agile for data science

Agile for Data Science: A Narrative of Iterative Success

Author: Dr. Evelyn Reed, PhD in Data Science, PMP certified, experienced Agile coach.

Publisher: Data Science Digest, a leading publisher of insightful articles and resources for data scientists and professionals in the field of data analytics.

Editor: Sarah Chen, MA in Journalism, specializing in technical writing and editing.

Keyword: agile for data science

Abstract: This article explores the transformative power of agile methodologies in the realm of data science. Through personal anecdotes, real-world case studies, and practical advice, we demonstrate how embracing agile for data science can lead to more efficient, collaborative, and ultimately, more successful projects. We delve into the challenges inherent in traditional data science approaches and showcase how agile principles offer a robust solution.

The Data Science Struggle: A Waterfall of Woes

My early career in data science was, to put it mildly, chaotic. We followed the traditional waterfall model: months of requirements gathering, followed by extensive modeling, then deployment – often to discover the model was addressing the wrong problem, or the data was insufficient. One project, predicting customer churn for a telecommunications company, took over a year. By the time we delivered the model, the customer churn patterns had significantly shifted, rendering much of our work obsolete. This experience, unfortunately, wasn't unique. Many data science projects fall victim to this "waterfall of woes," characterized by lengthy timelines, inflexible plans, and a high risk of failure.

Embracing Agile for Data Science: A Shift in Paradigm

Then, I discovered agile for data science. The shift was revolutionary. Agile, with its emphasis on iterative development, collaboration, and continuous feedback, offered a path to more efficient and adaptable data science projects. Instead of one massive, monolithic effort, agile projects are broken down into smaller, manageable sprints. Each sprint culminates in a demonstrable increment of progress, allowing for continuous feedback and adjustments along the way. This dramatically reduces the risk of building the wrong thing and allows for course correction early on.

Case Study 1: Fraud Detection with Agile

At my current company, we implemented agile for data science in our fraud detection system. Instead of spending months building a complex model, we started with a Minimum Viable Product (MVP) in just two weeks. This MVP, while basic, already identified a significant portion of fraudulent transactions. Each subsequent sprint incorporated user feedback, improved model accuracy, and added new features. The result? A more effective, adaptable, and cost-effective fraud detection system, deployed in a fraction of the time it would have taken using traditional methods.

Case Study 2: Personalized Recommendations – Agile's Adaptability

Another project involved creating a personalized recommendation engine for an e-commerce platform. Using agile for data science, we initially focused on a simple collaborative filtering algorithm. Through sprint reviews, we discovered that users valued diversity in recommendations more than strict personalization. This feedback allowed us to pivot and incorporate content-based filtering into our model, significantly improving user engagement. This flexibility, a hallmark of agile, would have been impossible with a rigid, pre-planned approach.

The Principles of Agile for Data Science

Agile for data science isn’t simply about adopting Scrum or Kanban; it’s about embedding agile principles into the data science lifecycle. This includes:

Iterative Development: Breaking down projects into small, manageable sprints.
Continuous Feedback: Regularly soliciting feedback from stakeholders and users.
Collaboration: Fostering close collaboration between data scientists, engineers, and business stakeholders.
Adaptability: Embracing change and adjusting plans as needed based on

feedback and new information.
Transparency: Maintaining open communication and visibility into project progress.

Overcoming the Challenges of Agile for Data Science

Implementing agile for data science isn't without its challenges. Data scientists often struggle with the seemingly unpredictable nature of agile, especially concerning data exploration and model development, which can be inherently less predictable than coding tasks. However, techniques like sprint planning poker and user story mapping can help mitigate these challenges, ensuring that sprint goals are realistic and attainable.

Conclusion

Agile for data science is no longer a niche methodology; it's a necessity for organizations striving for efficient, effective, and impactful data science initiatives. By embracing iterative development, continuous feedback, and strong collaboration, data science teams can significantly improve project outcomes, reduce risks, and deliver valuable insights faster and more effectively. The transition requires a cultural shift and a willingness to adapt, but the rewards—in terms of speed, quality, and overall success—are undeniable.

FAQs

1. What is the difference between traditional data science and agile data science? Traditional data science often follows a waterfall approach, with lengthy phases and limited flexibility, whereas agile data science embraces iterative development and continuous feedback.

1. Which agile framework is best for data science? Scrum and Kanban are both popular choices, and the best framework depends on the specific project and team dynamics.

1. How do you handle unpredictable data exploration in an agile sprint? Allocate time for exploration in each sprint, and use techniques like

sprint planning poker to estimate the effort involved.

1. How do you measure success in agile data science projects? Use metrics such as model accuracy, business impact, and user satisfaction, rather than solely focusing on code completion.

1. What are the common pitfalls of agile for data science? Lack of clear communication, insufficient stakeholder involvement, and resistance to change are common pitfalls.

1. How can I introduce agile to my existing data science team? Start with a pilot project, provide training, and foster a culture of collaboration and experimentation.

1. What tools can support agile for data science? Jira, Trello, and Azure DevOps are popular tools for managing agile projects.

1. How does agile for data science improve data quality? The iterative process allows for early identification and correction of data quality issues.

1. Is agile for data science suitable for all projects? While agile is highly effective for many projects, its suitability depends on the project's complexity, size, and the team's experience with agile methodologies.

Related Articles

1. "Data Science Sprints: A Practical Guide to Agile Development": This article provides a step-by-step guide to implementing agile sprints in data science projects.

1. "Agile Data Science: Overcoming the Challenges of Data Exploration": This article focuses on addressing the challenges of incorporating unpredictable data exploration into agile sprints.

1. "Measuring Success in Agile Data Science Projects: Key Metrics and KPIs": This article explores various metrics for measuring success in agile data science projects.

1. "Kanban for Data Science: A Flexible Approach to Agile Development": This article highlights the benefits of using the Kanban framework for data science projects.

1. "The Role of Collaboration in Agile Data Science Teams": This article emphasizes the importance of collaboration among data scientists, engineers, and stakeholders in agile environments.

1. "Building a Minimum Viable Product (MVP) in Data Science": This guide explains how to create an MVP using agile principles to test assumptions and validate ideas early.

1. "Agile Data Science and Data Governance: A Balancing Act": This article explores the intersection of agile practices and data governance in the data science field.

1. "Scaling Agile for Large Data Science Projects": This piece provides strategies for effectively implementing agile in larger, more complex data science initiatives.

