

Для ефективного функціонування програми була потрібна початкова база повідомлень. Вони були отримані наступним чином. Із Slack каналу programming_2016_2017 за допомогою Python-обгортки SlackClient для Slack API була стягнута історія повідомлень цього каналу у форматі json. (Було отримано не всі повідомлення каналу через обмеження їх кількості Slackом до 10 000 на Slack команду.) Цей json-файл був автоматично перетворений у тип даних словник за допомогою обгортки SlackClient. Цей словник містив два ключі. За ключем "messages" було отримано доступ до списку словників, де кожен словник містив ключі, які характеризували повідомлення. Характеристики: id автора повідомлення, час повідомлення, текст повідомлення, реакції на повідомлення, канал в якому було відправлено повідомлення. Отримані повідомлення були записані у текстовий файл кожне повідомлення в окремому рядку, кожне у форматі словника. Загалом обсяг усіх повідомлень займає 24kB дискового простору.

Також із GitHub репозиторію було отримано файл із stop-words(прийменники, займенники, сполучники, частки, прислівники) української мови. Загальний розмір 4.1kB. Ці дані необхідні для фільтрування отриманих повідомлень.

Було отримано ID Slack каналу programming_2016_2017 за його назвою, ID викладачів для подальшої інтеграції бота в Slack канал, ID бота для відокремлення його повідомлень від повідомлень інших користувачів.