

1. Record Nr.	UNINA9910463201903321
Autore	Verborgh Ruben
Titolo	Using OpenRefine : the essential OpenRefine guide that takes you from data analysis and error fixing to linking your dataset to the Web // Ruben Verborgh, Max De Wilde ; cover image by Aniket Sawant
Pubbl/distr/stampa	Birmingham, England : , : Packt Publishing, , 2013 ©2013
ISBN	1-78328-909-0
Edizione	[1st edition]
Descrizione fisica	1 online resource (114 p.)
Collana	Community Experience Distilled
Disciplina	006.312
Soggetti	Data mining Electronic data processing Electronic books.
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Sommario/riassunto	<p>With this book on OpenRefine, managing and cleaning your large datasets suddenly got a lot easier! With a cookbook approach and free datasheets included, you'll quickly and painlessly improve your data managing capabilities. Create links between your dataset and others in an instant Effectively transform data with regular expressions and the General Refine Expression Language Spot issues in your dataset and take effective action with just a few clicks In Detail Data is supposed to be the new gold, but how can you unlock the value in your data? Managing large datasets used to be a task for specialists, but you don't have to worry about inconsistencies or errors anymore. OpenRefine lets you clean, link, and publish your dataset in a breeze. Using OpenRefine takes you on a practical tour of all the handy features of this well-known data transformation tool. It is a hands-on recipe book that teaches you data techniques by example. Starting from the basics, it gradually transforms you into an OpenRefine expert. This book will teach you all the necessary skills to handle any large dataset and to turn it into high-quality data for the Web. After you learn how to analyze data and spot issues, we'll see how we can solve them to obtain</p>

a clean dataset. Messy and inconsistent data is recovered through advanced techniques such as automated clustering. We'll then show extract links from keyword and full-text fields using reconciliation and named-entity extraction. Using OpenRefine is more than a manual: it's a guide stuffed with tips and tricks to get the best out of your data.
