

TEAM NLTK

PROPERTY DATA

COLLECTED DATA

- ▶ Properties: 14 000+ (ImmoWeb)
- ▶ Whole Belgium
- ▶ 20 different property/sup-types
- ▶ No duplicates, CSV -format
- ▶ TBD: numerical value conversion

HOW IT'S DONE

- ▶ Selenium
 - ▶ Ad url
- ▶ BeautifulSoup
 - ▶ Collects attributes
- ▶ Analysis tool
 - ▶ Numeric values (one hot etc)
 - ▶ CSV file creation

TEXT

ROLES

- ▶ Jari: development
- ▶ Yass: revision
- ▶ Nemish: development
- ▶ Giovanna: revision

WHAT WENT WRONG?

- ▶ Address information
- ▶ Price-information issues