1. "Agile for Data Science: A Case Study in Financial Modeling": This article presents a detailed case study illustrating the application of agile methodologies in a financial modeling project.

agile for data science: Agile Data Science Russell Jurney, 2013-10-15 Mining big data requires a deep investment in people and time. How can you be sure you're building the right models? With this hands-on book, you'll learn a flexible toolset and methodology for building effective analytics applications with Hadoop. Using lightweight tools such as Python, Apache Pig, and the D3.js library, your team will create an agile environment for exploring data, starting with an example application to mine your own email inboxes. You'll learn an iterative approach that enables you to quickly change the kind of analysis you're doing, depending on what the data is telling you. All example code in this book is available as working Heroku apps. Create analytics applications by using the agile big data development methodology Build value from your data in a series of agile sprints, using the data-value stack Gain insight by using several data structures to extract multiple features from a single dataset Visualize data with charts, and expose different aspects through interactive reports Use historical data to predict the future, and translate predictions into action Get feedback from users after each sprint to keep your project on track

agile for data science: Agile Data Science 2.0 Russell Jurney, 2017-06-07 Data science teams looking to turn research into useful analytics applications require not only the right tools, but also the right approach if they're to succeed. With the revised second edition of this hands-on guide, up-and-coming data scientists will learn how to use the Agile Data Science development methodology to build data applications with Python, Apache Spark, Kafka, and other tools. Author Russell Jurney demonstrates how to compose a data platform for building, deploying, and refining analytics applications with Apache Kafka, MongoDB, ElasticSearch, d3.js, scikit-learn, and Apache Airflow. You'll learn an iterative approach that lets you quickly change the kind of analysis you're doing, depending on what the data is telling you. Publish data science work as a web application, and affect meaningful change in your organization. Build value from your data in a series of agile sprints, using the data-value pyramid Extract features for statistical models from a single dataset Visualize data with charts, and expose different aspects through interactive reports Use historical data to predict the future via classification and regression Translate predictions into actions Get feedback from users after each sprint to keep your project on track

agile for data science: *Practical DataOps* Harvinder Atwal, 2019-12-09 Gain a practical introduction to DataOps, a new discipline for delivering data science at scale inspired by practices at companies such as Facebook, Uber, LinkedIn, Twitter, and eBay. Organizations need more than the latest AI algorithms, hottest tools, and best people to turn data into insight-driven action and useful analytical data products. Processes and thinking employed to manage and use data in the 20th century are a bottleneck for working effectively with the variety of data and advanced analytical use cases that organizations have today. This book provides the approach and methods to ensure continuous rapid use of data to create analytical data products and steer decision making. Practical DataOps shows you how to optimize the data supply chain from diverse raw data sources to the final data product, whether the goal is a machine learning model or other data-orientated output. The book provides an approach to eliminate wasted effort and improve collaboration between data producers, data consumers, and the rest of the organization through the adoption of lean thinking and agile software development principles. This book helps you to improve the speed and accuracy of analytical application development through data management and DevOps practices that securely expand data access, and rapidly increase the number of reproducible data products through automation, testing, and integration. The book also shows how to collect feedback and monitor performance to manage and continuously improve your processes and output. What You Will LearnDevelop a data strategy for your organization to help it reach its long-term goals Recognize and eliminate barriers to delivering data to users at scale Work on the right things for the right stakeholders through agile collaboration Create trust in data via rigorous testing and effective data management Build a culture of learning and continuous improvement through monitoring

deployments and measuring outcomes Create cross-functional self-organizing teams focused on goals not reporting lines Build robust, trustworthy, data pipelines in support of AI, machine learning, and other analytical data products Who This Book Is For Data science and advanced analytics experts, CIOs, CDOs (chief data officers), chief analytics officers, business analysts, business team leaders, and IT professionals (data engineers, developers, architects, and DBAs) supporting data teams who want to dramatically increase the value their organization derives from data. The book is ideal for data professionals who want to overcome challenges of long delivery time, poor data quality, high maintenance costs, and scaling difficulties in getting data science output and machine learning into customer-facing production.

agile for data science: Agile Analytics Ken Collier, 2012 Using Agile methods, you can bring far greater innovation, value, and quality to any data warehousing (DW), business intelligence (BI), or analytics project. However, conventional Agile methods must be carefully adapted to address the unique characteristics of DW/BI projects. In Agile Analytics, Agile pioneer Ken Collier shows how to do just that. Collier introduces platform-agnostic Agile solutions for integrating infrastructures consisting of diverse operational, legacy, and specialty systems that mix commercial and custom code. Using working examples, he shows how to manage analytics development teams with widely diverse skill sets and how to support enormous and fast-growing data volumes. Collier's techniques offer optimal value whether your projects involve back-end data management, front-end business analysis, or both. Part I focuses on Agile project management techniques and delivery team coordination, introducing core practices that shape the way your Agile DW/BI project community can collaborate toward success Part II presents technical methods for enabling continuous delivery of business value at production-quality levels, including evolving superior designs; test-driven DW development; version control; and project automation Collier brings together proven solutions you can apply right now--whether you're an IT decision-maker, data warehouse professional, database administrator, business intelligence specialist, or database developer. With his help, you can mitigate project risk, improve business alignment, achieve better results--and have fun along the way.

agile for data science: *Agile Machine Learning* Eric Carter, Matthew Hurst, 2019-08-21 Build resilient applied machine learning teams that deliver better data products through adapting the guiding principles of the Agile Manifesto. Bringing together talented people to create a great applied machine learning team is no small feat. With developers and data scientists both contributing expertise in their respective fields, communication alone can be a challenge. Agile Machine Learning teaches you how to deliver superior data products through agile processes and to learn, by example, how to organize and manage a fast-paced team challenged with solving novel data problems at scale, in a production environment. The authors’ approach models the ground-breaking engineering principles described in the Agile Manifesto. The book provides further context, and contrasts the original principles with the requirements of systems that deliver a data product. What You'll Learn Effectively run a data engineering team that is metrics-focused, experiment-focused, and data-focused Make sound implementation and model exploration decisions based on the data and the metrics Know the importance of data wallowing: analyzing data in real time in a group setting Recognize the value of always being able to measure your current state objectively Understand data literacy, a key attribute of a reliable data engineer, from definitions to expectations Who This Book Is For Anyone who manages a machine learning team, or is responsible for creating production-ready inference components. Anyone responsible for data project workflow of sampling data; labeling, training, testing, improving, and maintaining models; and system and data metrics will also find this book useful. Readers should be familiar with software engineering and understand the basics of machine learning and working with data.

agile for data science: *Agile Machine Learning with DataRobot* Bipin Chadha, Sylvester Juwe, 2021-12-24 Leverage DataRobot's enterprise AI platform and automated decision intelligence to extract business value from data Key FeaturesGet well-versed with DataRobot features using real-world examplesUse this all-in-one platform to build, monitor, and deploy ML models for

handling the entire production life cycleMake use of advanced DataRobot capabilities to programmatically build and deploy a large number of ML modelsBook Description DataRobot enables data science teams to become more efficient and productive. This book helps you to address machine learning (ML) challenges with DataRobot's enterprise platform, enabling you to extract business value from data and rapidly create commercial impact for your organization. You'll begin by learning how to use DataRobot's features to perform data prep and cleansing tasks automatically. The book then covers best practices for building and deploying ML models, along with challenges faced while scaling them to handle complex business problems. Moving on, you'll perform exploratory data analysis (EDA) tasks to prepare your data to build ML models and ways to interpret results. You'll also discover how to analyze the model's predictions and turn them into actionable insights for business users. Next, you'll create model documentation for internal as well as compliance purposes and learn how the model gets deployed as an API. In addition, you'll find out how to operationalize and monitor the model's performance. Finally, you'll work with examples on time series forecasting, NLP, image processing, MLOps, and more using advanced DataRobot capabilities. By the end of this book, you'll have learned to use DataRobot's AutoML and MLOps features to scale ML model building by avoiding repetitive tasks and common errors. What you will learnUnderstand and solve business problems using DataRobotUse DataRobot to prepare your data and perform various data analysis tasks to start building modelsDevelop robust ML models and assess their results correctly before deploymentExplore various DataRobot functions and outputs to help you understand the models and select the one that best solves the business problemAnalyze a model's predictions and turn them into actionable insights for business usersUnderstand how DataRobot helps in governing, deploying, and maintaining ML modelsWho this book is for This book is for data scientists, data analysts, and data enthusiasts looking for a practical guide to building and deploying robust machine learning models using DataRobot. Experienced data scientists will also find this book helpful for rapidly exploring, building, and deploying a broader range of models. The book assumes a basic understanding of machine learning.

