Ibm Datastage Manuals

Deployment Guide for InfoSphere GuardiumThe AI LadderIBM Smart Analytics SystemInfoSphere DataStage Parallel Framework Standard PracticesBusiness Intelligence GuidebookSmarter Modeling of IBM InfoSphere Master Data Management SolutionsSystem Programmer's Guide to Workload ManagerSmarter Business: Dynamic Information with IBM InfoSphere Data Replication CDCIBM Cloud Private System Administrator's GuideIBM WebSphere Information Analyzer and Data Quality AssessmentMetadata Management with IBM InfoSphere Information ServerIBM Infosphere DatastageAccelerating Data Transformation with IBM DB2 Analytics Accelerator for z/OSMetadata Management with IBM InfoSphere Information ServerData Warehousing with the Informix Dynamic ServerIBM Watson Content Analytics: Discovering Actionable Insight from Your ContentCivil Engineer's Reference BookIBM Technical Computing CloudsIBM InfoSphere Replication Server and Data Event PublisherImplementing an InfoSphere Optim Data Growth SolutionInfoSphere DataStage Parallel Framework Standard PracticesIBM Information Server: Integration and Governance for Emerging Data Warehouse DemandsPython for Data AnalysisImplementing an InfoSphere Optim Data Growth SolutionBMC Control-M 7IBM InfoSphere DataStage Data Flow and Job DesignDB2 pureXML CookbookPrinciples of Data WranglingIBM Information Governance SolutionsMonetizing Your DataIBM Informix Developer's HandbookIBM InfoSphere Information Server Installation and Configuration GuideHarness the Power of Big

Data The IBM Big Data PlatformMigrating from Microsoft SQL Server to IBM InformixInfoSphere DataStage for Enterprise XML Data IntegrationInfoSphere DataStage for Enterprise XML Data IntegrationIBM Cognos 10 Report Studio CookbookBuilding Big Data and Analytics Solutions in the CloudIBM Data WarehousingBig Data For Dummies

Deployment Guide for InfoSphere Guardium

XML is one of the most common standards for the exchange of information. However, organizations find challenges in how to address the complexities of dealing with hierarchical data types, particularly as they scale to gigabytes and beyond. In this IBM® Redbooks® publication, we discuss and describe the new capabilities in IBM InfoSphere® DataStage® 8.5. These capabilities enable developers to more easily manage the design and processing requirements presented by the most challenging XML sources. Developers can use these capabilities to create powerful hierarchical transformations and to parse and compose XML data with high performance and scalability. Spanning both batch and real-time run times, these capabilities can be used to solve a broad range of business requirements. As part of the IBM InfoSphere Information Server 8.5 release, InfoSphere DataStage was enhanced with new hierarchical transformation capabilities called . XML Stage provides native XML schema support and powerful

XML transformation functionality. These capabilities are based on a unique state-of-the-art technology that allows you to parse and compose any complex XML structure from and to a relational form, as well as to a separate hierarchical form. This book is targeted at an audience of systems designers and developers who focus on implementing XML integration support in their environments.

The AI Ladder

A key task that any aspiring data-driven organization needs to learn is data wrangling, the process of converting raw data into something truly useful. This practical guide provides business analysts with an overview of various data wrangling techniques and tools, and puts the practice of data wrangling into context by asking, "What are you trying to do and why?" Wrangling data consumes roughly 50-80% of an analyst's time before any kind of analysis is possible. Written by key executives at Trifacta, this book walks you through the wrangling process by exploring several factors—time, granularity, scope, and structure—that you need to consider as you begin to work with data. You'll learn a shared language and a comprehensive understanding of data wrangling, with an emphasis on recent agile analytic processes used by many of today's data-driven organizations. Appreciate the importance—and the satisfaction—of wrangling data the right way. Understand what kind of data is available Choose which data to use and at what level of detail Meaningfully combine multiple sources of data Decide how to distill

the results to a size and shape that can drive downstream analysis

IBM Smart Analytics System

DB2 pureXML Cookbook Master the Power of the IBM Hybrid Data Server Hands-On Solutions and Best Practices for Developing and Managing XML Database Applications with DB2 More and more database developers and DBAs are being asked to develop applications and manage databases that involve XML data. Many are utilizing the highly praised DB2 pureXML technology from IBM. In the DB2 pureXML Cookbook, two leading experts from IBM offer the practical solutions and proven code samples that database professionals need to build better XML solutions faster. Organized by task, this book is packed with more than 700 easy-to-adapt "recipe-style" examples covering the entire application lifecycle–from planning and design through coding, optimization, and troubleshooting. This extraordinary library of recipes includes more than 250 XQuery and SQL/XML queries. With the authors' hands-on guidance, you'll learn how to combine pureXML "ingredients" to efficiently perform virtually any XML data management task, from the simplest to the most advanced. Coverage includes pureXML in DB2 9 for z/OS and DB2 9.1, 9.5, and 9.7 for Linux, UNIX, and Windows Best practices for designing XML data, applications, and storage objects Importing, exporting, loading, replicating, and federating XML data Querying XML data, from start to finish: XPath and XQuery data model and languages, SQL/XML, stored procedures,

