

- **web_scrapy folder**

Data crawling code from Chefkoch.

- **functions.py**

All the helper functions used in the following Jupyter notebook.

- **aggregate_data.ipynb**

We aggregate all the dataset we crawled from different categories and remove the duplicate recipes according to the unique recipe id

- **Section_3&4.ipynb**

The coding of section 3 and section 4.

- **Section_5.ipynb**

The coding of section 5 'what ingredients'. We trained word2vec and plotted figure 7.

- **Section_6.ipynb**

The coding of section 6 'Who is eating what'. Besides, the coding for figure 2 in section 3 is also included.