

1. Record Nr.	UNISA996691917003316
Autore	Baski Piotr
Titolo	Harmonizing language data : Standards for linguistic resources. // Piotr Baski
Pubbl/distr/stampa	LaVergne, : De Gruyter, 2025
ISBN	3-11-220953-2 3-11-220821-8
Descrizione fisica	1 online resource
Collana	Digital Linguistics, ; 4.
Classificazione	COM014000FOR000000LAN009000
Soggetti	Nonfiction Computer Technology Grammar & Language Usage Language Arts
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Title from eBook information screen..
Sommario/riassunto	Standards function as safeguards to ensure that data remains interpretable, uniformly queryable, and archivable over time – a critical challenge for digital humanists working with complex linguistic resources. This book provides an overview of essential standards for ensuring the sustainability of data in the Digital Humanities (DH). It addresses the selection of data encoding formats, methods of annotating primary data, and approaches to making resources findable and accessible. The focus is on various forms of linguistic data, such as texts, lexicons, or parallel arrangements (e.g., translations or transcribed recordings). The work explains the role of annotations and metadata in structuring and contextualizing data and examines the influence of diverse data formats, shaped by local academic or industrial practices. In contrast to neural language models, which often yield impressive but opaque results, DH projects aim for transparency, reproducibility, and sustainability. Achieving these goals requires interoperability – the seamless interaction between data and tools. The book demonstrates how clear guidelines and best practices help ensure the long-term usability of data. It offers digital humanists practical

approaches and well-founded standards to sustainably archive and efficiently utilize their data, making it an indispensable resource for the field.
