

## 2<sup>nd</sup> International Conference on **Big Data Analysis and Data Mining** November 30-December 01, 2015 San Antonio, USA



## Dominik Ślęzak

University of Warsaw, Poland

## **Big Data Analytics with Approximate SQL**

In many big data applications, original data sources are dynamically translated and stored in a relational format. This can be done in order to better adapt to the existing analytical interfaces and to represent the data in a way that is conceptually closer to the endusers who often find it natural to operate with entities, their attributes and relations. This fact surely raises a significant challenge, but also a great opportunity in front of RDBMS systems specialized in data analytics. As a case study with this respect, we consider Infobright's database software which is used worldwide in a number of areas requiring massive data processing. We discuss its architecture that utilizes the principles of rough sets and granular computing in combination with columnar stores and parallel computing. We also outline its development directions aimed at further scalability improvements. Particularly, we concentrate on an idea of introducing a new form of approximate SQL which can open brand new possibilities for big data analytics, both with regard to information storage and processing.

## **Biography**

Dominik Ślęzak received PhD in 2002 from University of Warsaw and DSc in 2011 from Polish Academy of Sciences. In 2005 he co-founded Info bright Inc., where he holds position of Chief Scientist. He is also an Associate Professor at Institute of Mathematics, University of Warsaw. He delivered invited talks at over 20 international conferences. He is co-author of over 150 papers and co-inventor in 5 granted US patents. He serves as Associate Editor for several scientific journals. In 2014 he served as general program chair of IEEE/WIC/ACM Web Intelligence Congress. From 2012-2014 he served as President of International Rough Set Society.

slezak@mimuw.edu.pl

Notes: