
Read Book Pdf Springer Pig Apache Beginning

Thank you enormously much for downloading **Pdf Springer Pig Apache Beginning**. Maybe you have knowledge that, people have look numerous times for their favorite books afterward this Pdf Springer Pig Apache Beginning, but stop going on in harmful downloads.

Rather than enjoying a fine PDF past a cup of coffee in the afternoon, instead they juggled taking into account some harmful virus inside their computer. **Pdf Springer Pig Apache Beginning** is within reach in our digital library an online permission to it is set as public hence you can download it instantly. Our digital library saves in multiple countries, allowing you to get the most less latency epoch to download any of our books considering this one. Merely said, the Pdf Springer Pig Apache Beginning is universally compatible bearing in mind any devices to read.

KEY=PIG - CRISTINA POWELL

Beginning Apache Pig Big Data Processing Made Easy Apress *Learn to use Apache Pig to develop lightweight big data applications easily and quickly. This book shows you many optimization techniques and covers every context where Pig is used in big data analytics. Beginning Apache Pig shows you how Pig is easy to learn and requires relatively little time to develop big data applications. The book is divided into four parts: the complete features of Apache Pig; integration with other tools; how to solve complex business problems; and optimization of tools. You'll discover topics such as MapReduce and why it cannot meet every business need; the features of Pig Latin such as data types for each load, store, joins, groups, and ordering; how Pig workflows can be created; submitting Pig jobs using Hue; and working with Oozie. You'll also see how to extend the framework by writing UDFs and custom load, store, and filter functions. Finally you'll cover different optimization techniques such as gathering statistics about a Pig script, joining strategies, parallelism, and the role of data formats in good performance. What You Will Learn*

- Use all the features of Apache Pig
- Integrate Apache Pig with other tools
- Extend Apache Pig
- Optimize Pig Latin code
- Solve different use cases for Pig Latin

Who This Book Is For All levels of IT professionals: architects, big data enthusiasts, engineers, developers, and big data administrators

Middleware 2013 ACM/IFIP/USENIX 14th International Middleware Conference, Beijing, China, December 9-13, 2013, Proceedings Springer *This book constitutes the refereed proceedings of the ACM/IFIP/USENIX 14th International Middleware Conference, held in Beijing, China, in December 2013. The 24 revised full papers presented were carefully reviewed and*

selected from 189 submissions. The papers cover a wide range of topics including design, implementation, deployment and evaluation of middleware for next-generation platforms such as cloud computing, social networks and large-scale storage and distributed systems. The middleware solutions introduced provide features such as availability, efficiency, scalability, fault-tolerance, trustworthy operation and support security and privacy needs. **Handbook of Systems Engineering and Risk Management in Control Systems, Communication, Space Technology, Missile, Security and Defense Operations CRC Press** This book provides multifaceted components and full practical perspectives of systems engineering and risk management in security and defense operations with a focus on infrastructure and manpower control systems, missile design, space technology, satellites, intercontinental ballistic missiles, and space security. While there are many existing selections of systems engineering and risk management textbooks, there is no existing work that connects systems engineering and risk management concepts to solidify its usability in the entire security and defense actions. With this book Dr. Anna M. Doro-on rectifies the current imbalance. She provides a comprehensive overview of systems engineering and risk management before moving to deeper practical engineering principles integrated with newly developed concepts and examples based on industry and government methodologies. The chapters also cover related points including design principles for defeating and deactivating improvised explosive devices and land mines and security measures against kinds of threats. The book is designed for systems engineers in practice, political risk professionals, managers, policy makers, engineers in other engineering fields, scientists, decision makers in industry and government and to serve as a reference work in systems engineering and risk management courses with focus on security and defense operations. **Software Engineering Frameworks for the Cloud Computing Paradigm Springer Science & Business Media** This book presents the latest research on Software Engineering Frameworks for the Cloud Computing Paradigm, drawn from an international selection of researchers and practitioners. The book offers both a discussion of relevant software engineering approaches and practical guidance on enterprise-wide software deployment in the cloud environment, together with real-world case studies. Features: presents the state of the art in software engineering approaches for developing cloud-suitable applications; discusses the impact of the cloud computing paradigm on software engineering; offers guidance and best practices for students and practitioners; examines the stages of the software development lifecycle, with a focus on the requirements engineering and testing of cloud-based applications; reviews the efficiency and performance of cloud-based applications; explores feature-driven and cloud-aided software design; provides relevant theoretical frameworks, practical approaches and future research directions. **Processing Big Data with Azure HDInsight Building Real-World Big Data Systems on Azure HDInsight Using the Hadoop Ecosystem Apress** Get a jump start on using Azure HDInsight and Hadoop Ecosystem components. As most Hadoop and Big Data projects are written in either Java, Scala, or Python, this book minimizes the effort to learn another language and is written from the perspective of a .NET developer. Hadoop components are covered, including Hive, Pig, HBase, Storm, and Spark on Azure HDInsight, and code samples are written in .NET only. Processing Big

