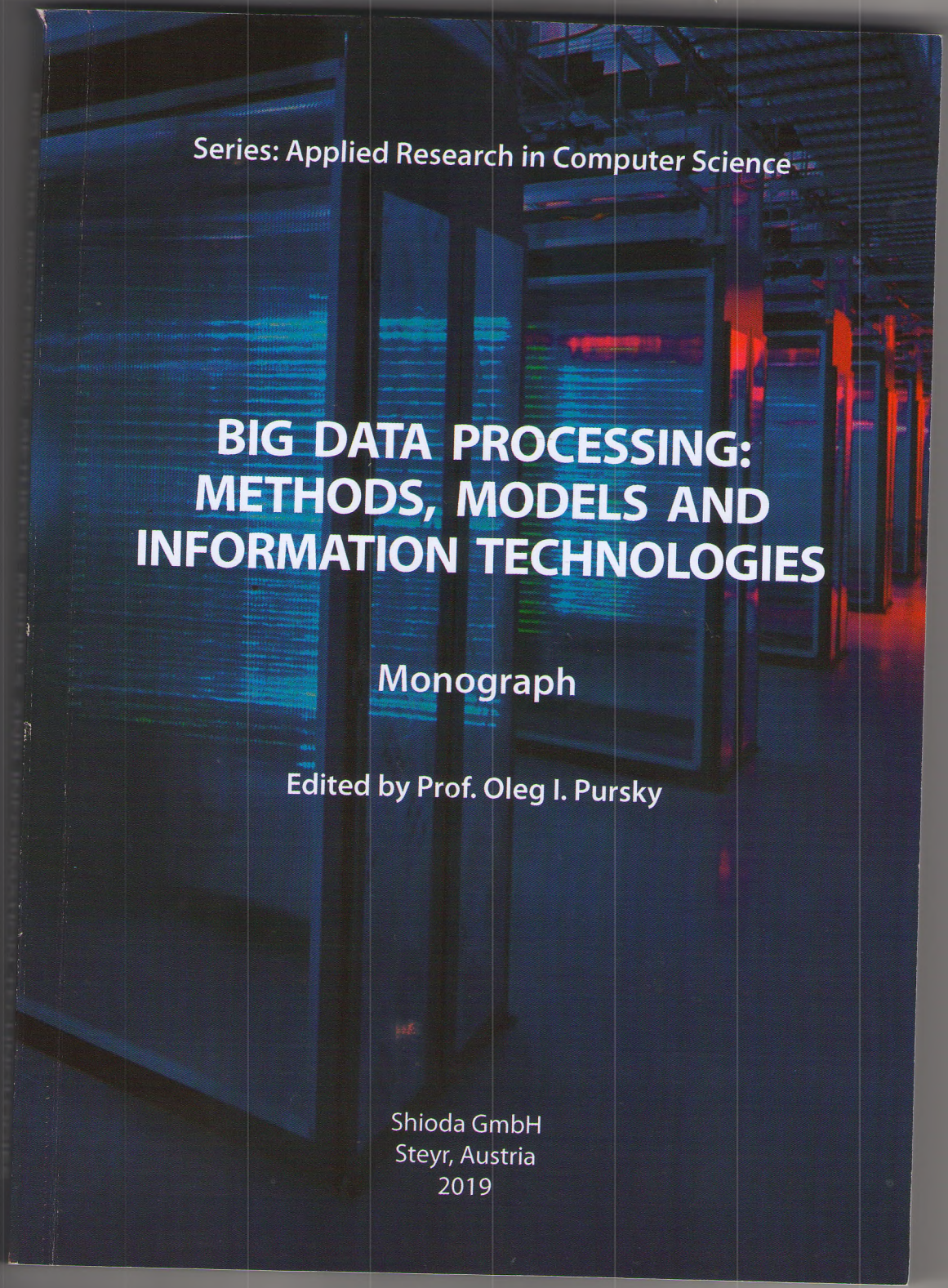

Series: Applied Research in Computer Science
BIG DATA PROCESSING: METHODS, MODELS AND INFORMATION TECHNOLOGIES
Monograph
Edited by Prof. Oleg I. Pursky
Shioda GmbH
Steyr, Austria
2019

Series: Applied Research in Computer Science

# BIG DATA PROCESSING: METHODS, MODELS AND INFORMATION TECHNOLOGIES

Monograph

Edited by Prof. Oleg I. Pursky

Shioda GmbH
Steyr, Austria
2019

Shioda GmbH, Steyr, Austria

# BIG DATA PROCESSING: METHODS, MODELS AND INFORMATION TECHNOLOGIES

Big Data processing: methods, models and information technologies: monograph. – edited by Oleg I. Pursky. – Shioda GmbH, Steyr, Austria, 2019. - 234 p.