agile for data science: Agile Data Warehousing Project Management Ralph Hughes, 2012-12-28 You have to make sense of enormous amounts of data, and while the notion of agile data warehousing might sound tricky, it can yield as much as a 3-to-1 speed advantage while cutting project costs in half. Bring this highly effective technique to your organization with the wisdom of agile data warehousing expert Ralph Hughes. Agile Data Warehousing Project Management will give you a thorough introduction to the method as you would practice it in the project room to build a serious data mart. Regardless of where you are today, this step-by-step implementation guide will prepare you to join or even lead a team in visualizing, building, and validating a single component to an enterprise data warehouse. - Provides a thorough grounding on the mechanics of Scrum as well as practical advice on keeping your team on track - Includes strategies for getting accurate and actionable requirements from a team's business partner - Revolutionary estimating techniques that make forecasting labor far more understandable and accurate - Demonstrates a blends of Agile methods to simplify team management and synchronize inputs across IT specialties - Enables you and your teams to start simple and progress steadily to world-class performance levels

agile for data science: Agile Data Warehouse Design Lawrence Corr, Jim Stagnitto, 2011-11 Agile Data Warehouse Design is a step-by-step guide for capturing data warehousing/business intelligence (DW/BI) requirements and turning them into high performance dimensional models in the most direct way: by modelstorming (data modeling + brainstorming) with BI stakeholders. This book describes BEAM□, an agile approach to dimensional modeling, for improving communication between data warehouse designers, BI stakeholders and the whole DW/BI development team. BEAM□ provides tools and techniques that will encourage DW/BI designers and developers to move away from their keyboards and entity relationship based tools and model interactively with their colleagues. The result is everyone thinks dimensionally from the outset! Developers understand how to efficiently implement dimensional modeling solutions. Business stakeholders feel ownership of the data warehouse they have created, and can already imagine how they will use it to answer their

business questions. Within this book, you will learn: □ Agile dimensional modeling using Business Event Analysis & Modeling (BEAM□) □ Modelstorming: data modeling that is quicker, more inclusive, more productive, and frankly more fun! □ Telling dimensional data stories using the 7Ws (who, what, when, where, how many, why and how) □ Modeling by example not abstraction; using data story themes, not crow's feet, to describe detail □ Storyboarding the data warehouse to discover conformed dimensions and plan iterative development □ Visual modeling: sketching timelines, charts and grids to model complex process measurement - simply □ Agile design documentation: enhancing star schemas with BEAM□ dimensional shorthand notation □ Solving difficult DW/BI performance and usability problems with proven dimensional design patterns Lawrence Corr is a data warehouse designer and educator. As Principal of DecisionOne Consulting, he helps clients to review and simplify their data warehouse designs, and advises vendors on visual data modeling techniques. He regularly teaches agile dimensional modeling courses worldwide and has taught dimensional DW/BI skills to thousands of students. Jim Stagnitto is a data warehouse and master data management architect specializing in the healthcare, financial services, and information service industries. He is the founder of the data warehousing and data mining consulting firm Llumino.

agile for data science: *Agile Data Warehousing* Ralph Hughes, 2008-07-14 Contains a six-stage plan for starting new warehouse projects and guiding programmers step-by-step until they become a world-class, Agile development team. It describes also how to avoid or contain the fierce opposition that radically new methods can encounter from the traditionally-minded IS departments found in many large companies.

agile for data science: How to Lead in Data Science Jike Chong, Yue Cathy Chang, 2021-12-28 A field guide for the unique challenges of data science leadership, filled with transformative insights, personal experiences, and industry examples. In How To Lead in Data Science you will learn: Best practices for leading projects while balancing complex trade-offs Specifying, prioritizing, and planning projects from vague requirements Navigating structural challenges in your organization Working through project failures with positivity and tenacity Growing your team with coaching, mentoring, and advising Crafting technology roadmaps and championing successful projects Driving diversity, inclusion, and belonging within teams Architecting a long-term business strategy and data roadmap as an executive Delivering a data-driven culture and structuring productive data science organizations How to Lead in Data Science is full of techniques for leading data science at every seniority level—from heading up a single project to overseeing a whole company's data strategy. Authors Jike Chong and Yue Cathy Chang share hard-won advice that they've developed building data teams for LinkedIn, Acorns, Yiren Digital, large asset-management firms, Fortune 50 companies, and more. You'll find advice on plotting your long-term career advancement, as well as quick wins you can put into practice right away. Carefully crafted assessments and interview scenarios encourage introspection, reveal personal blind spots, and highlight development areas. About the technology Lead your data science teams and projects to success! To make a consistent, meaningful impact as a data science leader, you must articulate technology roadmaps, plan effective project strategies, support diversity, and create a positive environment for professional growth. This book delivers the wisdom and practical skills you need to thrive as a data science leader at all levels, from team member to the C-suite. About the book How to Lead in Data Science shares unique leadership techniques from high-performance data teams. It’s filled with best practices for balancing project trade-offs and producing exceptional results, even when beginning with vague requirements or unclear expectations. You’ll find a clearly presented modern leadership framework based on current case studies, with insights reaching all the way to Aristotle and Confucius. As you read, you’ll build practical skills to grow and improve your team, your company’s data culture, and yourself. What's inside How to coach and mentor team members Navigate an organization’s structural challenges Secure commitments from other teams and partners Stay current with the technology landscape Advance your career About the reader For data science practitioners at all levels. About the author Dr. Jike Chong and Yue Cathy Chang build, lead, and grow high-performing data teams across

industries in public and private companies, such as Acorns, LinkedIn, large asset-management firms, and Fortune 50 companies. Table of Contents 1 What makes a successful data scientist? PART 1 THE TECH LEAD: CULTIVATING LEADERSHIP 2 Capabilities for leading projects 3 Virtues for leading projects PART 2 THE MANAGER: NURTURING A TEAM 4 Capabilities for leading people 5 Virtues for leading people PART 3 THE DIRECTOR: GOVERNING A FUNCTION 6 Capabilities for leading a function 7 Virtues for leading a function PART 4 THE EXECUTIVE: INSPIRING AN INDUSTRY 8 Capabilities for leading a company 9 Virtues for leading a company PART 5 THE LOOP AND THE FUTURE 10 Landscape, organization, opportunity, and practice 11 Leading in data science and a future outlook

agile for data science: Analytics Phil Simon, 2017-07-03 For years, organizations have struggled to make sense out of their data. IT projects designed to provide employees with dashboards, KPIs, and business-intelligence tools often take a year or more to reach the finish line...if they get there at all. This has always been a problem. Today, though, it's downright unacceptable. The world changes faster than ever. Speed has never been more important. By adhering to antiquated methods, firms lose the ability to see nascent trends—and act upon them until it's too late. But what if the process of turning raw data into meaningful insights didn't have to be so painful, time-consuming, and frustrating? What if there were a better way to do analytics? Fortunately, you're in luck... Analytics: The Agile Way is the eighth book from award-winning author and Arizona State University professor Phil Simon. Analytics: The Agile Way demonstrates how progressive organizations such as Google, Nextdoor, and others approach analytics in a fundamentally different way. They are applying the same Agile techniques that software developers have employed for years. They have replaced large batches in favor of smaller ones...and their results will astonish you. Through a series of case studies and examples, Analytics: The Agile Way demonstrates the benefits of this new analytics mind-set: superior access to information, quicker insights, and the ability to spot trends far ahead of your competitors.