Data with Azure HDInsight covers the fundamentals of big data, how businesses are using it to their advantage, and how Azure HDInsight fits into the big data world. This book introduces Hadoop and big data concepts and then dives into creating different solutions with HDInsight and the Hadoop Ecosystem. It covers concepts with real-world scenarios and code examples, making sure you get hands-on experience. The best way to utilize this book is to practice while reading. After reading this book you will be familiar with Azure HDInsight and how it can be utilized to build big data solutions, including batch processing, stream analytics, interactive processing, and storing and retrieving data in an efficient manner. What You'll Learn Understand the fundamentals of HDInsight and Hadoop Work with HDInsight cluster Query with Apache Hive and Apache Pig Store and retrieve data with Apache HBase Stream data processing using Apache Storm Work with Apache Spark Who This Book Is For Software developers, technical architects, data scientists/analysts, and Hadoop administrators who want to develop on Microsoft's managed Hadoop offering, HDInsight

Network Data Analytics A Hands-On Approach for Application Development Springer *In order to carry out data analytics, we need powerful and flexible computing software. However the software available for data analytics is often proprietary and can be expensive. This book reviews Apache tools, which are open source and easy to use. After providing an overview of the background of data analytics, covering the different types of analysis and the basics of using Hadoop as a tool, it focuses on different Hadoop ecosystem tools, like Apache Flume, Apache Spark, Apache Storm, Apache Hive, R, and Python, which can be used for different types of analysis. It then examines the different machine learning techniques that are useful for data analytics, and how to visualize data with different graphs and charts. Presenting data analytics from a practice-oriented viewpoint, the book discusses useful tools and approaches for data analytics, supported by concrete code examples. The book is a valuable reference resource for graduate students and professionals in related fields, and is also of interest to general readers with an understanding of data analytics.*

Proceedings of the International Congress on Information and Communication Technology ICICT 2015, Volume 1 Springer *This volume contains 69 papers presented at ICICT 2015: International Congress on Information and Communication Technology. The conference was held during 9th and 10th October, 2015, Udaipur, India and organized by CSI Udaipur Chapter, Division IV, SIG-WNS, SIG-e-Agriculture in association with ACM Udaipur Professional Chapter, The Institution of Engineers (India), Udaipur Local Centre and Mining Engineers Association of India, Rajasthan Udaipur Chapter. This volume contains papers mainly focused on ICT for Managerial Applications, E-governance, IOT and E-Mining.*

Beginning Apache Spark 3 With DataFrame, Spark SQL, Structured Streaming, and Spark Machine Learning Library Apress *Take a journey toward discovering, learning, and using Apache Spark 3.0. In this book, you will gain expertise on the powerful and efficient distributed data processing engine inside of Apache Spark; its user-friendly, comprehensive, and flexible programming model for processing data in batch and streaming; and the scalable machine learning algorithms and practical utilities to build machine learning applications. Beginning Apache Spark 3 begins by explaining different ways of interacting with Apache Spark, such as Spark Concepts and Architecture, and Spark Unified Stack. Next, it offers an overview of*

Spark SQL before moving on to its advanced features. It covers tips and techniques for dealing with performance issues, followed by an overview of the structured streaming processing engine. It concludes with a demonstration of how to develop machine learning applications using Spark MLlib and how to manage the machine learning development lifecycle. This book is packed with practical examples and code snippets to help you master concepts and features immediately after they are covered in each section. After reading this book, you will have the knowledge required to build your own big data pipelines, applications, and machine learning applications. What You Will Learn Master the Spark unified data analytics engine and its various components Work in tandem to provide a scalable, fault tolerant and performant data processing engine Leverage the user-friendly and flexible programming model to perform simple to complex data analytics using dataframe and Spark SQL Develop machine learning applications using Spark MLlib Manage the machine learning development lifecycle using MLflow Who This Book Is For Data scientists, data engineers and software developers. **Information Fusion for Cyber-Security Analytics Springer** This book highlights several gaps that have not been addressed in existing cyber security research. It first discusses the recent attack prediction techniques that utilize one or more aspects of information to create attack prediction models. The second part is dedicated to new trends on information fusion and their applicability to cyber security; in particular, graph data analytics for cyber security, unwanted traffic detection and control based on trust management software defined networks, security in wireless sensor networks & their applications, and emerging trends in security system design using the concept of social behavioral biometric. The book guides the design of new commercialized tools that can be introduced to improve the accuracy of existing attack prediction models. Furthermore, the book advances the use of Knowledge-based Intrusion Detection Systems (IDS) to complement existing IDS technologies. It is aimed towards cyber security researchers. **Proceedings of the International Conference on Business and Management Dynamics 2016: Sustainable economies in the information economy AOSIS** Cape Town, South Africa, 7 Sept. 2016 – 8 Sept. 2016. Theme: Sustainable economies in the information economy. Purpose: To share the quality academic papers presented at the International Conference on Business and Management Dynamics (ICBMD) held from 7 to 8 September 2016 at African Pride Crystal Hotel and Spa in Cape Town. As grey literature, the proceedings are the contributions made by researchers at the conference and are considered the written record of the work that was presented to fellow conference delegates. Methodology: The methodology used varies from researcher to researcher but are suitable for the studies conducted. Thus, on the one hand, studies that were subjective in nature used the interpretive paradigm, where the qualitative approach adopted made use of the interview method to collect data. On the other hand, studies that were objectively inclined adopted the positivist philosophy and used survey questionnaires to collect data. However, there were some academic papers which used mixed methodology because of the nature of the study. Whatever methodology used adhered to the ethos of the philosophies underpinning the methodology. Contribution made to scholarship: The articles come from individual researchers and each article in the proceedings is unique. Mostly, there is no general argument leading from one

contribution to the next. However, it is interesting to note that in the area of economic performance it was evident that real exchange rate and net foreign direct investment contribute more towards innovations in economic growth. With regard to human capital development, papers presented evidence that there exists a definite need to explore the phenomenon of personal branding as limited scientific academic research has been done within the field of personal branding or on elements of the topic. Thus, the outcome argues that personal branding has an influence on leadership style which in turn impacts on organisational performance and related hygiene factors. Furthermore, it was demonstrated that current methods or strategies for enforcing institutionalisation of knowledge sharing within an organisation have not been successful, and, as such, new strategies are needed to reinforce efforts to nurture and invigorate the institutionalisation of knowledge sharing within an organisation. With regard to technology and big data impact on organisational performance, it was evident that system performance, memory consumption and CPU utilisation can be used as criteria to compare and evaluate big data technologies to improve organisational performance. Most of the articles' contribution reemphasised technology education and training as a means of digitising business and improving effectiveness. Target audience: The target readership is academic researchers and business leaders who require access to the latest developments in the fields of economics, information management, business, education, development studies, social sciences and technology. It is also for policymakers and other stakeholders who need a better understanding of the impact of new developments on existing policies and regulations for their review or amendment.

Data Science and Big Data Computing Frameworks and Methodologies Springer This illuminating text/reference surveys the state of the art in data science, and provides practical guidance on big data analytics. Expert perspectives are provided by authoritative researchers and practitioners from around the world, discussing research developments and emerging trends, presenting case studies on helpful frameworks and innovative methodologies, and suggesting best practices for efficient and effective data analytics. Features: reviews a framework for fast data applications, a technique for complex event processing, and agglomerative approaches for the partitioning of networks; introduces a unified approach to data modeling and management, and a distributed computing perspective on interfacing physical and cyber worlds; presents techniques for machine learning for big data, and identifying duplicate records in data repositories; examines enabling technologies and tools for data mining; proposes frameworks for data extraction, and adaptive decision making and social media analysis.

New Horizons for a Data-Driven Economy A Roadmap for Usage and Exploitation of Big Data in Europe Springer In this book readers will find technological discussions on the existing and emerging technologies across the different stages of the big data value chain. They will learn about legal aspects of big data, the social impact, and about education needs and requirements. And they will discover the business perspective and how big data technology can be exploited to deliver value within different sectors of the economy. The book is structured in four parts: Part I "The Big Data Opportunity" explores the value potential of big data with a particular focus on the European context. It also describes the legal, business and social dimensions that need to be addressed, and briefly introduces the

European Commission's BIG project. Part II "The Big Data Value Chain" details the complete big data lifecycle from a technical point of view, ranging from data acquisition, analysis, curation and storage, to data usage and exploitation. Next, Part III "Usage and Exploitation of Big Data" illustrates the value creation possibilities of big data applications in various sectors, including industry, healthcare, finance, energy, media and public services. Finally, Part IV "A Roadmap for Big Data Research" identifies and prioritizes the cross-sectorial requirements for big data research, and outlines the most urgent and challenging technological, economic, political and societal issues for big data in Europe. This compendium summarizes more than two years of work performed by a leading group of major European research centers and industries in the context of the BIG project. It brings together research findings, forecasts and estimates related to this challenging technological context that is becoming the major axis of the new digitally transformed business environment.

Big Data Technologies and Applications Springer The objective of this book is to introduce the basic concepts of big data computing and then to describe the total solution of big data problems using HPCC, an open-source computing platform. The book comprises 15 chapters broken into three parts. The first part, Big Data Technologies, includes introductions to big data concepts and techniques; big data analytics; and visualization and learning techniques. The second part, LexisNexis Risk Solution to Big Data, focuses on specific technologies and techniques developed at LexisNexis to solve critical problems that use big data analytics. It covers the open source High Performance Computing Cluster (HPCC Systems®) platform and its architecture, as well as parallel data languages ECL and KEL, developed to effectively solve big data problems. The third part, Big Data Applications, describes various data intensive applications solved on HPCC Systems. It includes applications such as cyber security, social network analytics including fraud, Ebola spread modeling using big data analytics, unsupervised learning, and image classification. The book is intended for a wide variety of people including researchers, scientists, programmers, engineers, designers, developers, educators, and students. This book can also be beneficial for business managers, entrepreneurs, and investors.

