
Data Matching Concepts And Techniques For Record Linkage Entity Resolution And Duplicate Detection Data Centric Systems And Applications

Concepts and Techniques for Querying and Analyzing Process Data
 Principles of Data Integration
 Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection
 Concepts, Methodologies, Tools, and Applications
 Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications
 A User's Guide
 Data Quality and Record Linkage Techniques
 Registries for Evaluating Patient Outcomes
 Data Analytics in Medicine: Concepts, Methodologies, Tools, and Applications
 An Introduction to Duplicate Detection
 Data Matching
 Concepts, Methodologies, Tools, and Applications
 Real-time Linked Dataspaces
 Data Mining with Rattle and R
 Concepts and Resources for Managers
 Research Anthology on Privatizing and Securing Data
 Achievements and Opportunities
 Concepts, Principles, and Practices
 The Behavioral and Social Sciences
 Biometrics: Concepts, Methodologies, Tools, and Applications
 Concepts, Methodologies, Tools, and Applications
 Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications
 Database Technologies: Concepts, Methodologies, Tools, and Applications
 Data Deduplication Approaches
 Concepts, Methodologies, Tools, and Applications
 Flexible Imputation of Missing Data, Second Edition
 Knowledge Graphs and Big Data Processing
 Cognitive Analytics: Concepts, Methodologies, Tools, and Applications
 Handbook on Impact Evaluation
 Methodological Developments in Data Linkage
 Methods and Techniques for Practical Privacy-Preserving Information Sharing
 Decision Support Systems
 Concepts, Methodologies and Techniques
 Concepts, Strategies, and Challenges
 Sharing Clinical Trial Data
 Process Analytics
 Mining of Massive Datasets
 Data-Driven Policy Impact Evaluation
 Data Mining for Business Intelligence
 Encyclopedia of Research Design

Data Matching Concepts And Techniques For Record Linkage Entity Resolution And Duplicate Detection Data Centric Systems And Applications

Downloaded from ecobankpayservices.ecobank.com by guest

JONATHAN ZION

Concepts and Techniques for Querying and Analyzing Process Data World Bank Publications

Data sharing can accelerate new discoveries by avoiding duplicative trials, stimulating new ideas for research, and enabling the maximal scientific knowledge and benefits to be gained from the efforts of clinical trial participants and investigators. At the same time, sharing clinical trial data presents risks, burdens, and challenges. These include the need to protect the privacy and honor the consent of clinical trial participants; safeguard the legitimate economic interests of sponsors; and guard against invalid secondary analyses, which could undermine trust in clinical trials or otherwise harm public health. *Sharing Clinical Trial Data* presents activities and strategies for the responsible sharing of clinical trial data. With the goal of increasing scientific knowledge to lead to better therapies for patients, this book identifies guiding principles and makes recommendations to maximize the benefits and minimize risks. This report offers guidance on the types of clinical trial data available at different points in the process, the

points in the process at which each type of data should be shared, methods for sharing data, what groups should have access to data, and future knowledge and infrastructure needs. Responsible sharing of clinical trial data will allow other investigators to replicate published findings and carry out additional analyses, strengthen the evidence base for regulatory and clinical decisions, and increase the scientific knowledge gained from investments by the funders of clinical trials. The recommendations of *Sharing Clinical Trial Data* will be useful both now and well into the future as improved sharing of data leads to a stronger evidence base for treatment. This book will be of interest to stakeholders across the spectrum of research—from funders, to researchers, to journals, to physicians, and ultimately, to patients.

Principles of Data Integration SAGE

In the light of better and more detailed administrative databases, this open access book provides statistical tools for evaluating the effects of public policies advocated by governments and public institutions. Experts from academia, national statistics offices and various research centers present modern econometric methods for an efficient data-driven policy evaluation and monitoring, assess the causal effects of policy measures and report on best practices of successful data management and usage. Topics include data confidentiality, data linkage, and national practices in policy areas such as public health, education and employment. It offers scholars as well as practitioners from public administrations, consultancy firms and

nongovernmental organizations insights into counterfactual impact evaluation methods and the potential of data-based policy and program evaluation.

Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection SAGE Publications

Poor data quality can seriously hinder or damage the efficiency and effectiveness of organizations and businesses. The growing awareness of such repercussions has led to major public initiatives like the "Data Quality Act" in the USA and the "European 2003/98" directive of the European Parliament. Batini and Scannapieco present a comprehensive and systematic introduction to the wide set of issues related to data quality. They start with a detailed description of different data quality dimensions, like accuracy, completeness, and consistency, and their importance in different types of data, like federated data, web data, or time-dependent data, and in different data categories classified according to frequency of change, like stable, long-term, and frequently changing data. The book's extensive description of techniques and methodologies from core data quality research as well as from related fields like data mining, probability theory, statistical data analysis, and machine learning gives an excellent overview of the current state of the art. The presentation is completed by a short description and critical comparison of tools and practical methodologies, which will help readers to resolve their own quality problems. This book is an ideal combination of the soundness of theoretical foundations and the applicability of practical approaches. It is ideally suited for everyone – researchers, students, or professionals – interested in a comprehensive overview of data quality issues. In addition, it will serve as the basis for an introductory course or for self-study on this topic.

Concepts, Methodologies, Tools, and Applications IGI Global

This open access book is part of the LAMBDA Project (Learning, Applying, Multiplying Big Data Analytics), funded by the European Union, GA No. 809965. Data Analytics involves applying algorithmic processes to derive insights. Nowadays it is used in many industries to allow organizations and companies to make better decisions as well as to verify or disprove existing theories or models. The term data analytics is often used interchangeably with intelligence, statistics, reasoning, data mining, knowledge discovery, and others. The goal of this book is to introduce some of the definitions, methods, tools, frameworks, and solutions for big data processing, starting from the process of information extraction and knowledge representation, via knowledge processing and analytics to visualization, sense-making, and practical applications. Each chapter in this book addresses some pertinent aspect of the data processing chain, with a specific focus on understanding Enterprise Knowledge Graphs, Semantic Big Data Architectures, and Smart Data Analytics solutions. This book is addressed to graduate students from technical disciplines, to professional audiences following continuous education short courses, and to researchers from diverse areas following self-study courses. Basic skills in computer science, mathematics, and statistics are required.

Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications Springer

This open access book explores the dataspace paradigm as a best-effort approach to data management within data ecosystems. It establishes the theoretical foundations and principles of real-time linked dataspace as a data platform for intelligent systems. The book introduces a set of specialized best-effort techniques and models to enable loose administrative proximity and semantic integration for managing and processing events and streams. The book is divided into five major parts: Part I "Fundamentals and Concepts" details the motivation behind and core concepts of real-time linked dataspace, and establishes the need to evolve data management techniques in order to meet the challenges of enabling data ecosystems for intelligent systems within smart environments. Further, it explains the fundamental concepts of dataspace and the need for specialization in the processing of dynamic real-time data. Part II "Data Support Services" explores the design and evaluation of critical services, including catalog, entity management, query and search, data service discovery, and human-in-the-loop. In turn, Part III "Stream and Event Processing Services" addresses the design and evaluation of the specialized techniques created for real-time support services including complex event processing, event service composition, stream dissemination, stream matching, and approximate semantic matching. Part IV "Intelligent Systems and Applications" explores the use of real-time linked dataspace within real-world smart environments. In closing, Part V "Future Directions" outlines future research challenges for dataspace, data ecosystems, and intelligent systems. Readers will gain a detailed understanding of how the dataspace paradigm is now being used to enable data ecosystems for intelligent systems within smart environments. The book covers the fundamental theory, the creation of new techniques needed for support services, and lessons learned from real-world intelligent systems and applications focused on sustainability. Accordingly, it will benefit not only researchers and graduate students in the fields of data management, big data, and IoT, but also professionals who need to create advanced data management platforms for intelligent systems, smart environments, and data ecosystems.

