
Foundations Of Data Quality Management Synthesis Lectures On Data Man

Ten Steps to Quality Data and Trusted Information (TM)

16th Asia-Pacific Web Conference, APWeb 2014, Changsha, China, September 5-7, 2014. Proceedings

Security, Privacy, and Anonymity in Computation, Communication, and Storage
XXVI IJCIEOM, Rio de Janeiro, Brazil, July 8-11, 2020

The 1st International Conference on Advanced Intelligent System and Informatics
(AISI2015), November 28-30, 2015, Beni Suef, Egypt

Advances in Visual Informatics

Web-Age Information Management

Flexible Query Answering Systems

Data Cleaning

Invariant Probabilities of Markov-Feller Operators and Their Supports

A Practical Perspective

Non-Volatile Memory Database Management Systems

Handbook of Data Quality

Database Systems for Advanced Applications

DASFAA 2016 International Workshops: BDMS, BDQM, Mol, and SeCoP, Dallas, TX, USA, April 16-19, 2016, Proceedings

Database Systems for Advanced Applications

11th International Conference, CAiSE'99, Heidelberg, Germany, June 14-18, 1999, Proceedings

11th International Conference and Satellite Workshops, SpaCCS 2018, Melbourne, NSW, Australia, December 11-13, 2018, Proceedings

Confederated International Conferences: CoopIS and ODBASE 2014, Amantea, Italy, October 27-31, 2014. Proceedings

Data-Intensive Workflow Management

Executing Data Quality Projects

Web Information Systems Engineering - WISE 2017

Data Management in Machine Learning Systems

Essays dedicated to Peter Buneman

Computational Intelligence, Cyber Security and Computational Models

Advances in Conceptual Modeling

In Search of Elegance in the Theory and Practice of Computation
Advanced Information Systems Engineering
Web Technologies and Applications
14th International Conference, FQAS 2021, Bratislava, Slovakia, September 19–24,
2021, Proceedings
Proceedings of ICC3, 2013
Corporate Data Quality
Proceedings of the 3rd International Conference of Reliable Information and
Communication Technology (IRICT 2018)
Data Cleaning
Prerequisite for Successful Business Models
Natural Language Data Management and Interfaces
5th International Visual Informatics Conference, IVIC 2017, Bangi, Malaysia,
November 28–30, 2017, Proceedings
Data Processing on FPGAs
18th International Conference, Puschino, Russia, October 7-11, 2017, Proceedings,
Part I

WINTERS BRODY

Ten Steps to Quality Data and Trusted Information (TM) Springer

This book constitutes the workshop proceedings of the 21st International Conference on Database Systems for Advanced Applications, DASFAA 2016, held in Dallas, TX, USA, in April 2016. The volume contains 32 full papers (selected from 43 submissions) from 4 workshops, each focusing on a specific area that contributes to the main themes of DASFAA 2016:

The Third International Workshop on Semantic Computing and Personalization, SeCoP 2016; the Third International Workshop on Big Data Management and Service, BDMS 2016; the First International Workshop on Big Data Quality Management, BDQM 2016; and the Second International Workshop on Mobile of Internet, Mol 2016. *16th Asia-Pacific Web Conference, APWeb 2014, Changsha, China, September 5-7, 2014. Proceedings* IGI Global

This book constitutes the refereed proceedings of the 14th International Conference on Flexible Query Answering Systems, FQAS 2021, held virtually and in Bratislava, Slovakia, in September 2021. The 16 full papers and 1 perspective papers presented were carefully reviewed and selected from 17 submissions. They are organized in the following topical sections: model-based flexible query answering approaches and data-driven approaches. *Security, Privacy, and*

*Anonymity in
Computation,
Communication, and
Storage* Springer

This book presents the proceedings of the 3rd International Conference of Reliable Information and Communication Technology 2018 (IRICT 2018), which was held in Kuala Lumpur, Malaysia, on July 23–24, 2018. The main theme of the conference was “Data Science, AI and IoT Trends for the Fourth Industrial Revolution.” A total of 158 papers were submitted to the conference, of which

103 were accepted and considered for publication in this book. Several hot research topics are covered, including Advances in Data Science and Big Data Analytics, Artificial Intelligence and Soft Computing, Business Intelligence, Internet of Things (IoT) Technologies and Applications, Intelligent Communication Systems, Advances in Computer Vision, Health Informatics, Reliable Cloud Computing Environments, Recent Trends in Knowledge Management, Security

Issues in the Cyber World, and Advances in Information Systems Research, Theories and Methods.