Big Data: Concepts, Methodologies, Tools, and Applications Concepts, Methodologies, Tools, and Applications IGI Global The digital age has presented an exponential growth in the amount of data available to individuals looking to draw conclusions based on given or collected information across industries. Challenges associated with the analysis, security, sharing, storage, and visualization of large and complex data sets continue to plague data scientists and analysts alike as traditional data processing applications struggle to adequately manage big data. Big Data: Concepts, Methodologies, Tools, and Applications is a multi-volume compendium of research-based perspectives and solutions within the realm of large-scale and complex data sets. Taking a multidisciplinary approach, this publication presents exhaustive coverage of crucial topics in the field of big data including diverse applications, storage solutions, analysis techniques, and methods for searching and transferring large data sets, in addition to security issues. Emphasizing essential research in the field of data science, this publication is an ideal reference source for data analysts, IT professionals, researchers, and academics.

Guide to Big Data Applications Springer This handbook brings together a variety of approaches to the uses of big data in multiple fields,

primarily science, medicine, and business. This single resource features contributions from researchers around the world from a variety of fields, where they share their findings and experience. This book is intended to help spur further innovation in big data. The research is presented in a way that allows readers, regardless of their field of study, to learn from how applications have proven successful and how similar applications could be used in their own field. Contributions stem from researchers in fields such as physics, biology, energy, healthcare, and business. The contributors also discuss important topics such as fraud detection, privacy implications, legal perspectives, and ethical handling of big data. **Internet of Things Converging Technologies for Smart Environments and Integrated Ecosystems River Publishers** The book aims to provide a broad overview of various topics of the Internet of Things (IoT) from the research and development priorities to enabling technologies, architecture, security, privacy, interoperability and industrial applications. It is intended to be a stand-alone book in a series that covers the Internet of Things activities of the IERC - Internet of Things European Research Cluster - from technology to international cooperation and the global "state of play." The book builds on the ideas put forward by the European Research Cluster on the Internet of Things Strategic Research and Innovation Agenda and presents views and state of the art results on the challenges facing the research, development and deployment of IoT at the global level. Today we see the integration of Industrial, Business and Consumer Internet which is bringing together the Internet of People, Internet of Things, Internet of Energy, Internet of Vehicles, Internet of Media, Services and Enterprises in forming the backbone of the digital economy, the digital society and the foundation for the future knowledge and innovation based economy. These developments are supporting solutions for the emerging challenges of public health, aging population, environmental protection and climate change, the conservation of energy and scarce materials, enhancements to safety and security and the continuation and growth of economic prosperity. Penetration of smartphones and advances in nanoelectronics, cyber-physical systems, wireless communication, software, and Cloud computing technology will be the main drivers for IoT development. The IoT contribution is seen in the increased value of information created by the number of interconnections among things and the transformation of the processed information into knowledge shared into the Internet of Everything. The connected devices are part of ecosystems connecting people, processes, data, and things which are communicating in the Cloud using the increased storage and computing power while attempting to standardize communication and metadata. In this context, the next generation of Cloud computing technologies will need to be flexible enough to scale autonomously, adaptive enough to handle constantly changing connections and resilient enough to stand up to the huge flows of data that will occur. In 2025, analysts forecast that there will be six devices per human on the planet, which means around 50 billion more connected devices over the next 12 years. The Internet of Things market is connected to this anticipated device growth from industrial Machine to Machine (M2M) systems, smart meters and wireless sensors. Internet of Things technology will generate new services and new interfaces by creating smart environments and smart spaces with applications ranging from Smart Cities, Smart Transport, Buildings, Energy, Grid, to Smart Health and Life. **The Data Science**

Framework A View from the EDISON Project Springer Nature *This edited book first consolidates the results of the EU-funded EDISON project (Education for Data Intensive Science to Open New science frontiers), which developed training material and information to assist educators, trainers, employers, and research infrastructure managers in identifying, recruiting and inspiring the data science professionals of the future. It then deepens the presentation of the information and knowledge gained to allow for easier assimilation by the reader. The contributed chapters are presented in sequence, each chapter picking up from the end point of the previous one. After the initial book and project overview, the chapters present the relevant data science competencies and body of knowledge, the model curriculum required to teach the required foundations, profiles of professionals in this domain, and use cases and applications. The text is supported with appendices on related process models. The book can be used to develop new courses in data science, evaluate existing modules and courses, draft job descriptions, and plan and design efficient data-intensive research teams across scientific disciplines.*

Beginning Apache Spark 2 With Resilient Distributed Datasets, Spark SQL, Structured Streaming and Spark Machine Learning library Apress *Develop applications for the big data landscape with Spark and Hadoop. This book also explains the role of Spark in developing scalable machine learning and analytics applications with Cloud technologies. Beginning Apache Spark 2 gives you an introduction to Apache Spark and shows you how to work with it. Along the way, you'll discover resilient distributed datasets (RDDs); use Spark SQL for structured data; and learn stream processing and build real-time applications with Spark Structured Streaming. Furthermore, you'll learn the fundamentals of Spark ML for machine learning and much more. After you read this book, you will have the fundamentals to become proficient in using Apache Spark and know when and how to apply it to your big data applications.*

