat pomocí Apache Spark a samoučících neuronových sítí na jediném počítači	67
<i>David Andrešič, Petr Šaloun</i>	
Získávání dat z bibliografických databází	72
<i>Dalibor Fiala</i>	
SISel: Aviation Safety Powered by Semantic Technologies	77
<i>Martin Ledvinka, Petr Křemen, Bogdan Kostov, Miroslav Blaško</i>	
Projektové příspěvky	
Analýza zpravodajských textů a jejich komentářů napříč jazyky	84
<i>Josef Steinberger</i>	
First Insight into the Processing of the Historical Documents from the Period of Totalitarian Regimes	89
<i>Lucie Skorkovská, Petr Neduchal, Zbyněk Zajíc, Pavel Ircing, Luděk Müller, Lukáš Bureš</i>	
Projekt MONSOON – návrh platformy pro analýzu velkých dat v průmyslu	93
<i>Martin Sarnovský, Peter Bednár</i>	
Visionářské příspěvky	
Využití formálních gramatik v automatickém plánování – na cestě k sjednocujícímu modelu.....	99
<i>Roman Barták</i>	
Kripke style Dynamic model for Web Annotation with Similarity and Reliability	105
<i>Michal Kopecký, Marta Vomlelová, Peter Vojtáš</i>	

Obsah

Příspěvky o probíhajícím výzkumu

UWB at SemEval 2014 and 2016	111
<i>Tomáš Brychcín, Tomáš Hercig, Lukáš Svoboda, Michal Konkol</i>	
Univerzální řešení domén v relační databázi	116
<i>Martin Zima, Michal Nykl, Martin Dostal</i>	
Data integration for customer preference learning.....	120
<i>Michal Kopecký, Marta Vomlelová, Peter Vojtáš</i>	
Interaktívna vizualizácia hierarchických štruktúr	125
<i>Miroslav Smatana, Peter Butka</i>	
Vyhľadávanie významných konceptov v rámci konceptuálnej analýzy dát	131
<i>Miroslav Smatana, Peter Butka, Zuzana Čabalová</i>	
Hierarchické prístupy k modelovaniu témy v dokumentoch	136
<i>Miroslav Smatana, Peter Butka, Matúš Gore</i>	
Automatizace klasifikace evropských projektů pomocí klasifikátoru	141
<i>Ondřej Zamazal</i>	
Fokusevaná kategorizační síla webových ontologií	146
<i>Vojtěch Svátek, Ondřej Zamazal, Miroslav Vacura</i>	
Anotovanie slovníka pre analýzu sentimentu pomocou PSO	151
<i>Martin Mikula, Kristína Machová</i>	
Pokroky v analýze heterogenních neuroinformatických dat	157
<i>Ondřej Klempíř, Václav Čejka, Jan Tesař, Radim Krupička</i>	
Predikcia spotových cien elektriny	162
<i>Róbert Magyar, Viera Rozinajová</i>	
Exploračná analýza medicínskych záznamov	166
<i>František Babič, Michal Vadovský, Ján Paralič</i>	

Doktorandské symposium

Procedurální znalosti expertů a model GLIF	172
<i>Ondřej Říha</i>	
Personalizované odporúčanie využívajúce vizuálne stimuly	176
<i>Peter Gašpar, Michal Kompan, Mária Bieliková</i>	
The agent-based model of the dynamic spectrum access networks based on the bilateral bargaining	181
<i>Marcel Vološin, Eugen Šlapak, Juraj Gazda</i>	
Predikcia úpadku spoločností s ručením obmedzeným využitím metód pre rozpoznanie odľahlých bodov.....	187
<i>Peter Gnip, Martin Zoričák, Peter Drotár</i>	

Obsah

Detecting Anomalous Trajectories and Traffic Services	192
<i>Mazen Ismael</i>	
Applying Trusted Knowledge in Evaluation Phase of Data Mining.....	198
<i>Viktor Nekvapil</i>	
Učenie s prenosom medzi prirodzenými jazykmi	204
<i>Matúš Pikuliak, Marián Šimko, Mária Bieliková</i>	
Smerom k automatickej detekcii problémov s použiteľnosťou prostredníctvom sledovania pohľadu.....	209
<i>Martin Svrček, Mária Bieliková</i>	
Použitie spracovaných záznamov reči pacientov pre určenie štádia Parkinsonovej choroby	215
<i>Michal Vadovský, Ján Paralič</i>	
Analýza dát za účelom zlepšenia konkrétneho procesu logistickej firmy.....	221
<i>Miroslava Muchová, Ján Paralič</i>	
Information Extraction from the Web by Matching Visual Presentation Patterns	227
<i>Matej Minarik, Radek Burget</i>	
Rejstřík autorů.....	232