23.01.2023

**Описание чат-бота Telegram @dbSNP\_bot**

1. Описание концепции.

В настоящее время большую популярность набирают генетические тестирования, которые позволяют определить предрасположенность к ряду заболеваний, некоторым видам спорта, непереносимости ряда пищевых продуктов и пр. При проведении генетических тестов определяются генетические варианты, нуклеотидная последовательность которых отличается от референсного генома, так называемые альтернативные аллели. Но не каждый отличающийся от референса генетический вариант несет в себе клиническую значимость. Так, например, по рекомендациям ACMG (American College of Medical Genetics) генетический вариант не может считаться патогенным, если частота встречаемости его в популяции выше 5%. Но, к сожалению, в генетических паспортах некоторые компании декларируют такие варианты как риск-фактор или даже патогенные, что вводит в заблуждение пациентов. Чат-бот Telegram @dbSNP\_bot выдает информацию о генетическом варианте (по запросу rs) с частотой его встречаемости в популяциях с портала (<https://www.ncbi.nlm.nih.gov/snp/>) и благодаря этому генетики, клинические биоинформатики и просто любой человек, обладающий результатами генетического тестирования, смогут получить эту информацию со своего смартфона в режиме real-time.

1. Описание реализации

В ходе написания скрипта использовались библиотеки pytelegrambotapi, bs4, requests. Скрипт парсит 3 основные страницы: <https://www.ncbi.nlm.nih.gov/projects/SNP/get_html.cgi?whichHtml=overview>

<https://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi>

[https://www.ncbi.nlm.nih.gov/snp/{rs}#frequency\_tab](https://www.ncbi.nlm.nih.gov/snp/%7brs%7d#frequency_tab), где rs – сообщение от пользователя

Проверка существования последней ссылки, а соответственно корректности введенной rs, происходит по коду requests.get(url).status\_code (равно или не равно 200).

К сожалению, не получилось задеплоить код на Heroku, GoogleCloud (из-за отсутствия зарубежной карты), поэтому теперь код крутится на pythonanywhere by ANACONDA, но он иногда требует перезагрузки бота.

1. Идеи для масштабирования проекта.

В качестве развития проекта и расширения его функционала можно добавить ссылки и переход на них на соответствующие страницы ClinVar, OMIM (для упрощения интерпретации клинической значимости генетического варианта), изображения из геномного браузера, иллюстрирующие позицию этого варианта в геноме.