What You Will Learn *Understand Spark unified data processing platform How to run Spark in Spark Shell or Databricks Use and manipulate RDDs Deal with structured data using Spark SQL through its operations and advanced functions Build real-time applications using Spark Structured Streaming Develop intelligent applications with the Spark Machine Learning library Who This Book Is For* *Programmers and developers active in big data, Hadoop, and Java but who are new to the Apache Spark platform.*

Practical Hadoop Migration How to Integrate Your RDBMS with the Hadoop Ecosystem and Re-Architect Relational Applications to NoSQL Apress *Re-architect relational applications to NoSQL, integrate relational database management systems with the Hadoop ecosystem, and transform and migrate relational data to and from Hadoop components. This book covers the best-practice design approaches to re-architecting your relational applications and transforming your relational data to optimize concurrency, security, denormalization, and performance. Winner of IBM's 2012 Gerstner Award for his implementation of big data and data warehouse initiatives and author of Practical Hadoop Security, author Bhushan Lakhe walks you through the entire transition process. First, he lays out the criteria for deciding what blend of re-architecting, migration, and integration between RDBMS and HDFS best meets your transition objectives. Then he demonstrates how to design your transition model. Lakhe proceeds to cover the selection criteria for ETL tools, the implementation steps for migration with SQOOP- and Flume-based data transfers, and transition*

optimization techniques for tuning partitions, scheduling aggregations, and redesigning ETL. Finally, he assesses the pros and cons of data lakes and Lambda architecture as integrative solutions and illustrates their implementation with real-world case studies. Hadoop/NoSQL solutions do not offer by default certain relational technology features such as role-based access control, locking for concurrent updates, and various tools for measuring and enhancing performance. *Practical Hadoop Migration* shows how to use open-source tools to emulate such relational functionalities in Hadoop ecosystem components. *What You'll Learn* Decide whether you should migrate your relational applications to big data technologies or integrate them Transition your relational applications to Hadoop/NoSQL platforms in terms of logical design and physical implementation Discover RDBMS-to-HDFS integration, data transformation, and optimization techniques Consider when to use Lambda architecture and data lake solutions Select and implement Hadoop-based components and applications to speed transition, optimize integrated performance, and emulate relational functionalities *Who This Book Is For* Database developers, database administrators, enterprise architects, Hadoop/NoSQL developers, and IT leaders. Its secondary readership is project and program managers and advanced students of database and management information systems. **Specifying Big Data Benchmarks First Workshop, WBDB 2012, San Jose, CA, USA, May 8-9, 2012 and Second Workshop, WBDB 2012, Pune, India, December 17-18, 2012, Revised Selected Papers Springer** This book constitutes the thoroughly refereed revised selected papers of the First Workshop on Big Data Benchmarks, WBDB 2012, held in San Jose, CA, USA, in May 2012 and the Second Workshop on Big Data Benchmarks, WBDB 2012, held in Pune, India, in December 2012. The 14 revised papers presented were carefully reviewed and selected from 60 submissions. The papers are organized in topical sections on benchmarking, foundations and tools; domain specific benchmarking; benchmarking hardware and end-to-end big data benchmarks. **Networked Digital Technologies 4th International Conference, NDT 2012, Dubai, UAE, April 24-26, 2012. Proceedings, Part I Springer** This two-volume-set (CCIS 293 and CCIS 294) constitutes the refereed proceedings of the International Conference on Networked Digital Technologies, NDT 2012, held in Dubai, UAE, in April 2012. The 96 papers presented in the two volumes were carefully reviewed and selected from 228 submissions. The papers are organized in topical sections on collaborative systems for e-sciences; context-aware processing and ubiquitous systems; data and network mining; grid and cloud computing; information and data management; intelligent agent-based systems; internet modeling and design; mobile, ad hoc and sensor network management; peer-to-peer social networks; quality of service for networked systems; semantic Web and ontologies; security and access control; signal processing and computer vision for networked systems; social networks; Web services. **High Performance Computing 7th Latin American Conference, CARLA 2020, Cuenca, Ecuador, September 2-4, 2020, Revised Selected Papers Springer Nature** This book constitutes revised selected papers of the 7th Latin American High Performance Computing Conference, CARLA 2020, held in Cuenca, Ecuador, in September 2020. Due to the COVID-19 pandemic the conference was held in a virtual mode. The 15 revised full papers presented were carefully reviewed and selected out of 36 submissions. The

