1.	Record Nr.	UNINA9910826746603321
	Autore	Oakes Michael P.
	Titolo	Literary detective work on the computer / / Michael P. Oakes
	Pubbl/distr/stampa	Amsterdam (Netherlands) ; ; Philadelphia, Pennsylvania : , : John Benjamins Publishing Company, , 2014 ©2014
	ISBN	90-272-7013-9
	Descrizione fisica	1 online resource (293 p.)
	Collana	Natural Language Processing ; ; Volume 12
	Disciplina	410.285
	Soggetti	Computational linguistics - Research Imitation in literature Plagiarism
	Lingua di pubblicazione	Inglese
	Formato	Materiale a stampa
	Livello bibliografico	Monografia
	Note generali	Description based upon print version of record.
	Nota di bibliografia	Includes bibliographical references and index.
	Nota di contenuto	Literary Detective Work on the Computer; Editorial page; Title page; LCC data; Table of contents; Preface; 1. Author identification; 1. Introduction; 2. Feature selection; 2.1 Evaluation of feature sets for authorship attribution; 3. Inter-textual distances; 3.1 Manhattan distance and Euclidean distance; 3.2 Labbe and Labbe's measure; 3.3 Chi-squared distance; 3.4 The cosine similarity measure; 3.6 Burrows' Delta; 3.5 Kullback-Leibler Divergence (KLD); 3.7 Evaluation of feature-based measures for inter-textual distance; 3.8 Inter-textual distance by semantic similarity 3.9 Stemmatology as a measure of inter-textual distance4. Clustering techniques; 4.1 Introduction to factor analysis; 4.2 Matrix algebra; 4.3 Use of matrix algebra for PCA; 4.4 PCA case studies; 4.5 Correspondence analysis; 5. Comparisons of classifiers; 6. Other tasks related to authorship; 6.1 Stylochronometry; 6.2 Affect dictionaries and psychological profiling; 6.3 Evaluation of author profiling; 7. Conclusion; 2. Plagiarism and spam filtering; 1. Introduction; 2. Plagiarism detection software; 2.1 Collusion and plagiarism, external and intrinsic 2.2 Preprocessing of corpora and feature extraction2.3 Sequence comparison and exact match; 2.4 Source-suspicious document similarity measures; 2.5 Fingerprinting; 2.6 Language models; 2.7

Sommario/riassunto	Computational linguistics can be used to uncover mysteries in text which are not always obvious to visual inspection. For example, the computer analysis of writing style can show who might be the true author of a text in cases of disputed authorship or suspected plagiarism. The theoretical background to authorship attribution is presented in a step by step manner, and comprehensive reviews of the field are given in two specialist areas, the writings of William Shakespeare and his contemporaries, and the various writing styles seen in religious texts. The final chapter looks at the progress com