XXVI IJCIEOM, Rio de Janeiro, Brazil, July 8-11, 2020 Academic Press

The issue of data quality is as old as data itself. However, the proliferation of diverse, large-scale and often publically available data on the Web has increased the risk of poor data quality and misleading data interpretations. On the other hand, data is now

exposed at a much more strategic level e.g. through business intelligence systems, increasing manifold the stakes involved for individuals, corporations as well as government agencies. There, the lack of knowledge about data accuracy, currency or completeness can have erroneous and even catastrophic results. With these changes, traditional approaches to data management in general, and data quality control specifically, are challenged. There is an

evident need to incorporate data quality considerations into the whole data cycle, encompassing managerial/governance as well as technical aspects. Data quality experts from research and industry agree that a unified framework for data quality management should bring together organizational, architectural and computational approaches. Accordingly, Sadiq structured this handbook in four parts: Part I is on organizational

solutions, i.e. the development of data quality objectives for the organization, and the development of strategies to establish roles, processes, policies, and standards required to manage and ensure data quality. Part II, on architectural solutions, covers the technology landscape required to deploy developed data quality management processes, standards and policies. Part III, on computational solutions, presents effective and efficient tools and

techniques related to record linkage, lineage and provenance, data uncertainty, and advanced integrity constraints. Finally, Part IV is devoted to case studies of successful data quality initiatives that highlight the various aspects of data quality in action. The individual chapters present both an overview of the respective topic in terms of historical research and/or practice and state of the art, as well as specific techniques, methodologies and

frameworks developed by the individual contributors. Researchers and students of computer science, information systems, or business management as well as data professionals and practitioners will benefit most from this handbook by not only focusing on the various sections relevant to their research area or particular practical work, but by also studying chapters that they may initially consider not to be directly relevant to them, as there they will learn about new perspectives

and approaches. *The 1st International Conference on Advanced Intelligent System and Informatics (AISII2015), November 28-30, 2015, Beni Suef, Egypt* Springer Executing Data Quality Projects, Second Edition presents a structured yet flexible approach for creating, improving, sustaining and managing the quality of data and information within any organization. Studies show that data quality problems are costing businesses billions of dollars each year, with

poor data linked to waste and inefficiency, damaged credibility among customers and suppliers, and an organizational inability to make sound decisions. Help is here! This book describes a proven Ten Step approach that combines a conceptual framework for understanding information quality with techniques, tools, and instructions for practically putting the approach to work - with the end result of high-quality trusted data and information, so critical to today's data-

dependent organizations. The Ten Steps approach applies to all types of data and all types of organizations - for-profit in any industry, non-profit, government, education, healthcare, science, research, and medicine. This book includes numerous templates, detailed examples, and practical advice for executing every step. At the same time, readers are advised on how to select relevant steps and apply them in different ways to best address the many situations they will

face. The layout allows for quick reference with an easy-to-use format highlighting key concepts and definitions, important checkpoints, communication activities, best practices, and warnings. The experience of actual clients and users of the Ten Steps provide real examples of outputs for the steps plus highlighted, sidebar case studies called Ten Steps in Action. This book uses projects as the vehicle for data quality work and the word broadly to include: 1) focused data quality

improvement projects, such as improving data used in supply chain management, 2) data quality activities in other projects such as building new applications and migrating data from legacy systems, integrating data because of mergers and acquisitions, or untangling data due to organizational breakups, and 3) ad hoc use of data quality steps, techniques, or activities in the course of daily work. The Ten Steps approach can also be used to enrich an organization's

standard SDLC (whether sequential or Agile) and it complements general improvement methodologies such as six sigma or lean. No two data quality projects are the same but the flexible nature of the Ten Steps means the methodology can be applied to all. The new Second Edition highlights topics such as artificial intelligence and machine learning, Internet of Things, security and privacy, analytics, legal and regulatory requirements, data science, big data, data