agile for data science: Intelligence-Based Medicine Anthony C. Chang, 2020-06-27 Intelligence-Based Medicine: Data Science, Artificial Intelligence, and Human Cognition in Clinical Medicine and Healthcare provides a multidisciplinary and comprehensive survey of artificial intelligence concepts and methodologies with real life applications in healthcare and medicine. Authored by a senior physician-data scientist, the book presents an intellectual and academic interface between the medical and the data science domains that is symmetric and balanced. The content consists of basic concepts of artificial intelligence and its real-life applications in a myriad of medical areas as well as medical and surgical subspecialties. It brings section summaries to emphasize key concepts delineated in each section; mini-topics authored by world-renowned experts in the respective key areas for their personal perspective; and a compendium of practical resources, such as glossary, references, best articles, and top companies. The goal of the book is to inspire clinicians to embrace the artificial intelligence methodologies as well as to educate data scientists about the medical ecosystem, in order to create a transformational paradigm for healthcare and medicine by using this emerging new technology. - Covers a wide range of relevant topics from cloud computing, intelligent agents, to deep reinforcement learning and internet of everything - Presents the concepts of artificial intelligence and its applications in an easy-to-understand format accessible to clinicians and data scientists - Discusses how artificial intelligence can be utilized in a myriad of subspecialties and imagined of the future - Delineates the necessary elements for successful implementation of artificial intelligence in medicine and healthcare

agile for data science: Agile Processes in Software Engineering and Extreme Programming – Workshops Rashina Hoda, 2019-08-30 This open access book constitutes the research workshops, doctoral symposium and panel summaries presented at the 20th International Conference on Agile Software Development, XP 2019, held in Montreal, QC, Canada, in May 2019. XP is the premier agile software development conference combining research and practice. It is a hybrid forum where agile researchers, academics, practitioners, thought leaders, coaches, and trainers get together to present and discuss their most recent innovations, research results,

experiences, concerns, challenges, and trends. Following this history, for both researchers and seasoned practitioners XP 2019 provided an informal environment to network, share, and discover trends in Agile for the next 20 years. Research papers and talks submissions were invited for the three XP 2019 research workshops, namely, agile transformation, autonomous teams, and large scale agile. This book includes 15 related papers. In addition, a summary for each of the four panels at XP 2019 is included. The panels were on security and privacy; the impact of the agile manifesto on culture, education, and software practices; business agility – agile's next frontier; and Agile – the next 20 years.

agile for data science: Becoming an Agile Software Architect Rajesh R V, 2021-03-19 A guide to successfully operating in a lean-agile organization for solutions architects and enterprise architects Key FeaturesDevelop the right combination of processes and technical excellence to address architectural challengesExplore a range of architectural techniques to modernize legacy systemsDiscover how to design and continuously improve well-architected sustainable softwareBook Description Many organizations have embraced Agile methodologies to transform their ability to rapidly respond to constantly changing customer demands. However, in this melee, many enterprises often neglect to invest in architects by presuming architecture is not an intrinsic element of Agile software development. Since the role of an architect is not pre-defined in Agile, many organizations struggle to position architects, often resulting in friction with other roles or a failure to provide a clear learning path for architects to be productive. This book guides architects and organizations through new Agile ways of incrementally developing the architecture for delivering an uninterrupted, continuous flow of values that meets customer needs. You'll explore various aspects of Agile architecture and how it differs from traditional architecture. The book later covers Agile architects' responsibilities and how architects can add significant value by positioning themselves appropriately in the Agile flow of work. Through examples, you'll also learn concepts such as architectural decision backlog,the last responsible moment, value delivery, architecting for change, DevOps, and evolutionary collaboration. By the end of this Agile book, you'll be able to operate as an architect in Agile development initiatives and successfully architect reliable software systems. What you will learnAcquire clarity on the duties of architects in Agile developmentUnderstand architectural styles such as domain-driven design and microservicesIdentify the pitfalls of traditional architecture and learn how to develop solutionsUnderstand the principles of value and data-driven architectureDiscover DevOps and continuous delivery from an architect's perspectiveAdopt Lean-Agile documentation and governanceDevelop a set of personal and interpersonal qualitiesFind out how to lead the transformation to achieve organization-wide agilityWho this book is for This agile study guide is for architects currently working on agile development projects or aspiring to work on agile software delivery, irrespective of the methodology they are using. You will also find this book useful if you're a senior developer or a budding architect looking to understand an agile architect's role by embracing agile architecture strategies and a lean-agile mindset. To understand the concepts covered in this book easily, you need to have prior knowledge of basic agile development practices.

agile for data science: Agile Data Warehousing for the Enterprise Ralph Hughes, 2015-09-19 Building upon his earlier book that detailed agile data warehousing programming techniques for the Scrum master, Ralph's latest work illustrates the agile interpretations of the remaining software engineering disciplines: - Requirements management benefits from streamlined templates that not only define projects quickly, but ensure nothing essential is overlooked. - Data engineering receives two new hyper modeling techniques, yielding data warehouses that can be easily adapted when requirements change without having to invest in ruinously expensive data-conversion programs. - Quality assurance advances with not only a stereoscopic top-down and bottom-up planning method, but also the incorporation of the latest in automated test engines. Use this step-by-step guide to deepen your own application development skills through self-study, show your teammates the world's fastest and most reliable techniques for creating business intelligence systems, or ensure that the IT department working for you is building your next decision support system the right way. - Learn how to quickly define scope and architecture before programming

starts - Includes techniques of process and data engineering that enable iterative and incremental delivery - Demonstrates how to plan and execute quality assurance plans and includes a guide to continuous integration and automated regression testing - Presents program management strategies for coordinating multiple agile data mart projects so that over time an enterprise data warehouse emerges - Use the provided 120-day road map to establish a robust, agile data warehousing program

agile for data science: *Research Anthology on Agile Software, Software Development, and Testing* Management Association, Information Resources, 2021-11-26 Software development continues to be an ever-evolving field as organizations require new and innovative programs that can be implemented to make processes more efficient, productive, and cost-effective. Agile practices particularly have shown great benefits for improving the effectiveness of software development and its maintenance due to their ability to adapt to change. It is integral to remain up to date with the most emerging tactics and techniques involved in the development of new and innovative software. The Research Anthology on Agile Software, Software Development, and Testing is a comprehensive resource on the emerging trends of software development and testing. This text discusses the newest developments in agile software and its usage spanning multiple industries. Featuring a collection of insights from diverse authors, this research anthology offers international perspectives on agile software. Covering topics such as global software engineering, knowledge management, and product development, this comprehensive resource is valuable to software developers, software engineers, computer engineers, IT directors, students, managers, faculty, researchers, and academicians.