UDFs, and much more Avoiding common errors and inefficient XML queries Converting relational data to XML and vice versa Updating and transforming XML documents Defining and working with XML indexes Monitoring and optimizing the performance of XML queries and other operations Using XML Schemas to constrain and validate XML documents XML application development–including code samples for Java, .NET, C, COBOL,PL/1, PHP, and Perl

InfoSphere DataStage Parallel Framework Standard Practices

Today, organizations face tremendous challenges with data explosion and information governance. InfoSphereTM OptimTM solutions solve the data growth problem at the source by managing the enterprise application data. The Optim Data Growth solutions are consistent, scalable solutions that include comprehensive capabilities for managing enterprise application data across applications, databases, operating systems, and hardware platforms. You can align the management of your enterprise application data with your business objectives to improve application service levels, lower costs, and mitigate risk. In this IBM® Redbooks® publication, we describe the IBM InfoSphere Optim Data Growth solutions and a methodology that provides implementation guidance from requirements analysis through deployment and administration planning. We also discuss various implementation topics including system architecture design, sizing, scalability, security, performance, and automation. This book is intended to provide

various systems development professionals, Data Solution Architects, Data Administrators, Modelers, Data Analysts, Data Integrators, or anyone who has to analyze or integrate data structures, a broad understanding about IBM InfoSphere Optim Data Growth solutions. By being used in conjunction with the product manuals and online help, this book provides guidance about implementing an optimal solution for managing your enterprise application data.

Business Intelligence Guidebook

IBM® WatsonTM Content Analytics (Content Analytics) Version 3.0 (formerly known as IBM Content Analytics with Enterprise Search (ICAwES)) helps you to unlock the value of unstructured content to gain new actionable business insight and provides the enterprise search capability all in one product. Content Analytics comes with a set of tools and a robust user interface to empower you to better identify new revenue opportunities, improve customer satisfaction, detect problems early, and improve products, services, and offerings. To help you gain the most benefits from your unstructured content, this IBM Redbooks® publication provides in-depth information about the features and capabilities of Content Analytics, how the content analytics works, and how to perform effective and efficient content analytics on your content to discover actionable business insights. This book covers key concepts in content analytics, such as facets, frequency, deviation, correlation, trend, and sentimental analysis. It describes the content analytics

miner, and guides you on performing content analytics using views, dictionary lookup, and customization. The book also covers using IBM Content Analytics Studio for domain-specific content analytics, integrating with IBM Content Classification to get categories and new metadata, and interfacing with IBM Cognos® Business Intelligence (BI) to add values in BI reporting and analysis, and customizing the content analytics miner with APIs. In addition, the book describes how to use the enterprise search capability for the discovery and retrieval of documents using various query and visual navigation techniques, and customization of crawling, parsing, indexing, and runtime search to improve search results. The target audience of this book is decision makers, business users, and IT architects and specialists who want to understand and analyze their enterprise content to improve and enhance their business operations. It is also intended as a technical how-to guide for use with the online IBM Knowledge Center for configuring and performing content analytics and enterprise search with Content Analytics.

Smarter Modeling of IBM InfoSphere Master Data Management Solutions

In this IBM® Redbooks® publication, we discuss considerations, and describe a methodology, for transitioning from Microsoft® SQL Server 2008 to the Informix®

Dynamic Server. We focus on the topic areas of data, applications, and administration, providing information about the differences in features and functionality, including the data types, data manipulation language, data definition language, and stored procedures. Understanding the features and functionality of the two products assists you in developing a migration plan. We provide a conversion methodology and discuss the processes for migrating the database objects and data from SQL Server to Informix using various methods. We show the SQL differences between SQL Server and Informix and illustrate, with examples, how to convert tables, views, stored procedures, functions, and triggers. We provide script conversion samples for data loading. We describe application programming and conversion considerations. In addition, we discuss the Informix configuration, as well as the administration features and functions Informix provides to help DBAs manage the Informix database server after it is migrated. With this information, you can develop your required transition methodology, and you can plan and execute the conversion activities in an orderly and cost-effective manner.

System Programmer's Guide to Workload Manager

Civil Engineer's Reference Book, Fourth Edition provides civil engineers with reports on design and construction practices in the UK and overseas. It gives a concise presentation of theory and practice in the many branches of a civil

engineer's profession and it enables them to study a subject in greater depth. The book discusses some improvements in earlier practices, for example in surveying, geotechnics, water management, project management, underwater working, and the control and use of materials. Other changes covered are from the evolving needs of clients for almost all forms of construction, maintenance and repair. Another major change is the introduction of new national and Euro-codes based on limit state design, covering most aspects of structural engineering. The fourth edition incorporates these advances and, at the same time, gives greater prominence to the special problems relating to work overseas, with differing client requirements and climatic conditions. Chapters 1 to 10 provide engineers, at all levels of development, with 'lecture notes' on the basic theories of civil engineering. Chapters 11 to 44 cover the practice of design and construction in many of the fields of civil engineering. Civil engineers, architects, lawyers, mechanical engineers, insurers, clients, and students of civil engineering will find benefit in the use of this text.

