1. Record Nr. UNINA9910254835903321

Titolo	Big Data Factories : Collaborative Approaches / / edited by Sorin Adam Matei, Nicolas Jullien, Sean P. Goggins
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2017
ISBN	3-319-59186-X
Edizione	[1st ed. 2017.]
Descrizione fisica	1 online resource (VI, 141 p. 18 illus., 14 illus. in color.)
Collana	Computational Social Sciences, , 2509-9574
Disciplina	005.7
Soggetti	Data mining Big data Bioinformatics Application software Research—Moral and ethical aspects Data Mining and Knowledge Discovery Big Data/Analytics Computer Appl. in Social and Behavioral Sciences Research Ethics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references at the end of each chapters and index.
Nota di contenuto	Chapter1. Introduction -- Part 1: Theoretical Principles and Approaches to Data Factories -- Chapter2. Accessibility and Flexibility: Two Organizing Principles for Big Data Collaboration -- Chapter3. The Open Community Data Exchange: Advancing Data Sharing and Discovery in Open Online Community Science -- Part 2: Theoretical principles and ideas for designing and deploying data factory approaches -- Chapter4. Levels of Trace Data for Social and Behavioral Science Research -- Chapter5. The 10 Adoption Drivers of Open Source Software that Enables e-Research in Data Factories for Open Innovations -- Chapter6. Aligning online social collaboration data around social order: theoretical considerations and measures -- Part 3: Approaches in action through case studies of data based research, best practice scenarios, or educational briefs -- Chapter7. Lessons learned

from a decade of FLOSS data collection -- Chapter8. Teaching Students How (NOT) to Lie, Manipulate, and Mislead with Information Visualizations -- Chapter9. Democratizing Data Science: The Community Data Science Workshops and Classes.

Sommario/riassunto

The book proposes a systematic approach to big data collection, documentation and development of analytic procedures that foster collaboration on a large scale. This approach, designated as “data factoring” emphasizes the need to think of each individual dataset developed by an individual project as part of a broader data ecosystem, easily accessible and exploitable by parties not directly involved with data collection and documentation. Furthermore, data factoring uses and encourages pre-analytic operations that add value to big data sets, especially recombining and repurposing. The book proposes a research-development agenda that can undergird an ideal data factory approach. Several programmatic chapters discuss specialized issues involved in data factoring (documentation, meta-data specification, building flexible, yet comprehensive data ontologies, usability issues involved in collaborative tools, etc.). The book also presents case studies for data factoring and processing that can lead to building better scientific collaboration and data sharing strategies and tools. Finally, the book presents the teaching utility of data factoring and the ethical and privacy concerns related to it. Chapter 9 of this book is available open access under a CC BY 4.0 license at link.springer.com.