A User's Guide Springer Science & Business Media

For MIS specialists and nonspecialists alike, a comprehensive, readable, understandable guide to the concepts and applications of decision support systems.

Data Quality and Record Linkage Techniques IGI Global

Data matching (also known as record or data linkage, entity resolution, object identification, or field matching) is the task of identifying, matching and merging records that correspond to the same entities from several databases or even within one database. Based on research in various domains including applied statistics, health informatics, data mining, machine learning, artificial intelligence, database management, and digital libraries, significant advances have been achieved over the last decade in all aspects of the data matching process, especially on how to improve the accuracy of data matching, and its scalability to large databases. Peter Christen's book is divided into three parts: Part I, "Overview", introduces the subject by presenting several sample applications and their special challenges, as well as a general overview of a generic data matching process. Part II, "Steps of the Data Matching Process", then details its main steps like pre-processing, indexing, field and record comparison, classification, and quality evaluation. Lastly, part III, "Further Topics", deals with specific aspects like privacy, real-time matching, or matching unstructured data. Finally, it briefly describes the main features of many research and open source systems available today. By providing the reader with a broad range of data matching concepts and techniques and touching on all aspects of the data matching process, this book helps researchers as well as students specializing in data quality or data matching aspects to familiarize themselves with recent research advances and to identify open research

challenges in the area of data matching. To this end, each chapter of the book includes a final section that provides pointers to further background and research material. Practitioners will better understand the current state of the art in data matching as well as the internal workings and limitations of current systems. Especially, they will learn that it is often not feasible to simply implement an existing off-the-shelf data matching system without substantial adaption and customization. Such practical considerations are discussed for each of the major steps in the data matching process.

Registries for Evaluating Patient Outcomes Springer

Due to the growing use of web applications and communication devices, the use of data has increased throughout various industries, including business and healthcare. It is necessary to develop specific software programs that can analyze and interpret large amounts of data quickly in order to ensure adequate usage and predictive results. Cognitive Analytics: Concepts, Methodologies, Tools, and Applications provides emerging perspectives on the theoretical and practical aspects of data analysis tools and techniques. It also examines the incorporation of pattern management as well as decision-making and prediction processes through the use of data management and analysis. Highlighting a range of topics such as natural language processing, big data, and pattern recognition, this multi-volume book is ideally designed for information technology professionals, software developers, data analysts, graduate-level students, researchers, computer engineers, software engineers, IT specialists, and academicians.

Data Analytics in Medicine: Concepts, Methodologies, Tools, and Applications Springer Science & Business Media

Now in its second edition, this book focuses on practical algorithms for mining data from even the largest datasets.

An Introduction to Duplicate Detection Academic Press

With the proliferation of huge amounts of (heterogeneous) data on the Web, the importance of information retrieval (IR) has grown considerably over the last few years. Big players in the computer industry, such as Google, Microsoft and Yahoo!, are the primary contributors of technology for fast access to Web-based information; and searching capabilities are now integrated into most information systems, ranging from business management software and customer relationship systems to social networks and mobile phone applications. Ceri and his co-authors aim at taking their readers from the foundations of modern information retrieval to the most advanced challenges of Web IR. To this end, their book is divided into three parts. The first part addresses the principles of IR and provides a systematic and compact description of basic information retrieval techniques (including binary, vector space and probabilistic models as well as natural language search processing) before focusing on its application to the Web. Part two addresses the foundational aspects of Web IR by discussing the general architecture of search engines (with a focus on the crawling and indexing processes), describing link analysis methods (specifically Page Rank and HITS), addressing recommendation and diversification, and finally presenting advertising in search (the main source of revenues for search engines). The third and final part describes advanced aspects of Web search, each chapter providing a self-contained, up-to-date survey on current Web research directions. Topics in this part include meta-search and multi-domain search, semantic search, search in the context of multimedia data, and crowd search. The book is ideally suited to courses on information retrieval, as it covers all Web-independent foundational aspects. Its presentation is self-contained and does not require prior background knowledge. It can also be used in the context of classic courses on data management, allowing the instructor to cover both structured and unstructured data in various formats. Its classroom use is facilitated by a set of slides, which can be downloaded from www.search-computing.org.

