

1. Record Nr.	UNINA9910896524003321
Autore	Gupta Nikhil
Titolo	Databricks Data Intelligence Platform : Unlocking the GenAI Revolution // by Nikhil Gupta, Jason Yip
Pubbl/distr/stampa	Berkeley, CA : , : Apress : , : Imprint : Apress, , 2024
ISBN	979-88-6880-444-1
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (481 pages)
Altri autori (Persone)	YipJason
Disciplina	001.422 005.7
Soggetti	Quantitative research Database management Machine learning Electronic data processing - Management Sampling (Statistics) Data Analysis and Big Data Database Management Machine Learning IT Operations Methodology of Data Collection and Processing
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	1. Introduction -- 2. Lakehouse Platform -- 3. Databricks Platform Overview -- 4. Data Ingestion and Real-Time Analytics -- 5. Delta Lake: Deep Dive -- 6. Data Governance with Unity Catalog -- 7. Data Engineering and Analytics -- 8. Data Science, Machine Learning, and AI -- 9. Building GenAI Applications on Databricks Platform -- 10. Data Warehousing with DBSQL -- 11. Data Intelligence Platform -- 12. CI/CD and Application Development -- 13. Databricks Pricing and Observability using System Tables -- 14. Platform Security and Compliance -- 15. Advanced Topics and Industry Applications -- 16. Streaming Applications and HA/DR -- 17. Databricks in the Cloud Ecosystem -- 18. Conclusion.
Sommario/riassunto	This book is your comprehensive guide to building robust Generative AI

solutions using the Databricks Data Intelligence Platform. Databricks is the fastest-growing data platform offering unified analytics and AI capabilities within a single governance framework, enabling organizations to streamline their data processing workflows, from ingestion to visualization. Additionally, Databricks provides features to train a high-quality large language model (LLM), whether you are looking for Retrieval-Augmented Generation (RAG) or fine-tuning. Databricks offers a scalable and efficient solution for processing large volumes of both structured and unstructured data, facilitating advanced analytics, machine learning, and real-time processing. In today's GenAI world, Databricks plays a crucial role in empowering organizations to extract value from their data effectively, driving innovation and gaining a competitive edge in the digital age. This book will not only help you master the Data Intelligence Platform but also help power your enterprise to the next level with a bespoke LLM unique to your organization. Beginning with foundational principles, the book starts with a platform overview and explores features and best practices for ingestion, transformation, and storage with Delta Lake. Advanced topics include leveraging Databricks SQL for querying and visualizing large datasets, ensuring data governance and security with Unity Catalog, and deploying machine learning and LLMs using Databricks MLflow for GenAI. Through practical examples, insights, and best practices, this book equips solution architects and data engineers with the knowledge to design and implement scalable data solutions, making it an indispensable resource for modern enterprises. Whether you are new to Databricks and trying to learn a new platform, a seasoned practitioner building data pipelines, data science models, or GenAI applications, or even an executive who wants to communicate the value of Databricks to customers, this book is for you. With its extensive feature and best practice deep dives, it also serves as an excellent reference guide if you are preparing for Databricks certification exams.

What You Will Learn

- Foundational principles of Lakehouse architecture
- Key features including Unity Catalog, Databricks SQL (DBSQL), and Delta Live Tables
- Databricks Intelligence Platform and key functionalities
- Building and deploying GenAI Applications from data ingestion to model serving
- Databricks pricing, platform security, DBRX, and many more topics .