agile for data science: Clean Agile Robert C. Martin, 2019-09-12 Agile Values and Principles for a New Generation “In the journey to all things Agile, Uncle Bob has been there, done that, and has the both the t-shirt and the scars to show for it. This delightful book is part history, part personal stories, and all wisdom. If you want to understand what Agile is and how it came to be, this is the book for you.” –Grady Booch “Bob’s frustration colors every sentence of Clean Agile, but it’s a justified frustration. What is in the world of Agile development is nothing compared to what could be. This book is Bob’s perspective on what to focus on to get to that ‘what could be.’ And he’s been there, so it’s worth listening.” –Kent Beck “It’s good to read Uncle Bob’s take on Agile. Whether just beginning, or a seasoned Agilista, you would do well to read this book. I agree with almost all of it. It’s just some of the parts make me realize my own shortcomings, dammit. It made me double-check our code coverage (85.09%).” –Jon Kern Nearly twenty years after the Agile Manifesto was first presented, the legendary Robert C. Martin (“Uncle Bob”) reintroduces Agile values and principles for a new generation–programmers and nonprogrammers alike. Martin, author of Clean Code and other highly influential software development guides, was there at Agile’s founding. Now, in Clean Agile: Back to Basics, he strips away misunderstandings and distractions that over the years have made it harder to use Agile than was originally intended. Martin describes what Agile is in no uncertain terms: a small discipline that helps small teams manage small projects . . . with huge implications because every big project is comprised of many small projects. Drawing on his fifty years’ experience with projects of every conceivable type, he shows how Agile can help you bring true professionalism to software development. Get back to the basics–what Agile is, was, and should always be Understand the origins, and proper practice, of SCRUM Master essential business-facing Agile practices, from small releases and acceptance tests to whole-team communication Explore Agile team members’ relationships with each other, and with their product Rediscover indispensable Agile technical practices: TDD, refactoring, simple design, and pair programming Understand the central roles values and craftsmanship play in your Agile team’s success If you want Agile’s true benefits, there are no shortcuts: You need to do Agile right. Clean Agile: Back to Basics will show you how, whether you’re a developer, tester, manager, project manager, or customer. Register your book for convenient access to downloads, updates, and/or corrections as they become available. See inside book for details.

agile for data science: The Informed Company Dave Fowler, Matthew C. David, 2021-10-26 Learn how to manage a modern data stack and get the most out of data in your organization! Thanks

to the emergence of new technologies and the explosion of data in recent years, we need new practices for managing and getting value out of data. In the modern, data driven competitive landscape the best guess approach—reading blog posts here and there and patching together data practices without any real visibility—is no longer going to hack it. The Informed Company provides definitive direction on how best to leverage the modern data stack, including cloud computing, columnar storage, cloud ETL tools, and cloud BI tools. You'll learn how to work with Agile methods and set up processes that's right for your company to use your data as a key weapon for your success . . . You'll discover best practices for every stage, from querying production databases at a small startup all the way to setting up data marts for different business lines of an enterprise. In their work at Chartio, authors Fowler and David have learned that most businesspeople are almost completely self-taught when it comes to data. If they are using resources, those resources are outdated, so they're missing out on the latest cloud technologies and advances in data analytics. This book will firm up your understanding of data and bring you into the present with knowledge around what works and what doesn't. Discover the data stack strategies that are working for today's successful small, medium, and enterprise companies Learn the different Agile stages of data organization, and the right one for your team Learn how to maintain Data Lakes and Data Warehouses for effective, accessible data storage Gain the knowledge you need to architect Data Warehouses and Data Marts Understand your business's level of data sophistication and the steps you can take to get to level up your data The Informed Company is the definitive data book for anyone who wants to work faster and more nimbly, armed with actionable decision-making data.

agile for data science: *Driving Digital* Isaac Sacolick, 2017-08-24 Every organization makes plans for updating products, technologies, and business processes. But that’s not enough anymore for the twenty-first-century company. The race is now on for everyone to become a digital enterprise. For those individuals who have been charged with leading their company’s technology-driven change, the pressure is intense while the correct path forward unclear. Help has arrived! In Driving Digital, author Isaac Sacolick shares the lessons he’s learned over the years as he has successfully spearheaded multiple transformations and helped shape digital-business best practices. Readers no longer have to blindly trek through the mine field of their company’s digital transformation. In this thoroughly researched one-stop manual, learn how to: • Formulate a digital strategy • Transform business and IT practices • Align development and operations • Drive culture change • Bolster digital talent • Capture and track ROI • Develop innovative digital practices • Pilot emerging technologies • And more! Your company cannot avoid the digital disruption heading its way. The choice is yours: Will this mean the beginning of the end for your business, or will your digital practices be what catapults you into next-level success?

agile for data science: Big Data Strategies for Agile Business Bhuvan Unhelkar, 2017-09-13 Agile is a set of values, principles, techniques, and frameworks for the adaptable, incremental, and efficient delivery of work. Big Data is a rapidly growing field that encompasses crucial aspects of data such as its volume, velocity, variety, and veracity. This book outlines a strategic approach to Big Data that will render a business Agile. It discusses the important competencies required to streamline and focus on the analytics and presents a roadmap for implementing such analytics in business.

agile for data science: Business Intelligence and Agile Methodologies for Knowledge-Based Organizations: Cross-Disciplinary Applications Rahman El Sheikh, Asim Abdel, 2011-09-30 Business intelligence applications are of vital importance as they help organizations manage, develop, and communicate intangible assets such as information and knowledge. Organizations that have undertaken business intelligence initiatives have benefited from increases in revenue, as well as significant cost savings.Business Intelligence and Agile Methodologies for Knowledge-Based Organizations: Cross-Disciplinary Applications highlights the marriage between business intelligence and knowledge management through the use of agile methodologies. Through its fifteen chapters, this book offers perspectives on the integration between process modeling, agile methodologies, business intelligence, knowledge management, and

strategic management.

agile for data science: Data Analytics in Project Management Seweryn Spalek, J. Davidson Frame, Yanping Chen, Carl Pritchard, Alfonso Bucero, Werner Meyer, Ryan Legard, Michael Bragen, Klas Skogmar, Deanne Larson, Bert Brijs, 2019-01-01 Data Analytics in Project Management. Data analytics plays a crucial role in business analytics. Without a rigid approach to analyzing data, there is no way to glean insights from it. Business analytics ensures the expected value of change while that change is implemented by projects in the business environment. Due to the significant increase in the number of projects and the amount of data associated with them, it is crucial to understand the areas in which data analytics can be applied in project management. This book addresses data analytics in relation to key areas, approaches, and methods in project management. It examines: • Risk management • The role of the project management office (PMO) • Planning and resource management • Project portfolio management • Earned value method (EVM) • Big Data • Software support • Data mining • Decision-making • Agile project management Data analytics in project management is of increasing importance and extremely challenging. There is rapid multiplication of data volumes, and, at the same time, the structure of the data is more complex. Digging through exabytes and zettabytes of data is a technological challenge in and of itself. How project management creates value through data analytics is crucial. Data Analytics in Project Management addresses the most common issues of applying data analytics in project management. The book supports theory with numerous examples and case studies and is a resource for academics and practitioners alike. It is a thought-provoking examination of data analytics applications that is valuable for projects today and those in the future.

