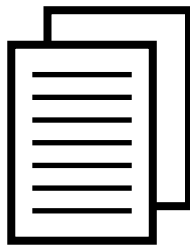


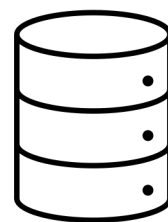
4,173
repositories

Downloading and
collecting data



Attributes of
770,656 issues

Selecting, converting,
and classifying data



Our dataset

Download issues of
selected repositories
from GitHub

Downloaded
issues

Get time
information

Count the number of
words and characters
in the text

Count the number of
movies and images
in the text

Focus target issues

Convert words to
tf-idf values

Classify into categories
by presence or absence
of movies and images