

22 November
17:00-18:15 CET

Future Trends in Databases:
From Data Science through Artificial
Intelligence to Quantum Computing

 ifip
60 YEARS

Moderated by:



Tiziana Catarci
Sapienza University



Monica Scannapieco
Italian National
Institute of Statistics



Oscar Romero
Universitat Politècnica
de Catalunya



Paolo Ceravolo
Milan University



Sven Groppe
Lübeck University



Philippe Cudre-Mauroux
Fribourg University


IFIP 60th Anniversary *Future of Information Processing* Event Series

IFIP Working Group 2.6 on Databases

- The Group was established in 1974
- Member of IFIP Technical Committee TC-2 on Software Engineering
- 19 members and 9 observers

Board

- chair: **Robert Wrembel** (Poznan University of Technology, Poland)
- vice-chair: **Paolo Ceravolo** (Università degli studi di Milano)
- secretary: **Antonia Azzini** (Cefriel of Milan)

 ifip
60 YEARS

Some of our recent activities



SIMPDA 2018
EIGHTH INTERNATIONAL SYMPOSIUM ON DATA-DRIVEN PROCESS DISCOVERY AND ANALYSIS
13-14 DECEMBER, 2018 - SEVILLE, SPAIN

Seminars

January 26, 2021

- Paolo Cervolo: Event Log Encoding: assessing the state of the art → [slides](#)
- Antonio Azzini: RADAR, Rich Advanced Design Approach For Reporting
- Robert Wrembel: On warehousing standard and big data → [slides](#)

February 17, 2021

- Oscar Romero: DEEDS - Data Engineering for Data Science: an overview of the Mario Curia project → [slides](#)
- Sven Groppe: Hybrid Multi-Model Multi-Platform (HM3P) Databases → [slides](#)
- Mustafa Jassar: From Wordnets to Linguistic Ontologies, for Knowledge Graphs → [slides](#)

April 14, 2021

- Kai-Uwe Sattler: Graph Data Management in Persistent Memory → [slides](#)
- Besim Blasi (invited speaker from BarcelonaTech): Learning the impact of data pre-processing in data analysis → [slides](#)
- Michał Bodrzyński: Data mesh in hybrid cloud

September 29, 2021

- Sylvio Barbon Jr.: Graph Data Management in Persistent Memory → [slides](#)
- Philippe Cudré-Mauroux: Data Management on New Hardware → [slides](#)
- Semra Maghoul: Graph Embeddings and Data Modeling: a primer → [slides](#)

Journal on Data Semantics (2018) 7:45–65
https://doi.org/10.1007/s10992-018-0289-2

ORIGINAL ARTICLE

Big Data Semantics

Paolo Cervolo¹ · Antonio Azzini² · Marco Angelini³ · Tiziana Catarci⁴ · Philippe Cudré-Mauroux⁵ · Ernesto Damiani⁶ · Alexandra Masala⁷ · Maurice Van Keulen⁸ · Mustafa Jassar⁹ · Giuseppe Santocchia¹⁰ · Kai-Uwe Sattler¹¹ · Monica Scamigliesi¹² · Manuel Wimmer¹³ · Robert Wrembel¹⁴ · Faai Zarkas¹⁵

Received: 17 April 2018 / Accepted: 20 April 2018 / Published online: 23 May 2018
© Springer Nature Switzerland AG 2018

Abstract
Big Data technology has discarded traditional data modeling approaches as no longer applicable to distributed data processing. It is, however, largely recognized that Big Data imposes level challenges in data and infrastructure management, indeed, multiple components and procedures must be coordinated to ensure a high level of data quality and accessibility for the application teams, e.g., data analytics and reporting. In this paper, the field of so-kind to authors by members of EDP-WG 2.6 on Data Semantics, we propose a review of the literature addressing these topics and discuss relevant challenges for future research. Based on our literature review, we argue that methods, principles, and perspectives developed by the Data Semantics community can significantly contribute to address Big Data challenges.

1 Introduction
The term “Big Data” is widely used to designate a discontinuity in data processing and analysis [1, 2]. The early literature described this discontinuity using the “3 V” acronym that highlights the unprecedented data Volume, Velocity (e.g., in terms of input data rate), Variety (in terms of data types, as well as non-uniform Velocity and Value of today’s applications) [3–6]. In other words, data-intensive applications require a data processing rate that may exceed the resources available on a single node and this condition is in general difficult to predict when dealing with online data streams [7]. On demand elastic computing platforms, such as Amazon AWS [8], and distributed processing frameworks, such as Apache Hadoop and Spark [9], have been developed as a technological solution for addressing these scalability issues. The attention of the research community has, accordingly, focused on processing functions [10, 11] and execution performance [12], giving less attention to other key features of information management, for example, reuse, reusability, and modularity.
Data and infrastructure management represent recurring challenges [5, 13] for Big Data. Due to the heterogeneous nature of distributed data, Big Data applications do not make use of traditional data modeling [14]. Distributed datasets and streams may consist of unstructured, semi-structured, or highly structured but still non-relational data items such as time series or nested records in which traditional data modeling techniques are prohibitive. Clearly, heterogeneous and/or weakly structured data make it difficult to design schemas in advance [15]. In addition, Big Data datasets may be processed only a few times or even once per use case, making it more expensive to load them into a database management system. In turn, data heterogeneity is related to the variety in data-intensive applications involving Cyber-Physical Systems (CPS) [16].

2 Related Work
johannes@inf.unica.it

¹ Università degli Studi di Milano, Milano, Italy
² Consorzio Inter-Ateneo Triestino, CTR, Udine, Italy
³ SAPIENZA University of Rome, Rome, Italy
⁴ University of Ferrara, Ferrara, Italy
⁵ EPFL, MARS Laboratory, Ave. Datan, CH-1200, Switzerland
⁶ Vienna University of Technology, WU, Austria
⁷ University of Pavia, Pavia, Italy
⁸ RWTH Aachen University, Aachen, Germany
⁹ Department for Manufacturing and Statistical Design, Milan National Institute of Statistics (Istat), Rome, Italy
¹⁰ Poznań University of Technology, Poznań, Poland
¹¹ TU Braunschweig, Braunschweig, Germany
¹² Aristotle University of Thessaloniki, Thessaloniki, Greece
¹³ Aristotle University of Thessaloniki, Thessaloniki, Greece
¹⁴ Aristotle University of Thessaloniki, Thessaloniki, Greece
¹⁵ Aristotle University of Thessaloniki, Thessaloniki, Greece



Panel Format

- Short intro by panelists (5min each)
- Q&A session

... and don't hesitate to contact Robert
(Robert.Wrembel@cs.put.poznan.pl)
if you are interested in joining 2.6!

<http://www.cs.put.poznan.pl/ifip-wg26/>