Data Matching John Wiley & Sons

This User's Guide is intended to support the design, implementation, analysis, interpretation, and quality evaluation of registries created to increase understanding of patient outcomes. For the purposes of this guide, a patient registry is an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes. A registry database is a file (or files) derived from the registry. Although registries can serve many purposes, this guide focuses on registries created for one or more of the following purposes: to describe the natural history of disease, to determine clinical effectiveness or cost-effectiveness of health care products and services, to measure or monitor safety and harm, and/or to measure quality of care. Registries are classified according to how their populations are defined. For example, product registries include patients who have been exposed to biopharmaceutical products or medical devices. Health services registries consist of patients who have had a common procedure, clinical encounter, or hospitalization. Disease or condition registries are defined by patients having the same diagnosis, such as cystic fibrosis or heart failure. The User's Guide was created by researchers affiliated with AHRQ's Effective Health Care Program, particularly those who participated in AHRQ's DEClDE (Developing Evidence to Inform Decisions About Effectiveness) program. Chapters were subject to multiple internal and external independent reviews.

Concepts, Methodologies, Tools, and Applications CRC Press

Public programs are designed to reach certain goals and beneficiaries. Methods to understand whether such programs actually work, as well as the level and nature of impacts on intended beneficiaries, are main themes of this book.

Real-time Linked Dataspaces SAGE

Fully updated to reflect the most recent changes in the field, the Second Edition of Propensity Score Analysis provides an accessible, systematic review of the origins, history, and statistical foundations of propensity score analysis, illustrating how it can be used for solving evaluation and causal-inference problems. With a strong focus on practical applications, the authors explore various strategies for employing PSA, discuss the use of PSA with alternative types of data, and delineate the limitations of PSA under a variety of constraints. Unlike existing textbooks on program evaluation and causal inference, this book delves into statistical concepts, formulas, and models within the context of a robust and engaging focus on application.

Data Mining with Rattle and R Springer

The contributors to Best Practices in Quantitative Methods envision quantitative methods in the 21st century, identify the best practices, and, where possible, demonstrate the superiority of their recommendations empirically. Editor Jason W. Osborne designed this book with the goal of providing readers with the most effective, evidence-based, modern quantitative methods and quantitative data analysis across the social and behavioral sciences. The text is divided into five main sections covering select best practices in Measurement, Research Design, Basics of Data Analysis,

Quantitative Methods, and Advanced Quantitative Methods. Each chapter contains a current and expansive review of the literature, a case for best practices in terms of method, outcomes, inferences, etc., and broad-ranging examples along with any empirical evidence to show why certain techniques are better. Key Features: Describes important implicit knowledge to readers: The chapters in this volume explain the important details of seemingly mundane aspects of quantitative research, making them accessible to readers and demonstrating why it is important to pay attention to these details. Compares and contrasts analytic techniques: The book examines instances where there are multiple options for doing things, and make recommendations as to what is the "best" choice—or choices, as what is best often depends on the circumstances. Offers new procedures to update and explicate traditional techniques: The featured scholars present and explain new options for data analysis, discussing the advantages and disadvantages of the new procedures in depth, describing how to perform them, and demonstrating their use. Intended Audience: Representing the vanguard of research methods for the 21st century, this book is an invaluable resource for graduate students and researchers who want a comprehensive, authoritative resource for practical and sound advice from leading experts in quantitative methods.