ISBN 978-3-953794-29-8

The presented in monograph research results create a scientific basis for data analysis systems improvement and can be used in the big data processing in the all field of activity in terms of creating decision support systems and development of applied information technologies. Monograph is intended for professionals in the field of data analysis, decision support systems, information systems and technologies, as well as for lecturers, postgraduates and students.

ISBN 978-3-953794-29-8





## CONTENT

## INTRODUCTION

Modern world economic conditions, economy globalization, acceleration of market development processes, large volumes of information and promptness of their processing, different sociopolitical factors demand a new approaches to development of adequate methodological solutions and tools in the field of big data processing, especially it concerns a new information systems and technologies of data processing management. The main technologies of this type include modern technologies for collecting and processing big amounts datasets.

Big Data processing techniques is a direction of research that determines ways to analyze large data sets, systematically extract information from, or otherwise, deal with data sets that are too large or complex to be study with by traditional data processing application software. The concept of Big Data processing involves working with large volumes and varied data from different sources in order to improve the efficiency of management decisions and enhance competitiveness. Big Data processing are effective in the conditions of continuous growth, distribution in numerous nodes of the computer network that are alternative to traditional database management systems. An important problem at using technology Big Data is the interpretation of data. It is necessary to understand which channels of information gathering can be used, how to evaluate it and how to correct incorrect data.

The practical value of presented in monograph research is in findings and results that create a scientific basis for data analysis systems improvement and can be used in the big data processing in the all field of activity in terms of creating decision support systems and the development of applied information technologies.

The monograph is intended for professionals in the field of data analysis, decision support systems, information systems and technologies, as well as for lecturers, postgraduates and students.

# MODELING OF THE INVARIANT DATA BASE OF THE ENTERPRISE MANAGEMENT SYSTEM

**Pavlo Demidov[1], Kateryna Palahuta[2]**

*[1]Computer Science Department*
*Kyiv National University of Trade and Economics, Ukraine*
*[2]Department of Program Engineering and Cyber Security*
*Kyiv National University of Trade and Economics, Ukraine*
*Corresponding author E-mail: dempvg@meta.ua*

Functioning in modern conditions of the automated information management system of the enterprise (AISUP) requires the use of a powerful relational server database (DB), as the processes of manufacturing and marketing of finished products belong to the class of the most complex, in which tasks are solved from the provision of production by all kinds resources, its organizations to marketing and sales of finished products. And, as a result, they require a complex form of organization and solution of a large number of information tasks.