Smarter Business: Dynamic Information with IBM InfoSphere Data Replication CDC

Transforming data from operational data models to purpose-oriented data structures has been commonplace for the last decades. Data transformations are

heavily used in all types of industries to provide information to various users at different levels. Depending on individual needs, the transformed data is stored in various different systems. Sending operational data to other systems for further processing is then required, and introduces much complexity to an existing information technology (IT) infrastructure. Although maintenance of additional hardware and software is one component, potential inconsistencies and individually managed refresh cycles are others. For decades, there was no simple and efficient way to perform data transformations on the source system of operational data. With IBM® DB2® Analytics Accelerator, DB2 for z/OS is now in a unique position to complete these transformations in an efficient and well-performing way. DB2 for z/OS completes these while connecting to the same platform as for operational transactions, helping you to minimize your efforts to manage existing IT infrastructure. Real-time analytics on incoming operational transactions is another demand. Creating a comprehensive scoring model to detect specific patterns inside your data can easily require multiple iterations and multiple hours to complete. By enabling a first set of analytical functionality in DB2 Analytics Accelerator, those dedicated mining algorithms can now be run on an accelerator to efficiently perform these modeling tasks. Given the speed of query processing on an accelerator, these modeling tasks can now be performed much quicker compared to traditional relational database management systems. This speed enables you to keep your scoring algorithms more up-to-date, and ultimately adapt more quickly to constantly changing customer behaviors. This IBM

Redbooks® publication describes the new table type that is introduced with DB2 Analytics Accelerator V4.1 PTF5 that enables more efficient data transformations. These tables are called accelerator-only tables, and can exist on an accelerator only. The tables benefit from the accelerator performance characteristics, while maintaining access through existing DB2 for z/OS application programming interfaces (APIs). Additionally, we describe the newly introduced analytical capabilities with DB2 Analytics Accelerator V5.1, putting you in the position to efficiently perform data modeling for online analytical requirements in your DB2 for z/OS environment. This book is intended for technical decision-makers who want to get a broad understanding about the analytical capabilities and accelerator-only tables of DB2 Analytics Accelerator. In addition, you learn about how these capabilities can be used to accelerate in-database transformations and in-database analytics in various environments and scenarios, including the following scenarios: Multi-step processing and reporting in IBM DB2 Query Management FacilityTM, IBM Campaign, or Microstrategy environments In-database transformations using IBM InfoSphere® DataStage® Ad hoc data analysis for data scientists In-database analytics using IBM SPSS® Modeler

IBM Cloud Private System Administrator's Guide

Between the high-level concepts of business intelligence and the nitty-gritty instructions for using vendors' tools lies the essential, yet poorly-understood layer

of architecture, design and process. Without this knowledge, Big Data is belittled – projects flounder, are late and go over budget. Business Intelligence Guidebook: From Data Integration to Analytics shines a bright light on an often neglected topic, arming you with the knowledge you need to design rock-solid business intelligence and data integration processes. Practicing consultant and adjunct BI professor Rick Sherman takes the guesswork out of creating systems that are cost-effective, reusable and essential for transforming raw data into valuable information for business decision-makers. After reading this book, you will be able to design the overall architecture for functioning business intelligence systems with the supporting data warehousing and data-integration applications. You will have the information you need to get a project launched, developed, managed and delivered on time and on budget – turning the deluge of data into actionable information that fuels business knowledge. Finally, you'll give your career a boost by demonstrating an essential knowledge that puts corporate BI projects on a fast-track to success. Provides practical guidelines for building successful BI, DW and data integration solutions. Explains underlying BI, DW and data integration design, architecture and processes in clear, accessible language. Includes the complete project development lifecycle that can be applied at large enterprises as well as at small to medium-sized businesses Describes best practices and pragmatic approaches so readers can put them into action. Companion website includes templates and examples, further discussion of key topics, instructor materials, and references to trusted industry sources.

IBM WebSphere Information Analyzer and Data Quality Assessment

Find the right big data solution for your business ororganization Big data management is one of the major challenges facingbusiness, industry, and not-for-profit organizations. Data setssuch as customer transactions for a mega-retailer, weather patternsmonitored by meteorologists, or social network activity can quicklyoutpace the capacity of traditional data management tools. If youneed to develop or manage big data solutions, you'll appreciate howthese four experts define, explain, and guide you through this newand often confusing concept. You'll learn what it is, why itmatters, and how to choose and implement solutions that work. Effectively managing big data is an issue of growing importanceto businesses, not-for-profit organizations, government, and ITprofessionals Authors are experts in information management, big data, and avariety of solutions Explains big data in detail and discusses how to select andimplement a solution, security concerns to consider, data storageand presentation issues, analytics, and much more Provides essential information in a no-nonsense,easy-to-understand style that is empowering Big Data For Dummies cuts through the confusion and helpsyou take charge of big data solutions for your organization.