Concepts and Resources for Managers IGI Global

This volume explores the scientific frontiers and leading edges of research across the fields of anthropology, economics, political science, psychology, sociology, history, business, education, geography, law, and psychiatry, as well as the newer, more specialized areas of artificial intelligence, child development, cognitive science, communications, demography, linguistics, and management and decision science. It includes recommendations concerning new resources, facilities, and programs that may be needed over the next several years to ensure rapid progress and provide a high level of returns to basic research.

Research Anthology on Privatizing and Securing Data Morgan & Claypool Publishers

This book starts with an introduction to process modeling and process paradigms, then explains how to query and analyze process models, and how to analyze the process execution data. In this way, readers receive a comprehensive overview of what is needed to identify, understand and improve business processes. The book chiefly focuses on concepts, techniques and methods. It covers a large body of knowledge on process analytics – including process data querying, analysis, matching and correlating process data and models – to help practitioners and researchers understand the underlying concepts, problems, methods, tools and techniques involved in modern process analytics. Following an introduction to basic business process and process analytics concepts, it describes the state of the art in this area before examining different analytics techniques in detail. In this regard, the book covers analytics over different levels of process abstractions, from process execution data and methods for linking and correlating process execution data, to inferring process models, querying process execution data and process models, and scalable process data analytics methods. In addition, it provides a review of commercial process analytics tools and their practical applications. The book is intended for a broad readership interested in business process management and process analytics. It provides researchers with an introduction to these fields by comprehensively classifying the current state of research, by describing in-depth techniques and methods, and by highlighting future research directions. Lecturers will find a wealth of material to choose from for a variety of courses, ranging from undergraduate courses in business process management to graduate courses in business process analytics. Lastly, it offers professionals a reference guide to the state of the art in commercial tools and techniques, complemented by many real-world use case scenarios.

Achievements and Opportunities IGI Global

Praise for the first edition: "This excellent text will be useful to every system engineer (SE) regardless of the domain. It covers ALL relevant SE material and does so in a very clear, methodical fashion. The breadth and depth of the author's presentation of SE principles and practices is outstanding." –Philip Allen This textbook presents a comprehensive, step-by-step guide to System Engineering analysis, design, and development via an integrated set of concepts, principles, practices, and methodologies. The methods presented in this text apply to any type of human system -- small, medium, and large organizational systems and system development projects delivering engineered systems or services across multiple business sectors such as medical, transportation, financial, educational, governmental, aerospace and defense, utilities, political, and charity, among others. Provides a common focal point for "bridging the gap" between and unifying System Users, System Acquirers, multi-discipline System Engineering, and Project, Functional, and Executive Management education, knowledge, and decision-making for developing systems, products, or services Each chapter provides definitions of key terms, guiding principles, examples, author's notes, real-world examples, and exercises, which highlight and reinforce key SE&D concepts and practices Addresses concepts employed in Model-Based Systems Engineering (MBSE), Model-Driven Design (MDD), Unified Modeling Language (UMLTM) / Systems Modeling Language (SysMLTM), and Agile/Spiral/V-Model Development such as user needs, stories, and use cases

analysis; specification development; system architecture development; User-Centric System Design (UCSD); interface definition & control; system integration & test; and Verification & Validation (V&V) Highlights/introduces a new 21st Century Systems Engineering & Development (SE&D) paradigm that is easy to understand and implement. Provides practices that are critical staging points for technical decision making such as Technical Strategy Development; Life Cycle requirements; Phases, Modes, & States; SE Process; Requirements Derivation; System Architecture Development, User-Centric System Design (UCSD); Engineering Standards, Coordinate Systems, and Conventions; et al. Thoroughly illustrated, with end-of-chapter exercises and numerous case studies and examples, Systems Engineering Analysis, Design, and Development, Second Edition is a primary textbook for multi-discipline, engineering, system analysis, and project management undergraduate/graduate level students and a valuable reference for professionals.