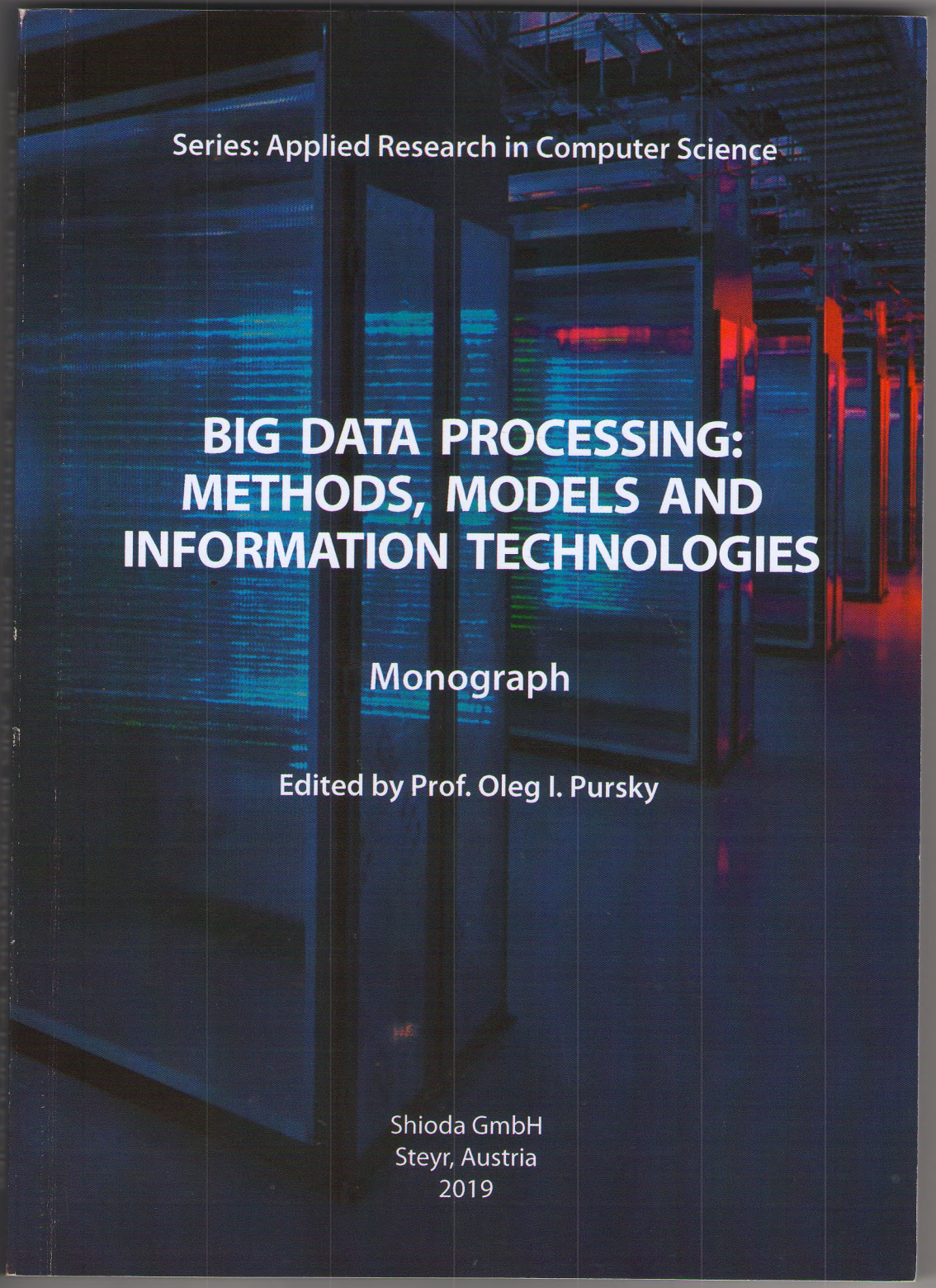
The architecture of the database, the composition of information objects, the links between them, the need for its expansion and development, its content and many other factors affect both the work of its own, and the entire information management system of the enterprise as a whole. The process of developing a database is iterative, branched and backward. At the same time, it is not possible to fully determine all requirements for the database at the initial stages of its development. They may be manifested, corrected and supplemented throughout the entire range from design to implementation, and in some cases during implementation. Refactoring, optimization and reengineering of the software code, as well as direct and reverse engineering of the database are necessary procedures that the developer has to do to improve the efficiency of the database and facilitate the understanding of the work of the programs.

The stages of designing data models require the definition of entities, keys, attributes, the normalization of relationships, the definition of relationships



between entities. An important step is also the choice of Case-systems for designing models, DBMS for the development and maintenance of databases, their configuration and combination into a single functioning system. Formation on the basis of the physical model of SQL-description allows you to automatically create information objects in the database of the chosen DBMS. Thus, the design and development of a real database for a production enterprise requires the implementation of a huge number of operations of various types.

The design and development of a relational database is devoted to a huge amount of research, which are presented in the form of scientific articles, monographs, manuals and other materials. In classical works [1-6], the following questions are usually considered: basic concepts, database architecture, data models, relational algebra, SQL language, database design (functional dependencies, normalization of relationships), data protection, distributed databases, object-relational and deductive DBMS. In the works over the past 5-7 years, the number of publications related to the work of databases on the Internet has increased, with the development of web application access to databases using object-oriented programming languages: PHP, Java and others. The summary of all this material is based on examples of small databases that are limited to two or three tables and links between them. This is correct in order to understand the above material [7, 8]. But, in this case, "behind the ship" there are problems with the development of an industrial database, the structure and content of which takes into account the experience of implementing ERP-system on dozens of objects of different types and sizes. The functional decomposition of the IP of an industrial enterprise, depending on the complexity of a particular enterprise, varies from 10 to 50 names of functional subsystems. The current ERP system of the enterprise must have such a powerful database and number of functional subsystems so that it can adapt to the work of the enterprise of different type and scale. This is also required for the development of medium



Series: Applied Research in Computer Science

# BIG DATA PROCESSING: METHODS, MODELS AND INFORMATION TECHNOLOGIES

Monograph

Edited by Prof. Oleg I. Pursky

Shioda GmbH
Steyr, Austria
2019