

1. Record Nr.	UNINA9910254570903321
Autore	Cook Joshua
Titolo	Docker for Data Science : Building Scalable and Extensible Data Infrastructure Around the Jupyter Notebook Server / / by Joshua Cook
Pubbl/distr/stampa	Berkeley, CA : , : Apress : , : Imprint : Apress, , 2017
ISBN	9781484230121 1484230124
Edizione	[1st ed. 2017.]
Descrizione fisica	1 online resource (XXI, 257 p. 97 illus., 76 illus. in color.)
Disciplina	005.1
Soggetti	Big data Artificial intelligence Open source software Computer programming Python (Computer program language) Big Data Artificial Intelligence Open Source Python
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di bibliografia	Includes bibliographical references.
Nota di contenuto	Chapter 1: Introduction -- Chapter 2: Docker -- Chapter 3: Interactive Programming -- Chapter 4: Docker Engine -- Chapter 5: The Dockerfile -- Chapter 6: Docker Hub -- Chapter 7: The Opinionated Jupyter Stacks -- Chapter 8: The Data Stores -- Chapter 9: Docker Compose -- Chapter 10: Interactive Development.
Sommario/riassunto	Learn Docker "infrastructure as code" technology to define a system for performing standard but non-trivial data tasks on medium- to large-scale data sets, using Jupyter as the master controller. It is not uncommon for a real-world data set to fail to be easily managed. The set may not fit well into access memory or may require prohibitively long processing. These are significant challenges to skilled software engineers and they can render the standard Jupyter system unusable. As a solution to this problem, Docker for Data

Science proposes using Docker. You will learn how to use existing pre-compiled public images created by the major open-source technologies—Python, Jupyter, Postgres—as well as using the Dockerfile to extend these images to suit your specific purposes. The Docker-Compose technology is examined and you will learn how it can be used to build a linked system with Python churning data behind the scenes and Jupyter managing these background tasks. Best practices in using existing images are explored as well as developing your own images to deploy state-of-the-art machine learning and optimization algorithms. What You'll Learn: Master interactive development using the Jupyter platform Run and build Docker containers from scratch and from publicly available open-source images Write infrastructure as code using the docker-compose tool and its docker-compose.yml file type Deploy a multi-service data science application across a cloud-based system.

---