papers included in this book are organized according to the topics on High Performance Computing Applications; High Performance Computing and Artificial Intelligence. **Handbook of Cloud Computing Springer Science & Business Media** Cloud computing has become a significant technology trend. Experts believe cloud computing is currently reshaping information technology and the IT marketplace. The advantages of using cloud computing include cost savings, speed to market, access to greater computing resources, high availability, and scalability. Handbook of Cloud Computing includes contributions from world experts in the field of cloud computing from academia, research laboratories and private industry. This book presents the systems, tools, and services of the leading providers of cloud computing; including Google, Yahoo, Amazon, IBM, and Microsoft. The basic concepts of cloud computing and cloud computing applications are also introduced. Current and future technologies applied in cloud computing are also discussed. Case studies, examples, and exercises are provided throughout. Handbook of Cloud Computing is intended for advanced-level students and researchers in computer science and electrical engineering as a reference book. This handbook is also beneficial to computer and system infrastructure designers, developers, business managers, entrepreneurs and investors within the cloud computing related industry. **Information Granularity, Big Data, and Computational Intelligence Springer** The recent pursuits emerging in the realm of big data processing, interpretation, collection and organization have emerged in numerous sectors including business, industry and government organizations. Data sets such as customer transactions for a mega-retailer, weather monitoring, intelligence gathering, quickly outpace the capacities of traditional techniques and tools of data analysis. The 3V (volume, variability and velocity) challenges led to the emergence of new techniques and tools in data visualization, acquisition, and serialization. Soft Computing being regarded as a plethora of technologies of fuzzy sets (or Granular Computing), neurocomputing and evolutionary optimization brings forward a number of unique features that might be instrumental to the development of concepts and algorithms to deal with big data. This carefully edited volume provides the reader with an updated, in-depth material on the emerging principles, conceptual underpinnings, algorithms and practice of Computational Intelligence in the realization of concepts and implementation of big data architectures, analysis, and interpretation as well as data analytics. The book is aimed at a broad audience of researchers and practitioners including those active in various disciplines in which big data, their analysis and optimization are of genuine relevance. One focal point is the systematic exposure of the concepts, design methodology, and detailed algorithms. In general, the volume adheres to the top-down strategy starting with the concepts and motivation and then proceeding with the detailed design that materializes in specific algorithms and representative applications. The material is self-contained and provides the reader with all necessary prerequisites and augments some parts with a step-by-step explanation of more advanced concepts supported by a significant amount of illustrative numeric material and some application scenarios to motivate the reader and make some abstract concepts more tangible. **Engineering Vibration, Communication and Information Processing Icoevci 2018, India Lecture Notes in Electrical En Intelligent Internet of Things From Device to Fog and Cloud Springer Nature** This holistic book is an

invaluable reference for addressing various practical challenges in architecting and engineering Intelligent IoT and eHealth solutions for industry practitioners, academic and researchers, as well as for engineers involved in product development. The first part provides a comprehensive guide to fundamentals, applications, challenges, technical and economic benefits, and promises of the Internet of Things using examples of real-world applications. It also addresses all important aspects of designing and engineering cutting-edge IoT solutions using a cross-layer approach from device to fog, and cloud covering standards, protocols, design principles, reference architectures, as well as all the underlying technologies, pillars, and components such as embedded systems, network, cloud computing, data storage, data processing, big data analytics, machine learning, distributed ledger technologies, and security. In addition, it discusses the effects of Intelligent IoT, which are reflected in new business models and digital transformation. The second part provides an insightful guide to the design and deployment of IoT solutions for smart healthcare as one of the most important applications of IoT. Therefore, the second part targets smart healthcare-wearable sensors, body area sensors, advanced pervasive healthcare systems, and big data analytics that are aimed at providing connected health interventions to individuals for healthier lifestyles.

Computational Methods and Data Engineering Proceedings of ICMDE 2020, Volume 2 Springer Nature *This book gathers selected high-quality research papers from the International Conference on Computational Methods and Data Engineering (ICMDE 2020), held at SRM University, Sonipat, Delhi-NCR, India. Focusing on cutting-edge technologies and the most dynamic areas of computational intelligence and data engineering, the respective contributions address topics including collective intelligence, intelligent transportation systems, fuzzy systems, data privacy and security, data mining, data warehousing, big data analytics, cloud computing, natural language processing, swarm intelligence, and speech processing.*

Introduction to Information Retrieval Cambridge University Press *Class-tested and coherent, this textbook teaches classical and web information retrieval, including web search and the related areas of text classification and text clustering from basic concepts. It gives an up-to-date treatment of all aspects of the design and implementation of systems for gathering, indexing, and searching documents; methods for evaluating systems; and an introduction to the use of machine learning methods on text collections. All the important ideas are explained using examples and figures, making it perfect for introductory courses in information retrieval for advanced undergraduates and graduate students in computer science. Based on feedback from extensive classroom experience, the book has been carefully structured in order to make teaching more natural and effective. Slides and additional exercises (with solutions for lecturers) are also available through the book's supporting website to help course instructors prepare their lectures.*

Open Learning Cultures A Guide to Quality, Evaluation, and Assessment for Future Learning Springer Science & Business Media *Today we are seeing a new form of blended learning: not only is technology enhancing the learning environment but formal and informal learning are combining and there is self- and peer-assessment of results. Open learning cultures are challenging the old and long-practiced methods used by educators and transforming learning into a more student-driven and independent activity, which uses online tools such as blogs, wikis*

or podcasts to connect resources, students and teachers in a novel way. While in higher education institutions most assessments are still tied to formal learning scenarios, teachers are more and more bound to recognize their students' informal learning processes and networks. This book will help teachers, lecturers and students to better understand how open learning landscapes work, how to define quality and create assessments in such environments, and how to apply these new measures. To this end, Ehlers first elaborates the technological background for more collaborative, distributed, informal, and self-guided learning. He covers the rise of social media for learning and shows how an architecture of participation can change learning activities. These new paradigms are then applied to learning and education to outline what open learning landscapes look like. Here he highlights the shift from knowledge transfer to competence development, the increase in lifelong learning, and the importance of informal learning, user generated content, and open educational resources. He then shows how to manage quality by presenting a step by step guide to developing customized quality concepts for open learning landscapes. Finally, several methods dealing with assessment in these new environments are presented, including guidelines, templates and use cases to exemplify the approaches. Overall, Ehlers argues for assessment as an integral part of learning processes, with quality assurance as a method of stimulating a quality culture and continuous quality development rather than as a simple controlling exercise.

