

# What is a word?

**It seems natural to think of a text as a sequence of words**

- A word is a meaningful sequence of characters

**How to find the boundaries of words?**

- In English we can split a sentence by spaces or punctuation

---

**Input:** Friends, Romans, Countrymen, lend me your ears;

**Output:** Friends Romans Countrymen lend me your ears

---

- In German there are compound words which are written without spaces
  - “Rechtsschutzversicherungsgesellschaften” stands for “insurance companies which provide legal protection”
- In Japanese there are no spaces at all!
  - But you can still read it right?