lakes, and cloud computing, among others, to show their dependence on data and information and why data quality is more relevant and critical now than ever before. Includes concrete instructions, numerous templates, and practical advice for executing every step of The Ten Steps approach. Contains real examples from around the world, gleaned from the author's consulting practice and from those who implemented based on her training courses and

the earlier edition of the book Allows for quick reference with an easy-to-use format highlighting key concepts and definitions, important checkpoints, communication activities, and best practices A companion Web site includes links to numerous data quality resources, including many of the templates featured in the text, quick summaries of key ideas from the Ten Steps methodology, and other tools and information that are available online

Advances in Visual Informatics Springer
The LNCS journal Transactions on Large-Scale Data and Knowledge-Centered Systems focuses on data management, knowledge discovery, and knowledge processing, which are core and hot topics in computer science. Since the 1990s, the Internet has become the main driving force behind application development in all domains. An increase in the demand for resource sharing (e.g., computing resources,

services, metadata, data sources) across different sites connected through networks has led to an evolution of data- and knowledge-management systems from centralized systems to decentralized systems enabling large-scale distributed applications providing high scalability. This, the 50th issue of Transactions on Large-Scale Data and Knowledge-Centered Systems, contains five fully revised selected regular papers. Topics covered include data anonymization, quasi-

identifier discovery methods, symbolic time series representation, detection of anomalies in time series, data quality management in biobanks, and the use of multi-agent technology in the design of intelligent systems for maritime transport.

Web-Age Information Management Morgan & Claypool

Foundations of Data Quality

Management Morgan & Claypool Publishers

Flexible Query Answering Systems Springer

This book constitutes the

thoroughly refereed proceedings of the 6th International Conference on Data Management Technologies and Applications, DATA 2017, held in Madrid, Spain, in July 2017. The 13 revised full papers were carefully reviewed and selected from 66 submissions. The papers deal with the following topics: databases, big data, data mining, data management, data security, and other aspects of information systems and technology involving advanced

applications of data.

Data Cleaning Morgan & Claypool Publishers

This volume gathers selected peer-reviewed papers presented at the XXVI International Joint Conference on Industrial Engineering and Operations Management (IJCIEOM), held on July 8-11, 2020 in Rio de Janeiro, Brazil. The respective chapters address a range of timely topics in industrial engineering, including operations and process management, global operations, managerial

economics, data science and stochastic optimization, logistics and supply chain management, quality management, product development, strategy and organizational engineering, knowledge and information management, work and human factors, sustainability, production engineering education, healthcare operations management, disaster management, and more. These topics broadly involve fields like operations,

manufacturing, industrial and production engineering, and management. Given its scope, the book offers a valuable resource for those engaged in optimization research, operations research, and practitioners alike. Invariant Probabilities of Markov-Feller Operators and Their Supports epubli This two volume set LNCS 9642 and LNCS 9643 constitutes the refereed proceedings of the 21st International Conference on Database Systems for Advanced Applications,

DASFAA 2016, held in Dallas, TX, USA, in April 2016. The 61 full papers presented were carefully reviewed and selected from a total of 183 submissions. The papers cover the following topics: crowdsourcing, data quality, entity identification, data mining and machine learning, recommendation, semantics computing and knowledge base, textual data, social networks, complex queries, similarity computing, graph databases, and miscellaneous, advanced

applications.

A Practical Perspective

Springer

This book constitutes the refereed proceedings of five workshops symposia, held at the 37th International Conference on Conceptual Modeling, ER 2018, in Xi'an, China, in October 2018. The 42 papers promote and disseminate research on theories of concepts underlying conceptual modeling, methods and tools for developing and communicating conceptual models, techniques for

transforming conceptual models into effective implementations, and the impact of conceptual modeling techniques on databases, business strategies and information systems. The following workshops are included in this volume: Emp-ER: Empirical Methods in Conceptual Modeling, MoBiD: Modeling and Management of Big Data, MREBA: Conceptual Modeling in Requirements and Business Analysis, QMMQ: Quality of Models and Models of Quality, SCME: Conceptual

Modeling Education.