agile for data science: Big Data Governance and Perspectives in Knowledge Management Strydom, Sheryl Kruger, Strydom, Moses, 2018-11-16 The world is witnessing the growth of a global movement facilitated by technology and social media. Fueled by information, this movement contains enormous potential to create more accountable, efficient, responsive, and effective governments and businesses, as well as spurring economic growth. Big Data Governance and Perspectives in Knowledge Management is a collection of innovative research on the methods and applications of applying robust processes around data, and aligning organizations and skillsets around those processes. Highlighting a range of topics including data analytics, prediction analysis, and software development, this book is ideally designed for academicians, researchers, information science professionals, software developers, computer engineers, graduate-level computer science students, policymakers, and managers seeking current research on the convergence of big data and information governance as two major trends in information management.

agile for data science: Agile Management for Software Engineering David J. Anderson, 2003-09-17 A breakthrough approach to managing agile software development, Agile methods might just be the alternative to outsourcing. However, agile development must scale in scope and discipline to be acceptable in the boardrooms of the Fortune 1000. In Agile Management for Software Engineering, David J. Anderson shows managers how to apply management science to gain the full business benefits of agility through application of the focused approach taught by Eli Goldratt in his Theory of Constraints. Whether you're using XP, Scrum, FDD, or another agile approach, you'll learn how to develop management discipline for all phases of the engineering process, implement realistic financial and production metrics, and focus on building software that delivers maximum customer value and outstanding business results.Coverage includes: Making the business case for agile methods: practical tools and disciplines How to choose an agile method for your next project Breakthrough application of Critical Chain Project Management and constraint-driven control of the flow of value Defines the four new roles for the agile manager in software projects—and competitive IT organizations Whether you're a development manager, project manager, team leader, or senior IT executive, this book will help you achieve all four of your most urgent challenges: lower cost, faster delivery, improved quality, and focused alignment with the business.

agile for data science: *Agile Database Techniques* Scott Ambler, 2012-09-17 Describes Agile

Modeling Driven Design (AMDD) and Test-Driven Design (TDD) approaches, database refactoring, database encapsulation strategies, and tools that support evolutionary techniques Agile software developers often use object and relational database (RDB) technology together and as a result must overcome the impedance mismatch The author covers techniques for mapping objects to RDBs and for implementing concurrency control, referential integrity, shared business logic, security access control, reports, and XML An agile foundation describes fundamental skills that all agile software developers require, particularly Agile DBAs Includes object modeling, UML data modeling, data normalization, class normalization, and how to deal with legacy databases Scott W. Ambler is author of Agile Modeling (0471202827), a contributing editor with Software Development (www.sdmagazine.com), and a featured speaker at software conferences worldwide

agile for data science: Agile Systems Engineering Bruce Powel Douglass, 2015-09-24 Agile Systems Engineering presents a vision of systems engineering where precise specification of requirements, structure, and behavior meet larger concerns as such as safety, security, reliability, and performance in an agile engineering context. World-renown author and speaker Dr. Bruce Powel Douglass incorporates agile methods and model-based systems engineering (MBSE) to define the properties of entire systems while avoiding errors that can occur when using traditional textual specifications. Dr. Douglass covers the lifecycle of systems development, including requirements, analysis, design, and the handoff to specific engineering disciplines. Throughout, Dr. Douglass couples agile methods with SysML and MBSE to arm system engineers with the conceptual and methodological tools they need to avoid specification defects and improve system quality while simultaneously reducing the effort and cost of systems engineering. - Identifies how the concepts and techniques of agile methods can be effectively applied in systems engineering context - Shows how to perform model-based functional analysis and tie these analyses back to system requirements and stakeholder needs, and forward to system architecture and interface definition - Provides a means by which the quality and correctness of systems engineering data can be assured (before the entire system is built!) - Explains agile system architectural specification and allocation of functionality to system components - Details how to transition engineering specification data to downstream engineers with no loss of fidelity - Includes detailed examples from across industries taken through their stages, including the Waldo industrial exoskeleton as a complex system

agile for data science: *Agile Software Requirements* Dean Leffingwell, 2010-12-27 “We need better approaches to understanding and managing software requirements, and Dean provides them in this book. He draws ideas from three very useful intellectual pools: classical management practices, Agile methods, and lean product development. By combining the strengths of these three approaches, he has produced something that works better than any one in isolation.” –From the Foreword by Don Reinertsen, President of Reinertsen & Associates; author of Managing the Design Factory; and leading expert on rapid product development Effective requirements discovery and analysis is a critical best practice for serious application development. Until now, however, requirements and Agile methods have rarely coexisted peacefully. For many enterprises considering Agile approaches, the absence of effective and scalable Agile requirements processes has been a showstopper for Agile adoption. In Agile Software Requirements, Dean Leffingwell shows exactly how to create effective requirements in Agile environments. Part I presents the “big picture” of Agile requirements in the enterprise, and describes an overall process model for Agile requirements at the project team, program, and portfolio levels Part II describes a simple and lightweight, yet comprehensive model that Agile project teams can use to manage requirements Part III shows how to develop Agile requirements for complex systems that require the cooperation of multiple teams Part IV guides enterprises in developing Agile requirements for ever-larger “systems of systems,” application suites, and product portfolios This book will help you leverage the benefits of Agile without sacrificing the value of effective requirements discovery and analysis. You’ll find proven solutions you can apply right now–whether you’re a software developer or tester, executive, project/program manager, architect, or team leader.

agile for data science: Agile Artificial Intelligence in Pharo Alexandre Bergel, 2020-06-20

Cover classical algorithms commonly used as artificial intelligence techniques and program agile artificial intelligence applications using Pharo. This book takes a practical approach by presenting the implementation details to illustrate the numerous concepts it explains. Along the way, you'll learn neural net fundamentals to set you up for practical examples such as the traveling salesman problem and cover genetic algorithms including a fun zoomorphic creature example. Furthermore, Practical Agile AI with Pharo finishes with a data classification application and two game applications including a Pong-like game and a Flappy Bird-like game. This book is informative and fun, giving you source code to play along with. You'll be able to take this source code and apply it to your own projects. What You Will LearnUse neurons, neural networks, learning theory, and moreWork with genetic algorithms Incorporate neural network principles when working towards neuroevolution Include neural network fundamentals when building three Pharo-based applications Who This Book Is For Coders and data scientists who are experienced programmers and have at least some prior experience with AI or deep learning. They may be new to Pharo programming, but some prior experience with it would be helpful.

agile for data science: Research Anthology on Privatizing and Securing Data Management Association, Information Resources, 2021-04-23 With the immense amount of data that is now available online, security concerns have been an issue from the start, and have grown as new technologies are increasingly integrated in data collection, storage, and transmission. Online cyber threats, cyber terrorism, hacking, and other cybercrimes have begun to take advantage of this information that can be easily accessed if not properly handled. New privacy and security measures have been developed to address this cause for concern and have become an essential area of research within the past few years and into the foreseeable future. The ways in which data is secured and privatized should be discussed in terms of the technologies being used, the methods and models for security that have been developed, and the ways in which risks can be detected, analyzed, and mitigated. The Research Anthology on Privatizing and Securing Data reveals the latest tools and technologies for privatizing and securing data across different technologies and industries. It takes a deeper dive into both risk detection and mitigation, including an analysis of cybercrimes and cyber threats, along with a sharper focus on the technologies and methods being actively implemented and utilized to secure data online. Highlighted topics include information governance and privacy, cybersecurity, data protection, challenges in big data, security threats, and more. This book is essential for data analysts, cybersecurity professionals, data scientists, security analysts, IT specialists, practitioners, researchers, academicians, and students interested in the latest trends and technologies for privatizing and securing data.

