

Modulo 2.2.3 – Segmentare il testo

Risorse e link

- Steven Bird, Ewan Klein e Edward Loper, *Natural Language Processing with Python*
<<https://www.nltk.org/book/>>
- Corpus Europarl
<<http://www.statmt.org/europarl/>>
- Robert Östling. *Bayesian Models for Multilingual Word Alignment*, 2016, Ph.D. thesis
<<https://www.diva-portal.org/smash/get/diva2:798117/FULLTEXT01.pdf>>
- Università degli Studi di Palermo, Tokenizzazione (slide)
<http://www1.unipa.it/sorce/didattica/sei1213/SEI1213_05_tokenizzazione.pdf>
- Voce “Linguistica computazionale”, in *Wikipedia*
<https://it.wikipedia.org/wiki/Linguistica_computazionale>
- Voce “Text segmentation”, in *Wikipedia*
<https://en.wikipedia.org/wiki/Text_segmentation>
- Voce “Traduzione assistita dal computer”, in *Wikipedia*
<https://it.wikipedia.org/wiki/Traduzione_assistita_dal_computer>
- Voce “Corpus Europarl” in *Wikipedia*
<https://en.wikipedia.org/wiki/Europarl_Corpus>
- Word Tokenization with Python NLTK
<<https://text-processing.com/demo/tokenize/>>