Metadata Management with IBM InfoSphere Information Server

To make better informed business decisions, better serve clients, and increase operational efficiencies, you must be aware of changes to key data as they occur. In addition, you must enable the immediate delivery of this information to the people and processes that need to act upon it. This ability to sense and respond to data changes is fundamental to dynamic warehousing, master data management, and many other key initiatives. A major challenge in providing this type of environment is determining how to tie all the independent systems together and process the immense data flow requirements. IBM® InfoSphere® Change Data Capture (InfoSphere CDC) can respond to that challenge, providing programming-free data integration, and eliminating redundant data transfer, to minimize the impact on production systems. In this IBM Redbooks® publication, we show you examples of how InfoSphere CDC can be used to implement integrated systems, to keep those systems updated immediately as changes occur, and to use your existing infrastructure and scale up as your workload grows. InfoSphere CDC can also enhance your investment in other software, such as IBM DataStage® and IBM QualityStage®, IBM InfoSphere Warehouse, and IBM InfoSphere Master Data Management Server, enabling real-time and event-driven processes. Enable the integration of your critical data and make it immediately available as your business needs it.

IBM Infosphere Datastage

XML is one of the most common standards for the exchange of information. However, organizations find challenges in how to address the complexities of dealing with hierarchical data types, particularly as they scale to gigabytes and beyond. In this IBM® Redbooks® publication, we discuss and describe the new capabilities in IBM InfoSphere® DataStage® 8.5. These capabilities enable developers to more easily manage the design and processing requirements presented by the most challenging XML sources. Developers can use these capabilities to create powerful hierarchical transformations and to parse and compose XML data with high performance and scalability. Spanning both batch and real-time run times, these capabilities can be used to solve a broad range of business requirements. As part of the IBM InfoSphere Information Server 8.5 release, InfoSphere DataStage was enhanced with new hierarchical transformation capabilities called . XML Stage provides native XML schema support and powerful XML transformation functionality. These capabilities are based on a unique state-of-the-art technology that allows you to parse and compose any complex XML structure from and to a relational form, as well as to a separate hierarchical form. This book is targeted at an audience of systems designers and developers who focus on implementing XML integration support in their environments.

Accelerating Data Transformation with IBM DB2 Analytics Accelerator for z/OS

In this IBM® Redbooks® publication, we present guidelines for the development of highly efficient and scalable information integration applications with InfoSphereTM DataStage® (DS) parallel jobs. InfoSphere DataStage is at the core of IBM Information Server, providing components that yield a high degree of freedom. For any particular problem there might be multiple solutions, which tend to be influenced by personal preferences, background, and previous experience. All too often, those solutions yield less than optimal, and non-scalable, implementations. This book includes a comprehensive detailed description of the components available, and descriptions on how to use them to obtain scalable and efficient solutions, for both batch and real-time scenarios. The advice provided in this document is the result of the combined proven experience from a number of expert practitioners in the field of high performance information integration, evolved over several years. This book is intended for IT architects, Information Management specialists, and Information Integration specialists responsible for delivering cost-effective IBM InfoSphere DataStage performance on all platforms.

Metadata Management with IBM InfoSphere Information Server

What do you know about your data? And how do you know what you know about your data? Information governance initiatives address corporate concerns about the quality and reliability of information in planning and decision-making processes. Metadata management refers to the tools, processes, and environment

that are provided so that organizations can reliably and easily share, locate, and retrieve information from these systems. Enterprise-wide information integration projects integrate data from these systems to one location to generate required reports and analysis. During this type of implementation process, metadata management must be provided along each step to ensure that the final reports and analysis are from the right data sources, are complete, and have quality. This IBM® Redbooks® publication introduces the information governance initiative and highlights the immediate needs for metadata management. It explains how IBM InfoSphereTM Information Server provides a single unified platform and a collection of product modules and components so that organizations can understand, cleanse, transform, and deliver trustworthy and context-rich information. It describes a typical implementation process. It explains how InfoSphere Information Server provides the functions that are required to implement such a solution and, more importantly, to achieve metadata management. This book is for business leaders and IT architects with an overview of metadata management in information integration solution space. It also provides key technical details that IT professionals can use in a solution planning, design, and implementation process.

Data Warehousing with the Informix Dynamic Server

IBM® Informix® is a low-administration, easy-to-use, and embeddable database that is ideal for application development. It supports a wide range of development

platforms, such as JavaTM, .NET, PHP, and web services, enabling developers to build database applications in the language of their choice. Informix is designed to handle RDBMS data and XML without modification and can be extended easily to handle new data sets. This IBM Redbooks® publication provides fundamentals of Informix application development. It covers the Informix Client installation and configuration for application development environments. It discusses the skills and techniques for building Informix applications with Java, ESQL/C, OLE DB, .NET, PHP, Ruby on Rails, DataBlade®, and Hibernate. The book uses code examples to demonstrate how to develop an Informix application with various drivers, APIs, and interfaces. It also provides application development troubleshooting and considerations for performance. This book is intended for developers who use IBM Informix for application development. Although some of the topics that we discuss are highly technical, the information in the book might also be helpful for managers or database administrators who are looking to better understand their Informix development environment.