agile for data science: *Encyclopedia of Data Science and Machine Learning* Wang, John, 2023-01-20 Big data and machine learning are driving the Fourth Industrial Revolution. With the age of big data upon us, we risk drowning in a flood of digital data. Big data has now become a critical part of both the business world and daily life, as the synthesis and synergy of machine learning and big data has enormous potential. Big data and machine learning are projected to not only maximize citizen wealth, but also promote societal health. As big data continues to evolve and the demand for professionals in the field increases, access to the most current information about the concepts, issues, trends, and technologies in this interdisciplinary area is needed. The Encyclopedia of Data Science and Machine Learning examines current, state-of-the-art research in the areas of data science, machine learning, data mining, and more. It provides an international forum for experts within these fields to advance the knowledge and practice in all facets of big data and machine learning, emphasizing emerging theories, principals, models, processes, and applications to inspire and circulate innovative findings into research, business, and communities. Covering topics such as benefit management, recommendation system analysis, and global software development, this expansive reference provides a dynamic resource for data scientists, data analysts, computer scientists, technical managers, corporate executives, students and educators of higher education, government officials, researchers, and academicians.

agile for data science: Agile Approaches for Successfully Managing and Executing

Projects in the Fourth Industrial Revolution Bolat, Hür Bersam, Temur, Gül Tekin, 2019-03-15 Communication between man and machine is vital to completing projects in the current day and age. Without this constant connectiveness as we enter an era of big data, project completion will result in utter failure. Agile Approaches for Successfully Managing and Executing Projects in the Fourth Industrial Revolution addresses changes wrought by Industry 4.0 and its effects on project management as well as adaptations and adjustments that will need to be made within project life cycles and project risk management. Highlighting such topics as agile planning, cloud projects, and organization structure, it is designed for project managers, executive management, students, and academicians.

agile for data science: Data Analytics with Hadoop Benjamin Bengfort, Jenny Kim, 2016-06 Ready to use statistical and machine-learning techniques across large data sets? This practical guide shows you why the Hadoop ecosystem is perfect for the job. Instead of deployment, operations, or software development usually associated with distributed computing, you'll focus on particular analyses you can build, the data warehousing techniques that Hadoop provides, and higher order data workflows this framework can produce. Data scientists and analysts will learn how to perform a wide range of techniques, from writing MapReduce and Spark applications with Python to using advanced modeling and data management with Spark MLlib, Hive, and HBase. You'll also learn about the analytical processes and data systems available to build and empower data products that can handle—and actually require—huge amounts of data. Understand core concepts behind Hadoop and cluster computing Use design patterns and parallel analytical algorithms to create distributed data analysis jobs Learn about data management, mining, and warehousing in a distributed context using Apache Hive and HBase Use Sqoop and Apache Flume to ingest data from relational databases Program complex Hadoop and Spark applications with Apache Pig and Spark DataFrames Perform machine learning techniques such as classification, clustering, and collaborative filtering with Spark's MLlib

agile for data science: Agile Project Delivery Aaron A. Blair, 2020-12-18 Agile Project Delivery reviews how different Agile methods can be applied to project delivery in complex corporate environments beyond the Agile Manifesto's original scope of software development. Taking readers through a typical project lifecycle, the text demonstrates how Agile techniques can be applied to each phase of a project using valuable tools and examples. Agile Project Delivery covers various approaches that are used across the many methodologies and frameworks that are part of the Agile family, including Scrum, XP, and Crystal, as well as some of Agile's influences, such as Lean and Kanban. Agile Project Delivery also provides readers with advanced instructions for using Atlassian's industry-leading Agile software, Jira. Bridging the gap between Agile methodology and application, this concise guide features practical delivery approaches, engaging case studies, useful templates to assist in Agile application, and chapter discussion questions to reinforce understanding on how to harness the benefits of Agile. With a focus on settings outside of software development and an accessible, pragmatic approach, Agile Project Delivery is an invaluable resource for students in any project management course, as well as for both aspiring and experienced project practitioners.

agile for data science: Choose Your WoW! Scott W. Ambler, Mark Lines, 2020 Hundreds of organizations around the world have already benefited from Disciplined Agile Delivery (DAD). Disciplined Agile (DA) is the only comprehensive tool kit available for guidance on building high-performance agile teams and optimizing your way of working (WoW). As a hybrid of all the leading agile and lean approaches, it provides hundreds of strategies to help you make better decisions within your agile teams, balancing self-organization with the realities and constraints of your unique enterprise context. The highlights of this handbook include: #1. As the official source of knowledge on DAD, it includes greatly improved and enhanced strategies with a revised set of goal diagrams based upon learnings from applying DAD in the field. #2 It is an essential handbook to help coaches and teams make better decisions in their daily work, providing a wealth of ideas for experimenting with agile and lean techniques while providing specific guidance and trade-offs for

those it depends questions. #3 It makes a perfect study guide for Disciplined Agile certification. Why fail fast (as our industry likes to recommend) when you can learn quickly on your journey to high performance? With this handbook, you can make better decisions based upon proven, context-based strategies, leading to earlier success and better outcomes--

agile for data science: Succeeding with Agile Mike Cohn, 2010 Proven, 100% Practical Guidance for Making Scrum and Agile Work in Any Organization This is the definitive, realistic, actionable guide to starting fast with Scrum and agile-and then succeeding over the long haul. Leading agile consultant and practitioner Mike Cohn presents detailed recommendations, powerful tips, and real-world case studies drawn from his unparalleled experience helping hundreds of software organizations make Scrum and agile work. Succeeding with Agile is for pragmatic software professionals who want real answers to the most difficult challenges they face in implementing Scrum. Cohn covers every facet of the transition: getting started, helping individuals transition to new roles, structuring teams, scaling up, working with a distributed team, and finally, implementing effective metrics and continuous improvement. Throughout, Cohn presents Things to Try Now sections based on his most successful advice. Complementary Objection sections reproduce typical conversations with those resisting change and offer practical guidance for addressing their concerns. Coverage includes Practical ways to get started immediately-and get good fast Overcoming individual resistance to the changes Scrum requires Staffing Scrum projects and building effective teams Establishing improvement communities of people who are passionate about driving change Choosing which agile technical practices to use or experiment with Leading self-organizing teams Making the most of Scrum sprints, planning, and quality techniques Scaling Scrum to distributed, multiteam projects Using Scrum on projects with complex sequential processes or challenging compliance and governance requirements Understanding Scrum's impact on HR, facilities, and project management Whether you've completed a few sprints or multiple agile projects and whatever your role-manager, developer, coach, ScrumMaster, product owner, analyst, team lead, or project lead-this book will help you succeed with your very next project. Then, it will help you go much further: It will help you transform your entire development organization.

