

1. Record Nr.	UNINA9910350287003321
Autore	Dash Niladri Sekhar
Titolo	Utility and Application of Language Corpora // by Niladri Sekhar Dash, L. Ramamoorthy
Pubbl/distr/stampa	Singapore : , : Springer Singapore : , : Imprint : Springer, , 2019
ISBN	981-13-1801-8
Edizione	[1st ed. 2019.]
Descrizione fisica	1 online resource (XXX, 290 p. 40 illus., 1 illus. in color.)
Disciplina	410
Soggetti	Corpora (Linguistics) Natural language processing (Computer science) Learning Instruction Corpus Linguistics Natural Language Processing (NLP) Learning & Instruction
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	1. Introduction -- 2. Digital Speech Corpora -- 3. Language Corpora: Indian Scenario -- 4. Issues in Corpus Generation -- 5. Process of Corpus -- 6. Corpus Sanitation and Pre-Editing -- 7. Statistical Studies on Corpus -- 8. Corpus Text Processing -- 9. Corpus as Primary Resource for ELT -- 10. Corpus as Secondary resource for ELT -- 11. Corpus and Lexicography -- 12. Corpus and Dialect Study Chapter -- 13. Corpus and Word Sense Disambiguation -- 14. Corpus and Language Technology -- 15. Corpus and Other Branches of Linguistics -- 16. Corpora: Future Indian Needs.
Sommario/riassunto	This book discusses some of the basic issues relating to corpus generation and the methods normally used to generate a corpus. Since corpus-related research goes beyond corpus generation, the book also addresses other major topics connected with the use and application of language corpora, namely, corpus readiness in the context of corpus sanitation and pre-editing of corpus texts; the application of statistical methods; and various text processing techniques. Importantly, it explores how corpora can be used as a primary or secondary resource

in English language teaching, in creating dictionaries, in word sense disambiguation, in various language technologies, and in other branches of linguistics. Lastly, the book sheds light on the status quo of corpus generation in Indian languages and identifies current and future needs. Discussing various technical issues in the field in a lucid manner, providing extensive new diagrams and charts for easy comprehension, and using simplified English, the book is an ideal resource for non-native English readers. Written by academics with many years of experience teaching and researching corpus linguistics, its focus on Indian languages and on English corpora makes it applicable to graduate and postgraduate students of applied linguistics, computational linguistics and language processing in South Asia and across countries where English is spoken as a first or second language.

---