IBM Watson Content Analytics: Discovering Actionable Insight from Your Content

IBM® InfoSphere® Guardium® provides the simplest, most robust solution for data security and data privacy by assuring the integrity of trusted information in

your data center. InfoSphere Guardium helps you reduce support costs by automating the entire compliance auditing process across heterogeneous environments. InfoSphere Guardium offers a flexible and scalable solution to support varying customer architecture requirements. This IBM Redbooks® publication provides a guide for deploying the Guardium solutions. This book also provides a roadmap process for implementing an InfoSphere Guardium solution that is based on years of experience and best practices that were collected from various Guardium experts. We describe planning, installation, configuration, monitoring, and administrating an InfoSphere Guardium environment. We also describe use cases and how InfoSphere Guardium integrates with other IBM products. The guidance can help you successfully deploy and manage an IBM InfoSphere Guardium system. This book is intended for the system administrators and support staff who are responsible for deploying or supporting an InfoSphere Guardium environment.

Civil Engineer's Reference Book

Boost your Big Data IQ! Gain insight into how to govern and consume IBM's unique in-motion and at-rest Big Data analytic capabilities Big Data represents a new era of computing—an inflection point of opportunity where data in any format may be explored and utilized for breakthrough insights—whether that data is in-place, in-motion, or at-rest. IBM is uniquely positioned to help clients navigate this

transformation. This book reveals how IBM is infusing open source Big Data technologies with IBM innovation that manifest in a platform capable of "changing the game." The four defining characteristics of Big Data—volume, variety, velocity, and veracity—are discussed. You'll understand how IBM is fully committed to Hadoop and integrating it into the enterprise. Hear about how organizations are taking inventories of their existing Big Data assets, with search capabilities that help organizations discover what they could already know, and extend their reach into new data territories for unprecedented model accuracy and discovery. In this book you will also learn not just about the technologies that make up the IBM Big Data platform, but when to leverage its purpose-built engines for analytics on data in-motion and data at-rest. And you'll gain an understanding of how and when to govern Big Data, and how IBM's industry-leading InfoSphere integration and governance portfolio helps you understand, govern, and effectively utilize Big Data. Industry use cases are also included in this practical guide.

IBM Technical Computing Clouds

This IBM® Redbooks® publication is intended for business leaders and IT architects who are responsible for building and extending their data warehouse and Business Intelligence infrastructure. It provides an overview of powerful new capabilities of Information Server in the areas of big data, statistical models, data governance and data quality. The book also provides key technical details that IT

professionals can use in solution planning, design, and implementation.

IBM InfoSphere Replication Server and Data Event Publisher

Implementing an InfoSphere Optim Data Growth Solution

This IBM® Redbooks® publication presents a development approach for master data management projects, and in particular, those projects based on IBM InfoSphere® MDM Server. The target audience for this book includes Enterprise Architects, Information, Integration and Solution Architects and Designers, Developers, and Product Managers. Master data management combines a set of processes and tools that defines and manages the non-transactional data entities of an organization. Master data management can provide processes for collecting, consolidating, persisting, and distributing this data throughout an organization. IBM InfoSphere Master Data Management Server creates trusted views of master data that can improve applications and business processes. You can use it to gain control over business information by managing and maintaining a complete and accurate view of master data. You also can use InfoSphere MDM Server to extract maximum value from master data by centralizing multiple data domains. InfoSphere MDM Server provides a comprehensive set of prebuilt business services

that support a full range of master data management functionality.

InfoSphere DataStage Parallel Framework Standard Practices

The IBM Informix® Dynamic Server (IDS) has the tools to build a powerful data warehouse infrastructure platform to lower costs and increase profits by doing more with your existing operational data and infrastructure. The Informix Warehouse Feature simplifies the process for design and deployment of a high performance data warehouse. With a state-of-the-art extract, load, and transform (ELT) tool and an Eclipse-based GUI environment that is easy to use, this comprehensive platform provides the foundation you need to cost effectively build and deploy the data warehousing infrastructure, using the IBM Informix Dynamic Server, and needed to enable the development and use of next-generation analytic solutions . This IBM® Redbooks® publication describes the technical information and demonstrates the functions and capabilities of the Informix Dynamic Server Warehouse Feature. It can help you understand how to develop a data warehousing architecture and infrastructure to meet your particular requirements, with the Informix Dynamic Server. It can also enable you to transform and manage your operational data, and use it to populate your data warehouse. With that new data warehousing environment, you can support the data analysis and decision-making that are required as you monitor and manage your business processes, and help you meet your business performance management goals, objectives, and

measurements.

IBM Information Server: Integration and Governance for Emerging Data Warehouse Demands

The IBM® Smart Analytics System is a fully-integrated and scalable data warehouse solution that combines software, server, and storage resources to offer optimal business intelligence and information management performance for enterprises. This IBM Redbooks® publication introduces the architecture and components of the IBM Smart Analytics System family. We describe the installation and configuration of the IBM Smart Analytics System and show how to manage the systems effectively to deliver an enterprise class service. This book explains the importance of integrating the IBM Smart Analytics System with the existing IT environment, as well as how to leverage investments in security, monitoring, and backup infrastructure. We discuss the monitoring tools for both operating systems and DB2®. Advance configuration, performance troubleshooting, and tuning techniques are also discussed. This book is targeted at the architects and specialists who need to know the concepts and the detailed instructions for a successful Smart Analytics System implementation and operation.