agile for data science: Data Science at the Command Line Jeroen Janssens, 2014-09-25 This hands-on guide demonstrates how the flexibility of the command line can help you become a more efficient and productive data scientist. You'll learn how to combine small, yet powerful, command-line tools to quickly obtain, scrub, explore, and model your data. To get you started—whether you're on Windows, OS X, or Linux—author Jeroen Janssens introduces the Data Science Toolbox, an easy-to-install virtual environment packed with over 80 command-line tools. Discover why the command line is an agile, scalable, and extensible technology. Even if you're already comfortable processing data with, say, Python or R, you'll greatly improve your data science workflow by also leveraging the power of the command line. Obtain data from websites, APIs, databases, and spreadsheets Perform scrub operations on plain text, CSV, HTML/XML, and JSON Explore data, compute descriptive statistics, and create visualizations Manage your data science workflow using Drake Create reusable tools from one-liners and existing Python or R code Parallelize and distribute data-intensive pipelines using GNU Parallel Model data with dimensionality reduction, clustering, regression, and classification algorithms

agile for data science: An Introduction to Agile Data Engineering Using Data Vault 2. 0 Kent Graziano, 2015-11-22 The world of data warehousing is changing. Big Data & Agile are hot topics. But companies still need to collect, report, and analyze their data. Usually this requires some form of data warehousing or business intelligence system. So how do we do that in the modern IT landscape in a way that allows us to be agile and either deal directly or indirectly with unstructured and semi structured data?The Data Vault System of Business Intelligence provides a method and approach to modeling your enterprise data warehouse (EDW) that is agile, flexible, and scalable. This book will give you a short introduction to Agile Data Engineering for Data Warehousing and Data Vault 2.0. I will explain why you should be trying to become Agile, some of the history and rationale for Data Vault 2.0, and then show you the basics for how to build a data warehouse model using the Data

Vault 2.0 standards.In addition, I will cover some details about the Business Data Vault (what it is) and then how to build a virtual Information Mart off your Data Vault and Business Vault using the Data Vault 2.0 architecture.So if you want to start learning about Agile Data Engineering with Data Vault 2.0, this book is for you.

agile for data science: *Data Science on the Google Cloud Platform* Valliappa Lakshmanan, 2017-12-12 Learn how easy it is to apply sophisticated statistical and machine learning methods to real-world problems when you build on top of the Google Cloud Platform (GCP). This hands-on guide shows developers entering the data science field how to implement an end-to-end data pipeline, using statistical and machine learning methods and tools on GCP. Through the course of the book, you'll work through a sample business decision by employing a variety of data science approaches. Follow along by implementing these statistical and machine learning solutions in your own project on GCP, and discover how this platform provides a transformative and more collaborative way of doing data science. You'll learn how to: Automate and schedule data ingest, using an App Engine application Create and populate a dashboard in Google Data Studio Build a real-time analysis pipeline to carry out streaming analytics Conduct interactive data exploration with Google BigQuery Create a Bayesian model on a Cloud Dataproc cluster Build a logistic regression machine-learning model with Spark Compute time-aggregate features with a Cloud Dataflow pipeline Create a high-performing prediction model with TensorFlow Use your deployed model as a microservice you can access from both batch and real-time pipelines

agile for data science: Lean-Agile Software Development Alan Shalloway, Guy Beaver, James R. Trott, 2009-10-22 Agile techniques have demonstrated immense potential for developing more effective, higher-quality software. However,scaling these techniques to the enterprise presents many challenges. The solution is to integrate the principles and practices of Lean Software Development with Agile's ideology and methods. By doing so, software organizations leverage Lean's powerful capabilities for "optimizing the whole" and managing complex enterprise projects. A combined "Lean-Agile" approach can dramatically improve both developer productivity and the software's business value.In this book, three expert Lean software consultants draw from their unparalleled experience to gather all the insights, knowledge, and new skills you need to succeed with Lean-Agile development. Lean-Agile Software Development shows how to extend Scrum processes with an Enterprise view based on Lean principles. The authors present crucial technical insight into emergent design, and demonstrate how to apply it to make iterative development more effective. They also identify several common development "anti-patterns" that can work against your goals, and they offer actionable, proven alternatives. Lean-Agile Software Development shows how to Transition to Lean Software Development quickly and successfully Manage the initiation of product enhancements Help project managers work together to manage product portfolios more effectively Manage dependencies across the software development organization and with its partners and colleagues Integrate development and QA roles to improve quality and eliminate waste Determine best practices for different software development teams The book's companion Web site, www.netobjectives.com/lasd, provides updates, links to related materials, and support for discussions of the book's content.

agile for data science: Data Science at the Command Line Jeroen Janssens, 2021-08-17 This thoroughly revised guide demonstrates how the flexibility of the command line can help you become a more efficient and productive data scientist. You'll learn how to combine small yet powerful command-line tools to quickly obtain, scrub, explore, and model your data. To get you started, author Jeroen Janssens provides a Docker image packed with over 100 Unix power tools--useful whether you work with Windows, macOS, or Linux. You'll quickly discover why the command line is an agile, scalable, and extensible technology. Even if you're comfortable processing data with Python or R, you'll learn how to greatly improve your data science workflow by leveraging the command line's power. This book is ideal for data scientists, analysts, engineers, system administrators, and researchers. Obtain data from websites, APIs, databases, and spreadsheets Perform scrub operations on text, CSV, HTML, XML, and JSON files Explore data, compute descriptive statistics,

and create visualizations Manage your data science workflow Create your own tools from one-liners and existing Python or R code Parallelize and distribute data-intensive pipelines Model data with dimensionality reduction, regression, and classification algorithms Leverage the command line from Python, Jupyter, R, RStudio, and Apache Spark

Additional Resources

什么是 **Agile Software Development** (敏捷软件开发)？ - 知乎
Apr 16, 2014 · 既然题主问的是“Agile Methodology”,那么便应该比限定在“软件开发”领域要更加宽泛。本回答从“敏捷开发”出发，尝试解读究竟什么才是“敏捷”。一、从“敏捷开发”说起 “敏捷”概念 ...

什么是芯片领域的“敏捷设计（*Agile Development* - 知乎
什么是芯片领域的“敏捷设计（Agile Development）”？引用矽说公众号对DARPA资助项目的解说；也有提到RISCV，CHISEL等字眼。敏捷设计与超高效计算芯片，DARPA为未来半导体发 ...

请问路由器双频合一开了好还是不开好？ - 知乎
说实在的。。。这个问题要看具体场景，没什么确定性的答案。就我自己而言，一般都是开着的。除非是我自己这边设备很多，要做隔离优化网络的时候，否则不会手动去把双频分开来。双频 ...

什么是 **Agile Software Development** (敏捷软件开发)？ - 知乎
Apr 16, 2014 · 既然题主问的是“Agile Methodology”,那么便应该比限定在“软件开发”领域要更加宽泛。本回答从“敏捷开发”出发，尝试解读究竟什么才是“敏捷”。一、从“敏捷开发”说起 “敏捷” ...

什么是芯片领域的“敏捷设计（Agile Development - 知乎
什么是芯片领域的“敏捷设计（Agile Development）”？引用矽说公众号对DARPA资助项目的解说；也有提到RISCV，CHISEL等字眼。敏捷设计与超高效计算芯片，DARPA为未来半导体发 ...

请问路由器双频合一开了好还是不开好？ - 知乎
说实在的。。。这个问题要看具体场景，没什么确定性的答案。就我自己而言，一般都是开着的。除非是我自己这边设备很多，要做隔离优化网络的时候，否则不会手动去把双频分开来。双 ...

Agile For Data Science

Agile For Data Science

Related Articles

- agile practice guide pmi pdf
- agents of shield parents guide
- agriculture word search answer key

Back to Home