Biomedical Applications Based on Natural and Artificial Computing International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2017, Corunna, Spain, June 19-23, 2017, Proceedings, Part II Springer The two volumes LNCS 10337 and 10338 constitute the proceedings of the International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2017, held in Corunna, Spain, in June 2017. The total of 102 full papers was carefully reviewed and selected from 194 submissions during two rounds of reviewing and improvement. The papers are organized in two volumes, one on natural and artificial computation for biomedicine and neuroscience, addressing topics such as theoretical neural computation; models; natural computing in bioinformatics; physiological computing in affective smart environments; emotions; as well as signal processing and machine learning applied to biomedical and neuroscience applications. The second volume deals with biomedical applications, based on natural and artificial computing and addresses topics such as biomedical applications; mobile brain computer interaction; human robot interaction; deep learning; machine learning applied to big data analysis; computational intelligence in data coding and transmission; and applications.

Tools and Algorithms for the Construction and Analysis of Systems 22nd International Conference, TACAS 2016, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2016, Eindhoven, The Netherlands, April 2-8, 2016, Proceedings Springer This book constitutes the proceedings of the 22nd International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2016, which took place in Eindhoven, The Netherlands, in April 2016, held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2016. The 44 full papers presented in this volume were carefully reviewed and selected from 175 submissions. They were organized in topical sections named: abstraction and

verification; probabilistic and stochastic systems; synthesis; tool papers; concurrency; tool demos; languages and automata; security; optimization; and competition on software verification - SV-COMP. **Current Trends in Web Engineering ICWE 2013 International Workshops ComposableWeb, QWE, MDWE, DMSSW, EMotions, CSE, SSN, and PhD Symposium, Aalborg, Denmark, July 8-12, 2013. Revised Selected Papers Springer** This book constitutes the thoroughly refereed post-proceedings of the seven workshops and the PhD Symposium that were co-located with the 13th International Conference on Web Engineering, ICWE 2013, held in Aalborg, Denmark, in July 2013. The papers cover research in topics such as social data management; cloud service engineering; agile web development and quality management in web engineering. **Intelligent Tools for Building a Scientific Information Platform: From Research to Implementation Springer** This book is a selection of results obtained within three years of research performed under SYNAT—a nation-wide scientific project aiming at creating an infrastructure for scientific content storage and sharing for academia, education and open knowledge society in Poland. The book is intended to be the last of the series related to the SYNAT project. The previous books, titled “Intelligent Tools for Building a Scientific Information Platform” and “Intelligent Tools for Building a Scientific Information Platform: Advanced Architectures and Solutions”, were published as volumes 390 and 467 in Springer's Studies in Computational Intelligence. Its contents is based on the SYNAT 2013 Workshop held in Warsaw. The papers included in this volume present an overview and insight into information retrieval, repository systems, text processing, ontology-based systems, text mining, multimedia data processing and advanced software engineering, addressing the problems of implementing intelligent tools for building a scientific information platform. **New Frontiers in Information and Software as Services Service and Application Design Challenges in the Cloud Springer Science & Business Media** The increasing costs of creating and maintaining infrastructures for delivering services to consumers have led to the emergence of cloud based third party service providers renting networks, computation power, storage, and even entire software application suites. On the other hand, service customers demand competitive pricing, service level agreements, and increased flexibility and scalability. Service consumers also expect process and data security, 24/7 service availability, and compliance with privacy regulations. This book focuses on such challenges associated with the design, implementation, deployment, and management of data and software as a service. The 12 papers presented in this volume were contributed by leaders in academia and industry, and were reviewed and supervised by an expert editorial board. They describe cutting-edge approaches in areas like service design, service security, service optimization, and service migration. **Essentials of Business Analytics An Introduction to the Methodology and its Applications Springer** This comprehensive edited volume is the first of its kind, designed to serve as a textbook for long-duration business analytics programs. It can also be used as a guide to the field by practitioners. The book has contributions from experts in top universities and industry. The editors have taken extreme care to ensure continuity across the chapters. The material is organized into three parts: A) Tools, B) Models and C) Applications. In Part A, the tools used by business analysts are described in detail. In Part B, these tools are applied to construct models used to solve

business problems. Part C contains detailed applications in various functional areas of business and several case studies. Supporting material can be found in the appendices that develop the pre-requisites for the main text. Every chapter has a business orientation. Typically, each chapter begins with the description of business problems that are transformed into data questions; and methodology is developed to solve these questions. Data analysis is conducted using widely used software, the output and results are clearly explained at each stage of development. These are finally transformed into a business solution. The companion website provides examples, data sets and sample code for each chapter.