Python for Data Analysis

Presents case studies and instructions on how to solve data analysis problems using Python.

Implementing an InfoSphere Optim Data Growth Solution

This easy IBM InfoSphere DataStage self-assessment will make you the established IBM InfoSphere DataStage domain master by revealing just what you need to know to be fluent and ready for any IBM InfoSphere DataStage challenge. How do I reduce the effort in the IBM InfoSphere DataStage work to be done to get problems solved? How can I ensure that plans of action include every IBM InfoSphere DataStage task and that every IBM InfoSphere DataStage outcome is in place? How will I save time investigating strategic and tactical options and ensuring IBM InfoSphere DataStage opportunity costs are low? How can I deliver tailored IBM InfoSphere DataStage advise instantly with structured going-forward plans? There's no better guide through these mind-expanding questions than acclaimed best-selling author Gerard Blokdyk. Blokdyk ensures all IBM InfoSphere DataStage essentials are covered, from every angle: the IBM InfoSphere DataStage self-assessment shows succinctly and clearly that what needs to be clarified to organize the business/project activities and processes so that IBM InfoSphere DataStage outcomes are achieved. Contains extensive criteria grounded in past and current successful projects and activities by experienced IBM InfoSphere DataStage practitioners. Their mastery, combined with the uncommon elegance of

the self-assessment, provides its superior value to you in knowing how to ensure the outcome of any efforts in IBM InfoSphere DataStage are maximized with professional results. Your purchase includes access to the $249 value IBM InfoSphere DataStage self-assessment dashboard download which gives you your dynamically prioritized projects-ready tool and shows your organization exactly what to do next. Your exclusive instant access details can be found in your book.

BMC Control-M 7

In this IBM® Redbooks® publication, we present guidelines for the development of highly efficient and scalable information integration applications with InfoSphereTM DataStage® (DS) parallel jobs. InfoSphere DataStage is at the core of IBM Information Server, providing components that yield a high degree of freedom. For any particular problem there might be multiple solutions, which tend to be influenced by personal preferences, background, and previous experience. All too often, those solutions yield less than optimal, and non-scalable, implementations. This book includes a comprehensive detailed description of the components available, and descriptions on how to use them to obtain scalable and efficient solutions, for both batch and real-time scenarios. The advice provided in this document is the result of the combined proven experience from a number of expert practitioners in the field of high performance information integration, evolved over several years. This book is intended for IT architects, Information

Management specialists, and Information Integration specialists responsible for delivering cost-effective IBM InfoSphere DataStage performance on all platforms.

IBM InfoSphere DataStage Data Flow and Job Design

This Cookbook contains step-by-step instructions for Report Studio 10.1 users to author effective reports. The book is designed in a way that you can refer to it chapter by chapter, look at the list of recipes and read them in no particular order.The Cognos 10.1 Report Studio Cookbook is for you if you are a Business Intelligence Developer who is working on IBM Cognos 10 Report Studio and wants to author impressive reports by putting to use what this tool has to offer. It is also ideal you are a Business Analyst or Power User who authors his own reports and wants to look beyond the conventional features of IBM Cognos 10 Report Studio.This book assumes that you are familiar with the architecture of IBM Cognos 10. You should also have basic knowledge of IBM Cognos Report Studio and can do the basic report authoring tasks.

DB2 pureXML Cookbook

This IBM® Redbooks® publication highlights IBM Technical Computing as a flexible infrastructure for clients looking to reduce capital and operational expenditures,

optimize energy usage, or re-use the infrastructure. This book strengthens IBM SmartCloud® solutions, in particular IBM Technical Computing clouds, with a well-defined and documented deployment model within an IBM System x® or an IBM Flex SystemTM. This provides clients with a cost-effective, highly scalable, robust solution with a planned foundation for scaling, capacity, resilience, optimization, automation, and monitoring. This book is targeted toward technical professionals (consultants, technical support staff, IT Architects, and IT Specialists) responsible for providing cloud-computing solutions and support.

Principles of Data Wrangling

Reviews planning and designing architecture and implementing the data warehouse. Includes discussions on how and why to apply IBM tools. Offers tips, tricks, and workarounds to ensure maximum performance. Companion Web site includes technical notes, product updates, corrections, and links to relevant material and training.

IBM Information Governance Solutions

IBM Information Server is a revolutionary new software platform that helps organizations derive more value from the complex heterogeneous information that

is spread across their systems. It enables organizations to integrate disparate data and deliver trusted information wherever and whenever needed, in line and in context, to specific people, applications, and processes. IBM WebSphere Information Analyzer is a data profiling and analysis tool that is a critical component of IBM Information Server. It is designed to help business and data analysts understand the content, quality, and structure of their data sources by automating the data discovery process. Bundled with IBM WebSphere Information Analyzer is AuditStage, a data rule monitoring tool that is designed to help business and data analysts validate data and assess ongoing data quality trends. This book describes a usage scenario that covers all dimensions of profiling, rule building, deployment, and quality monitoring through a data integration life cycle.