Concepts, Principles, and Practices SAGE

In the age of data science, the rapidly increasing amount of data is a major concern in numerous applications of computing operations and data storage. Duplicated data or redundant data is a main challenge in the field of data science research. Data Deduplication Approaches: Concepts, Strategies, and Challenges shows readers the various methods that can be used to eliminate multiple copies of the same files as well as duplicated segments or chunks of data within the associated files. Due to ever-increasing data duplication, its deduplication has become an especially useful field of research for storage environments, in particular persistent data storage. Data Deduplication Approaches provides readers with an overview of the concepts and background of data deduplication approaches, then proceeds to demonstrate in technical detail the strategies and challenges of real-time implementations of handling big data, data science, data backup, and recovery. The book also includes future research directions, case studies, and real-world applications of data deduplication, focusing on reduced storage, backup, recovery, and reliability. Includes data deduplication methods for a wide variety of applications Includes concepts and implementation strategies that will help the reader to use the suggested methods Provides a robust set of methods that will help readers to appropriately and judiciously use the suitable methods for their applications Focuses on reduced storage, backup, recovery, and reliability, which are the most important aspects of implementing data deduplication approaches Includes case studies

The Behavioral and Social Sciences Springer Nature

How do you approach answering queries when your data is stored in multiple databases that were designed independently by different people? This is first comprehensive book on data integration and is written by three of the most respected experts in the field. This book provides an extensive introduction to the theory and concepts underlying today's data integration techniques, with detailed, instruction for their application using concrete examples throughout to explain the concepts. Data integration is the problem of answering queries that span multiple data sources (e.g., databases, web pages). Data integration problems surface in multiple contexts, including enterprise information integration, query processing on the Web, coordination between government agencies and collaboration between scientists. In some cases, data integration is the key bottleneck to making progress in a field. The authors provide a working knowledge of data integration concepts and techniques, giving you the tools you need to develop a complete and concise package of algorithms and applications. Offers a range of data integration solutions enabling you to focus on what is most relevant to the problem at hand Enables you to build your own algorithms and implement your own data integration applications

Biometrics: Concepts, Methodologies, Tools, and Applications IGI Global

Data Mining: Concepts and Techniques provides the concepts and techniques in processing gathered data or information, which will be used in various applications. Specifically, it explains data mining and the tools used in discovering knowledge from the collected data. This book is referred as the knowledge discovery from data (KDD). It focuses on the feasibility, usefulness, effectiveness, and scalability of techniques of large data sets. After describing data mining, this edition explains the methods of knowing, preprocessing, processing, and warehousing data. It then presents information about data warehouses, online analytical processing (OLAP), and data cube technology. Then, the methods involved in mining frequent patterns, associations, and correlations for large data sets are described. The book details the methods for data classification and introduces the concepts and methods for data clustering. The remaining chapters discuss the outlier detection and the trends, applications, and research frontiers in data mining. This book is intended for Computer Science students, application developers, business professionals, and researchers who seek information on data mining. Presents dozens of algorithms and implementation examples, all in pseudo-code and suitable for use in real-world, large-scale data mining projects Addresses advanced topics such as mining object-relational databases, spatial databases, multimedia databases, time-series databases, text databases, the World Wide Web, and applications in several fields Provides a comprehensive, practical look at the concepts and techniques you need to get the most out of your data

Related with Data Matching Concepts And Techniques For Record Linkage Entity Resolution And Duplicate Detection Data Centric Systems And Applications:

© [Data Matching Concepts And Techniques For Record Linkage Entity Resolution And Duplicate Detection Data Centric Systems And Applications The Nature Of Science Answer Key Chapter 1](#)

© [Data Matching Concepts And Techniques For Record Linkage Entity Resolution And Duplicate Detection Data Centric Systems And Applications The More Complicated The Technology That An Organization Uses](#)

© [Data Matching Concepts And Techniques For Record Linkage Entity Resolution And Duplicate Detection Data Centric Systems And Applications The Most Documented Person In History](#)