Non-Volatile Memory Database Management Systems Springer

This Festschrift volume, published in honour of Peter Buneman, contains contributions written by some of his colleagues, former students, and friends. In celebration of his distinguished career a colloquium was held in Edinburgh, Scotland, 27-29 October, 2013. The articles presented herein belong to some of the many areas of Peter's research interests.

Handbook of Data

Quality Morgan & Claypool Publishers

This book contains cutting-edge research material presented by researchers, engineers, developers, and practitioners from academia and industry at the International Conference on Computational Intelligence, Cyber Security and Computational Models (ICC3) organized by PSG College of Technology, Coimbatore, India during December 19–21, 2013. The materials in the book

include theory and applications to provide design, analysis, and modeling of the key areas. The book will be useful material for students, researchers, professionals, as well as academicians in understanding current research trends and findings and future scope of research in computational intelligence, cyber security, and computational models. *Database Systems for Advanced Applications* Morgan & Claypool

Publishers

This book covers invariant probabilities for a large class of discrete-time homogeneous Markov processes known as Feller processes. These Feller processes appear in the study of iterated function systems with probabilities, convolution operators, and certain time series. From the reviews: "A very useful reference for researchers wishing to enter the area of stationary Markov processes both from a probabilistic and a dynamical point of view."

--MONATSHEFTE FÜR
MATHEMATIK
DASFAA 2016
International Workshops:
BDMS, BDQM, Mol, and
SeCoP, Dallas, TX, USA,
April 16-19, 2016,
Proceedings Springer
Continuous improvements
in technological
applications have allowed
more opportunities to
develop automated
systems. This not only
leads to higher success in
smart data analysis, but it
increases the overall
probability of
technological progression.
The Handbook of

Research on Machine
Learning Innovations and
Trends is a key resource
on the latest advances
and research regarding
the vast range of
advanced systems and
applications involved in
machine intelligence.
Highlighting
multidisciplinary studies
on decision theory,
intelligent search, and
multi-agent systems, this
publication is an ideal
reference source for
professionals and
researchers working in
the field of machine
learning and its

applications.
Database Systems for
Advanced Applications
Springer
This book constitutes the
refereed proceedings of
the 15th International
Conference on Web-Age
Information Management,
WAIM 2014, held in
Macau, China, in June
2014. The 48 revised full
papers presented
together with 35 short
papers were carefully
reviewed and selected
from numerous
submissions. The papers
are organized in topical
sections on information

retrieval; recommender systems; query processing and optimization; data mining; data and information quality; information extraction; mobile and pervasive computing; stream, time-series; security and privacy; semantic web; cloud computing; new hardware; crowdsourcing; social computing.

11th International Conference, CAiSE'99, Heidelberg, Germany, June 14-18, 1999, Proceedings Springer Science & Business Media

Workflows may be defined as abstractions used to model the coherent flow of activities in the context of an in silico scientific experiment. They are employed in many domains of science such as bioinformatics, astronomy, and engineering. Such workflows usually present a considerable number of activities and activations (i.e., tasks associated with activities) and may need a long time for execution. Due to the continuous need to store and process data efficiently (making

them data-intensive workflows), high-performance computing environments allied to parallelization techniques are used to run these workflows. At the beginning of the 2010s, cloud technologies emerged as a promising environment to run scientific workflows. By using clouds, scientists have expanded beyond single parallel computers to hundreds or even thousands of virtual machines. More recently, Data-Intensive Scalable Computing (DISC)

frameworks (e.g., Apache Spark and Hadoop) and environments emerged and are being used to execute data-intensive workflows. DISC environments are composed of processors and disks in large-commodity computing clusters connected using high-speed communications switches and networks. The main advantage of DISC frameworks is that they support and grant efficient in-memory data management for large-scale applications, such as

data-intensive workflows. However, the execution of workflows in cloud and DISC environments raise many challenges such as scheduling workflow activities and activations, managing produced data, collecting provenance data, etc. Several existing approaches deal with the challenges mentioned earlier. This way, there is a real need for understanding how to manage these workflows and various big data platforms that have been developed and introduced. As such, this

book can help researchers understand how linking workflow management with Data-Intensive Scalable Computing can help in understanding and analyzing scientific big data. In this book, we aim to identify and distill the body of work on workflow management in clouds and DISC environments. We start by discussing the basic principles of data-intensive scientific workflows. Next, we present two workflows that are executed in a single site and multi-site clouds taking advantage

of provenance. Afterward, we go towards workflow management in DISC environments, and we present, in detail, solutions that enable the optimized execution of the workflow using frameworks such as Apache Spark and its extensions.