Monetizing Your Data

AI may be the greatest opportunity of our time, with the potential to add nearly $16 trillion to the global economy over the next decade. But so far adoption has been much slower than anticipated. With this practical report, business leaders will discover where they are in their AI journey and learn the steps they still need to take to implement an AI solution in their organization. Author Rob Thomas, general manager of data and AI at IBM, introduces C-suite executives and line-of-business professionals to the AI Ladder--a unified, prescriptive approach to help them accelerate their AI journey, no matter where they are in the process. This succinct

report explores AI's drivers, value, and opportunity, as well as the adoption challenges organizations now face. Learn why you can't have AI without IA (information architecture) Unify data within a multicloud environment to unlock its value Make data accessible by collecting every type, regardless of where it lives Create a business-ready analytics foundation with built-in governance, protection, and compliance Benefit from AI models by learning smarter ways to analyze data Operationalize AI across multiple departments and within various processes

IBM Informix Developer's Handbook

Transforming data into revenue generating strategies and actions Organizations are swamped with data—collected from web traffic, point of sale systems, enterprise resource planning systems, and more, but what to do with it? Monetizing your Data provides a framework and path for business managers to convert ever-increasing volumes of data into revenue generating actions through three disciplines: decision architecture, data science, and guided analytics. There are large gaps between understanding a business problem and knowing which data is relevant to the problem and how to leverage that data to drive significant financial performance. Using a proven methodology developed in the field through delivering meaningful solutions to Fortune 500 companies, this book gives you the analytical tools, methods, and techniques to transform data you already have into information into insights that drive winning decisions. Beginning with an

explanation of the analytical cycle, this book guides you through the process of developing value generating strategies that can translate into big returns. The companion website, www.monetizingyourdata.com, provides templates, checklists, and examples to help you apply the methodology in your environment, and the expert author team provides authoritative guidance every step of the way. This book shows you how to use your data to: Monetize your data to drive revenue and cut costs Connect your data to decisions that drive action and deliver value Develop analytic tools to guide managers up and down the ladder to better decisions Turning data into action is key; data can be a valuable competitive advantage, but only if you understand how to organize it, structure it, and uncover the actionable information hidden within it through decision architecture and guided analytics. From multinational corporations to single-owner small businesses, companies of every size and structure stand to benefit from these tools, methods, and techniques; Monetizing your Data walks you through the translation and transformation to help you leverage your data into value creating strategies.

IBM InfoSphere Information Server Installation and Configuration Guide

This IBM® RedpaperTM publication provides suggestions, hints and tips, directions,

installation steps, checklists of prerequisites, and configuration information collected from several IBM InfoSphere® Information Server experts. It is intended to minimize the time required to successfully install and configure InfoSphere Information Server. The information in this document is based on field experiences of experts who have implemented InfoSphere Information Server. As such, it is intended to supplement, and not replace, the product documentation. Discover the proven choices and combinations for installing InfoSphere Information Server that have been the most successful for the IBM InfoSphere Center Of Excellence. This paper includes a broad range of customer needs and experiences, with a focus on the following areas: InfoSphere Information Server architecture Checklists Prerequisites Configuration choices that work well together This paper is based on thousands of hours of production systems experience, from which you can now reap significant benefits.

Harness the Power of Big Data The IBM Big Data Platform

Design, implement, and monitor a successful Q replication and Event Publishing project with IBM InfoSphere Replication Server and Data Event Publisher using this book and eBook.

Migrating from Microsoft SQL Server to IBM Informix

Big data is currently one of the most critical emerging technologies. Organizations around the world are looking to exploit the explosive growth of data to unlock previously hidden insights in the hope of creating new revenue streams, gaining operational efficiencies, and obtaining greater understanding of customer needs. It is important to think of big data and analytics together. Big data is the term used to describe the recent explosion of different types of data from disparate sources. Analytics is about examining data to derive interesting and relevant trends and patterns, which can be used to inform decisions, optimize processes, and even drive new business models. With today's deluge of data comes the problems of processing that data, obtaining the correct skills to manage and analyze that data, and establishing rules to govern the data's use and distribution. The big data technology stack is ever growing and sometimes confusing, even more so when we add the complexities of setting up big data environments with large up-front investments. Cloud computing seems to be a perfect vehicle for hosting big data workloads. However, working on big data in the cloud brings its own challenge of reconciling two contradictory design principles. Cloud computing is based on the concepts of consolidation and resource pooling, but big data systems (such as Hadoop) are built on the shared nothing principle, where each node is independent and self-sufficient. A solution architecture that can allow these mutually exclusive principles to coexist is required to truly exploit the elasticity and ease-of-use of cloud computing for big data environments. This IBM® RedpaperTM publication is aimed at chief architects, line-of-business executives, and CIOs to provide an

understanding of the cloud-related challenges they face and give prescriptive guidance for how to realize the benefits of big data solutions quickly and cost-effectively.