Big Data Related Technologies, Challenges and Future Prospects Springer This Springer Brief provides a comprehensive overview of the background and recent developments of big data. The value chain of big data is divided into four phases: data generation, data acquisition, data storage and data analysis. For each phase, the book introduces the general background, discusses technical challenges and reviews the latest advances. Technologies under discussion include cloud computing, Internet of Things, data centers, Hadoop and more. The authors also explore several representative applications of big data such as enterprise management, online social networks, healthcare and medical applications, collective intelligence and smart grids. This book concludes with a thoughtful discussion of possible research directions and development trends in the field. *Big Data: Related Technologies, Challenges and Future Prospects* is a concise yet thorough examination of this exciting area. It is designed for researchers and professionals interested in big data or related research. Advanced-level students in computer science and electrical engineering will also find this book useful.

High-Performance Modelling and Simulation for Big Data Applications Selected Results of the COST Action IC1406 cHiPSet Springer This open access book was prepared as a Final Publication of the COST Action IC1406 “High-Performance Modelling and Simulation for Big Data Applications (cHiPSet)” project. Long considered important pillars of the scientific method, Modelling and Simulation have evolved from traditional discrete numerical methods to complex data-intensive continuous analytical optimisations. Resolution, scale, and accuracy have become essential to predict and analyse natural and complex systems in science and engineering. When their level of abstraction raises to have a better discernment of the domain at hand, their representation gets increasingly demanding for computational and data resources. On the other hand, High Performance Computing typically entails the effective use of parallel and distributed processing units coupled with efficient storage, communication and visualisation systems to underpin complex data-intensive applications in distinct scientific and technical domains. It is then arguably required to have a seamless interaction of High Performance Computing with Modelling and Simulation in order to store, compute, analyse, and visualise large data sets in science and engineering. Funded by the European Commission, cHiPSet has provided a dynamic trans-European forum for their members and distinguished guests to openly discuss novel perspectives and topics of interests for these two communities. This cHiPSet compendium presents a set of selected case studies related to healthcare, biological data, computational advertising, multimedia, finance, bioinformatics, and telecommunications.

Big Data Analytics with Applications in Insider Threat Detection CRC Press

Today's malware mutates randomly to avoid detection, but reactively adaptive malware is more intelligent, learning and adapting to new computer defenses on the fly. Using the same algorithms that antivirus software uses to detect viruses, reactively adaptive malware deploys those algorithms to outwit antivirus defenses and to go undetected. This book provides details of the tools, the types of malware the tools will detect, implementation of the tools in a cloud computing framework and the applications for insider threat detection. **An Introduction to Data Everything You Need to Know About AI, Big Data and Data Science Springer** This book reflects the author's years of hands-on experience as an academic and practitioner. It is primarily intended for executives, managers and practitioners who want to redefine the way they think about artificial intelligence (AI) and other exponential technologies. Accordingly the book, which is structured as a collection of largely self-contained articles, includes both general strategic reflections and detailed sector-specific information. More concretely, it shares insights into what it means to work with AI and how to do it more efficiently; what it means to hire a data scientist and what new roles there are in the field; how to use AI in specific industries such as finance or insurance; how AI interacts with other technologies such as blockchain; and, in closing, a review of the use of AI in venture capital, as well as a snapshot of acceleration programs for AI companies. **Secure Data Science Integrating Cyber Security and Data Science CRC Press** Secure data science, which integrates cyber security and data science, is becoming one of the critical areas in both cyber security and data science. This is because the novel data science techniques being developed have applications in solving such cyber security problems as intrusion detection, malware analysis, and insider threat detection. However, the data science techniques being applied not only for cyber security but also for every application area—including healthcare, finance, manufacturing, and marketing—could be attacked by malware. Furthermore, due to the power of data science, it is now possible to infer highly private and sensitive information from public data, which could result in the violation of individual privacy. This is the first such book that provides a comprehensive overview of integrating both cyber security and data science and discusses both theory and practice in secure data science. After an overview of security and privacy for big data services as well as cloud computing, this book describes applications of data science for cyber security applications. It also discusses such applications of data science as malware analysis and insider threat detection. Then this book addresses trends in adversarial machine learning and provides solutions to the attacks on the data science techniques. In particular, it discusses some emerging trends in carrying out trustworthy analytics so that the analytics techniques can be secured against malicious attacks. Then it focuses on the privacy threats due to the collection of massive amounts of data and potential solutions. Following a discussion on the integration of services computing, including cloud-based services for secure data science, it looks at applications of secure data science to information sharing and social media. This book is a useful resource for researchers, software developers, educators, and managers who want to understand both the high level concepts and the technical details on the design and implementation of secure data science-based systems. It can also be used as a reference book for a graduate course in secure data science. Furthermore, this book provides numerous references that would be helpful for the reader

to get more details about secure data science.