11th International Conference and Satellite Workshops, SpACCS 2018, Melbourne, NSW, Australia, December 11-13, 2018, Proceedings Springer Science & Business Media

Data is the foundation of the digital economy. Industry 4.0 and digital services are producing so far unknown quantities of data and make new business models possible. Under these circumstances, data quality has become the critical factor for success. This book presents a holistic approach for data quality management and presents ten case studies about this issue. It is intended for practitioners dealing with data quality management and data governance as well as for

scientists. The book was written at the Competence Center Corporate Data Quality (CC CDQ) in close cooperation between researchers from the University of St. Gallen and Fraunhofer IML as well as many representatives from more than 20 major corporations. Chapter 1 introduces the role of data in the digitization of business and society and describes the most important business drivers for data quality. It presents the Framework

for Corporate Data Quality Management and introduces essential terms and concepts. Chapter 2 presents practical, successful examples of the management of the quality of master data based on ten cases studies that were conducted by the CC CDQ. The case studies cover every aspect of the Framework for Corporate Data Quality Management. Chapter 3 describes selected tools for master data quality management. The three tools have been

distinguished through their broad applicability (method for DQM strategy development and DQM maturity assessment) and their high level of innovation (Corporate Data League). Chapter 4 summarizes the essential factors for the successful management of the master data quality and provides a checklist of immediate measures that should be addressed immediately after the start of a data quality management project. This guarantees a quick start into the topic and

provides initial recommendations for actions to be taken by project and line managers. Please also check out the book's homepage at <http://www.cdq-book.org/>
Confederated International Conferences: CoopIS and ODBASE 2014, Amantea, Italy, October 27-31, 2014. Proceedings Springer
The significance of big data can be observed in any decision-making process as it is often used for forecasting and

predictive analytics. Additionally, big data can be used to build a holistic view of an enterprise through a collection and analysis of large data sets retrospectively. As the data deluge deepens, new methods for analyzing, comprehending, and making use of big data become necessary. *Enterprise Big Data Engineering, Analytics, and Management* presents novel methodologies and practical approaches to engineering, managing, and analyzing large-scale

data sets with a focus on enterprise applications and implementation. Featuring essential big data concepts including data mining, artificial intelligence, and information extraction, this publication provides a platform for retargeting the current research available in the field. Data analysts, IT professionals, researchers, and graduate-level students will find the timely research presented in this publication essential to furthering their knowledge in the field.

Data-Intensive Workflow Management

Springer

Roughly a decade ago, power consumption and heat dissipation concerns forced the semiconductor industry to radically change its course, shifting from sequential to parallel computing. Unfortunately, improving performance of applications has now become much more difficult than in the good old days of frequency scaling. This is also affecting databases and data processing applications in general,

and has led to the popularity of so-called data appliances—specialized data processing engines, where software and hardware are sold together in a closed box. Field-programmable gate arrays (FPGAs) increasingly play an important role in such systems. FPGAs are attractive because the performance gains of specialized hardware can be significant, while power consumption is much less than that of commodity processors.

On the other hand, FPGAs are way more flexible than hard-wired circuits (ASICs) and can be integrated into complex systems in many different ways, e.g., directly in the network for a high-frequency trading application. This book gives an introduction to FPGA technology targeted at a database audience. In the first few chapters, we explain in detail the inner workings of FPGAs. Then we discuss techniques and design patterns that help mapping algorithms to FPGA hardware so that

the inherent parallelism of these devices can be leveraged in an optimal way. Finally, the book will illustrate a number of concrete examples that exploit different advantages of FPGAs for data processing. Table of Contents: Preface / Introduction / A Primer in Hardware Design / FPGAs / FPGA Programming Models / Data Stream Processing / Accelerated DB Operators / Secure Data Processing / Conclusions / Bibliography / Authors' Biographies / Index