InfoSphere DataStage for Enterprise XML Data Integration

Today, organizations face tremendous challenges with data explosion and information governance. InfoSphereTM OptimTM solutions solve the data growth problem at the source by managing the enterprise application data. The Optim Data Growth solutions are consistent, scalable solutions that include comprehensive capabilities for managing enterprise application data across applications, databases, operating systems, and hardware platforms. You can align the management of your enterprise application data with your business objectives to improve application service levels, lower costs, and mitigate risk. In this IBM® Redbooks® publication, we describe the IBM InfoSphere Optim Data Growth solutions and a methodology that provides implementation guidance from requirements analysis through deployment and administration planning. We also discuss various implementation topics including system architecture design, sizing, scalability, security, performance, and automation. This book is intended to provide various systems development professionals, Data Solution Architects, Data Administrators, Modelers, Data Analysts, Data Integrators, or anyone who has to analyze or integrate data structures, a broad understanding about IBM InfoSphere

Optim Data Growth solutions. By being used in conjunction with the product manuals and online help, this book provides guidance about implementing an optimal solution for managing your enterprise application data.

InfoSphere DataStage for Enterprise XML Data Integration

What do you know about your data? And how do you know what you know about your data? Information governance initiatives address corporate concerns about the quality and reliability of information in planning and decision-making processes. Metadata management refers to the tools, processes, and environment that are provided so that organizations can reliably and easily share, locate, and retrieve information from these systems. Enterprise-wide information integration projects integrate data from these systems to one location to generate required reports and analysis. During this type of implementation process, metadata management must be provided along each step to ensure that the final reports and analysis are from the right data sources, are complete, and have quality. This IBM® Redbooks® publication introduces the information governance initiative and highlights the immediate needs for metadata management. It explains how IBM InfoSphereTM Information Server provides a single unified platform and a collection of product modules and components so that organizations can understand, cleanse, transform, and deliver trustworthy and context-rich information. It describes a typical implementation process. It explains how InfoSphere Information

Server provides the functions that are required to implement such a solution and, more importantly, to achieve metadata management. This book is for business leaders and IT architects with an overview of metadata management in information integration solution space. It also provides key technical details that IT professionals can use in a solution planning, design, and implementation process.

IBM Cognos 10 Report Studio Cookbook

IBM® Cloud Private is an application platform for developing and managing containerized applications across hybrid cloud environments, on-premises and public clouds. It is an integrated environment for managing containers that includes the container orchestrator Kubernetes, a private image registry, a management console, and monitoring frameworks. This IBM Redbooks covers tasks performed by IBM Cloud Private system administrators such as installation for high availability, configuration, backup and restore, using persistent volumes, networking, security, logging and monitoring. Istio integration, troubleshooting and so on. As part of this project we also developed several code examples and you can download those from the IBM Redbooks GitHub location: https://github.com/IBMRedbooks. The authors team has many years of experience in implementing IBM Cloud Private and other cloud solutions in production environments, so throughout this document we took the approach of providing you the recommended practices in those areas. If you are an IBM Cloud Private system

administrator, this book is for you. If you are developing applications on IBM Cloud Private, you can see the IBM Redbooks publication IBM Cloud Private Application Developer's Guide, SG24-8441.

Building Big Data and Analytics Solutions in the Cloud

IBM Data Warehousing

Master one of the world?s most powerful enterprise workload automation tools? BMC Control-M 7 - using this book and eBook.

Big Data For Dummies

Managing information within the enterprise has always been a vital and important task to support the day-to-day business operations and to enable analysis of that data for decision making to better manage and grow the business for improved profitability. To do all that, clearly the data must be accurate and organized so it is accessible and understandable to all who need it. That task has grown in importance as the volume of enterprise data has been growing significantly (analyst estimates of 40 - 50% growth per year are not uncommon) over the years.

However, most of that data has been what we call "structured" data, which is the type that can fit neatly into rows and columns and be more easily analyzed. Now we are in the era of "big data." This significantly increases the volume of data available, but it is in a form called "unstructured" data. That is, data from sources that are not as easily organized, such as data from emails, spreadsheets, sensors, video, audio, and social media sites. There is valuable information in all that data but it calls for new processes to enable it to be analyzed. All this has brought with it a renewed and critical need to manage and organize that data with clarity of meaning, understandability, and interoperability. That is, you must be able to integrate this data when it is from within an enterprise but also importantly when it is from many different external sources. What is described here has been and is being done to varying extents. It is called "information governance." Governing this information however has proven to be challenging. But without governance, much of the data can be less useful and perhaps even used incorrectly, significantly impacting enterprise decision making. So we must also respect the needs for information security, consistency, and validity or else suffer the potential economic and legal consequences. Implementing sound governance practices needs to be an integral part of the information control in our organizations. This IBM® Redbooks® publication focuses on the building blocks of a solid governance program. It examines some familiar governance initiative scenarios, identifying how they underpin key governance initiatives, such as Master Data Management, Quality Management, Security and Privacy, and Information Lifecycle Management. IBM

Information Management and Governance solutions provide a comprehensive suite to help organizations better understand and build their governance solutions. The book also identifies new and innovative approaches that are developed by IBM practice leaders that can help as you implement the foundation capabilities in your organizations.

ROMANCE ACTION & ADVENTURE MYSTERY & THRILLER BIOGRAPHIES & HISTORY CHILDREN'S YOUNG ADULT FANTASY HISTORICAL FICTION HORROR LITERARY FICTION NON-FICTION SCIENCE FICTION