Linguistic resources for NLP – Part 1 (C. Bosco)

September, 23 2024

Exercises on Tokenization

В

How many tokens you see in the following sentences? Count the tokens and write the number.

- Early in the morning, in San Francisco, several men are working. = 13
- To go to Mary's cottage you must turn left. = 11
- Dr. Verdi didn't like very cheap restaurants (like McDonald's). = 14
- That New York-based family often travelled to Italy. = 8
- Our meeting will be at $\frac{2pm}{pm}$ on $\frac{09}{25} = 9$
- The most intense moment was when he sang "Shape of you". =14