

**L0 Study Project Report**

Project: **Telegram Bot and Data Processing**

Project by:

* **Fatulla Bashirov**
* **Toghrul Rzayev**
* **Elgun Hasanov**
* **Sanan Guluzade**

Supervisor:

**Professor Safi Khaled**

Date:

May 2022

**Introduction:**

Studying how to parse raw data and present the processed data. For presenting our processed data, we will learn how to use Telegram bot API programmed in python. Our goal is to achieve this in the most efficient way possible.

**Methodology:**

Manipulation of timetables is processed through **XML** files at UFAZ. The Extensible Markup Language (XML) is a simple text-based format for representing structured information. It is easy to exchange data in this format between systems and applications. However, **.xml** was not created to display data.

To display the timetable students and teachers at UFAZ use the EduPage platform. Our idea was to use Telegram Bot to access the timetable most comfortably.

Our first problem was turning this .xml file format into another format, which our program can easily use. Elgun studied several python modules for this but then we scrapped this idea.

We decided to directly gather information from .xml file and store them in python classes with our python program. The work with the XML file was performed by Toghrul.

**Tools used to process the data**:

* xml.etree.ElementTree module implements API for parsing XML data (Python)
* Classes for sorting and using the data in Python

After parsing the timetable from XML, we had all lessons for the semester. Getting current, next lessons requires the exact time at that moment to give the right information. This was done by the **datetime** module. Studying and implementing this module was the part of Sanan. Sanan wrote some functions to determine the lesson period according to the time.

The group of student is important for the program as this information is used to find the appropriate classes that have the timetable information. Information about the students is stored in a JSON file, which was done by Elgun. To recognize the student we used the telegram IDs of the user.

After we processed the data from the timetable file, created a database and wrote a program to determine the time at that moment, we started to work with displaying the data. The commands of our bot were written by Fatulla

* **/start –** *to start the bot. If the ID of the student is not found in database it is asked to set a group*
* **/setgroup –** *to change the group of a user*
* **/current –** *to get the classroom, lesson subject and teacher of the current lesson*
* **/next –** *to get the information about the next lesson*
* **/tomorrow –** *to get the lessons for tomorrow*
* **/today –** *to get today’s lessons*
* **/nextweek –** *to obtain all the lessons for next week*
* **/subscribe –** *to get messages about next lessons*

The idea of the last command is to get notifications at the end of each lesson about the next lesson so that the student will know the classroom where the lesson will be.

\*Most of the work was distributed between the members of group in March and completed during the Nowruz holidays.

**Results and discussions**

In our end program, we were able to successfully parse XML data and use this data in the python program. This required us to learn and implement several modules. All these steps were written in several python programs. These programs were imported and used between each other. As future perspectives of our project, we think to automate the process of updating XML file, add absence recorder and exam grades to the database. The idea is to get the absence records and GPA anytime the student wants. Toghrul and Fatulla already created the shared mailbox (*auto-grades@ufaz.az*) to receive Excel files with grades of all groups and studied the **openpyxl** module to parse the grades.

**Conclusion**

In the era of digitalization, everything depends on data and it is very important to process the raw data and display it in a convenient form.

Databases are the main part of projects and companies. We tried to use Excel file before but it was hard to update the sheets and extract the data from such file. Toghrul found a better solution – JSON, which is easy to use in Python.

**References**

* To study the documentation of TelegramBotAPI

<https://www.youtube.com/watch?v=NwBWW8cNCP4&t=245s>

<https://www.youtube.com/watch?v=xFoUNDRVBYM&t=355s>

<https://pypi.org/project/pyTelegramBotAPI/>

* To learn how to use the datetime module

<https://www.programiz.com/python-programming/datetime>

* For parsing the data using xml.etree.ElementTree

<https://www.youtube.com/watch?v=j0xr0-IAqyk>

<https://www.youtube.com/watch?v=r6dyk68gymk>

<https://www.youtube.com/watch?v=1JblVElt5K0>

* Working with JSON file

<https://www.youtube.com/watch?v=9N6a-VLBa2I>

<https://www.youtube.com/watch?v=-51jxlQaxyA>