МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ «МИФИ»

ИНСТИТУТ ФИНАНСОВОЙ И ЭКОНОМИЧЕСКОЙ БЕЗОПАСНОСТИ **КАФЕДРА ФИНАНСОВОГО МОНИТОРИНГА**

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №3 НА ТЕМУ: МЕТОДЫ ИЗВЛЕЧЕНИЯ БОЛЬШИХ ДАННЫХ ПО КУРСУ «РИАС»

Выполнил: Нуритдинхо	оджаева А. (С20-702).
Преподаватель: Прохор	оов И.В.
Оценка	Дата

Цель: Освоить методы работы с большими данными, используя сервисы API ВКонтакте и API Яндекс. Диск.

Задание1 Технология извлечения данных из социальной сети ВКонтакте Функциональная модель:



Рисунок 1 Функциональная модель технологии извлечения данных из социальной сети

Для выполнения лабораторной работы создали Standalone приложение перейдя по ссылке: https://vk.com/dev, после чего выполнен следующий запрос в поисковой строке: https://oauth.vk.com/authorize?client_id=51920987&scope=photos,audio,video,docs,notes,pages, <a href="https://oauthorize.new.notifications.stats.notifications.notification

После перехода по ссылке ВКонтакте просит подтвердить доступ к аккаунту, заием получили следующую запись в адресной строке:

https://oauth.vk.com/blank.html#access_token=vk1.a.x3wkE8A5yrPoMDRoXtpA7gcCv Hh-

ybVaPk_YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHSkNjXtJLpjFXAALdB3dyL7ezk JHm2YmsfIoL0Q6J08Mw-__McGCT5XLSIB5pXnhICkgivUf_XT-

dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-

<u>fPL1CA&expires_in=0&user_id=353421386</u>, где красным выделен access.token, а голубым – срок его действия в секундах (ноль означает бессрочность).

После получения access.token можно приступить в выполнению запросов API.

1. Определение участников, которые на момент исследования были онлайн: api.vk.com/method/users.search?online=1&group_id=30444828&access_token=vk1.a.x3wkE8A 5yrPoMDRoXtpA7gcCvHh-

ybVaPk_YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHSkNjXtJLpjFXAALdB3dyL7ezk JHm2YmsfIoL0Q6J08Mw-__McGCT5XLSIB5pXnhICkgivUf_XT-

dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-

fPL1CA&v=5.199&p1=v1

Результат в формате JSON:

Рисунок 2 Пользователи онлайн на момент выполнения запроса

2. Все идентификаторы пользователей в сообществе:

https://api.vk.com/method/groups.getMembers?group_id=iroothelp&v=5.131&access_token=vk 1.a.x3wkE8A5yrPoMDRoXtpA7gcCvHh-

ybVaPk YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHSkNjXtJLpjFXAALdB3dyL7ezk JHm2YmsfIoL0Q6J08Mw- McGCT5XLSIB5pXnhICkgivUf XT-

dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-fPL1CA

Результат в формате JSON:

("response": ("count": 319, "items":
[395596, 808040, 1526555, 1600185, 1642754, 165821, 1777410, 1853723, 1871159, 2014798, 2365939, 2512092, 2871142, 2985339, 3677041, 4076937, 4539553, 46098644, 5387044, 5625295, 5689851, 6010551, 7673899, 7844890, 8186180, 8351090, 89090775, 8961884, 9143929, 9225116, 9263556, 2209042, 10136676, 11108557, 11122066, 11144099, 11640964, 11872687, 1191366, 12281891, 12390580, 12366170, 13747740, 13846877, 14533409, 1454819, 15569094, 116347094, 17162777, 17267366, 180606202, 2022235, 236739293, 2246489092, 2712162232, 74519099, 27922323, 235119999, 2745084, 2391927, 13945894, 145046427, 416656456, 43956444, 49956444, 49671366, 515991385, 5355524, 58610210, 53409393, 59975761, 60335159, 622304593, 66109433, 60760932, 70071070, 78801992, 70177086, 81064312, 48528934, 860712188, 86563396, 86680996, 86775238, 87595424, 87821814, 90380425, 9051739, 9315173, 931

Рисунок 3 Все идентификаторы пользователей в сообществе

3. Вывод последних пяти записей сообщества:

https://api.vk.com/method/wall.get?filter=all&domain=iroothelp&v=5.131&access_token=vk1.a .x3wkE8A5yrPoMDRoXtpA7gcCvHh-

ybVaPk_YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHSkNjXtJLpjFXAALdB3dyL7ezk JHm2YmsfIoL0Q6J08Mw-__McGCT5XLSIB5pXnhICkgivUf_XT-

dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-fPL1CA

Результат в формате JSON:

32.us=rapi.com//impft/fpcK/7ReQ_x1QQgInO31aTW-13Y_VXXE6hA6Q(ym63739Re8o.jpg?
size=1280x604&quality=96&sign=fcc50647die55bd65a389ed522da95de&c_uniq_tag=12RVmY0skrps0R0j6lcu65q_LaeGesb82hE9vJiwcqU&type=album*}], "text":"", "user_id":100, "web_view_token":"3d62ee230687f94834", "has_tags":false}}
("tvee":"onl":"", "onl":"

Рисунок 4 Последние записи сообщества

4. Все друзья в сообществе, которые онлайн:

URL:

https://api.vk.com/method/groups.getMembers?group_id=iroothelp&v=5.131&fields=online&fil ter=friends&access_token=vk1.a.x3wkE8A5yrPoMDRoXtpA7gcCvHh-

ybVaPk_YgNQx9VaJF5sA1cWtPqDg3AtTaAwtuDRas1kvizHSkNjXtJLpjFXAALdB3dyL7ezk JHm2YmsfIoL0Q6J08Mw-__McGCT5XLSIB5pXnhICkgivUf_XT-

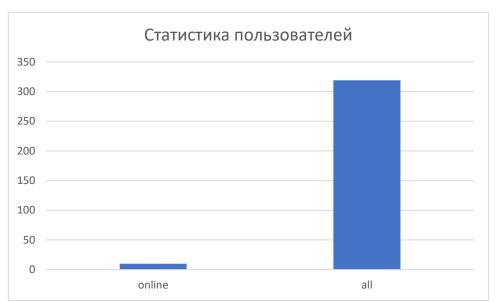
dGzbPwRSdpADvVmbiN2cox6p0TpGsXh3QMbcgY1HWPhncJC3Yv1g1hYU3-fPL1CA

Результат в формате JSON:

$$\{"response": \{"count":0,"items":[],"next_from":""\}\}$$

Рисунок 5 Друзья в сообществе, которые онлайн

Построим гистограмму полученных данных API ВКонтакте, с использованием excel:



Задание 2
Работа с облачным сервисом Яндекс. Диск. Технология работы с данными, хранящимися в облачном хранилище из приложения.

Функциональная модель:



Рисунок 6 Функциональная модель технологии извлечения данных из Яндекс.Диска

Для выполнения лабораторной работы создали приложение перейдя по ссылке: https://oauth.yandex.ru/, после чего выполнен следующий запрос в поисковой строке:

 $\frac{https://oauth.yandex.ru/authorize?response\ type=token\&client\ id=b9088f68f6344106b0434074e48014c}{4}$

После перехода по ссылке ВКонтакте просит подтвердить доступ к аккаунту, затем получили следующую запись в адресной строке:

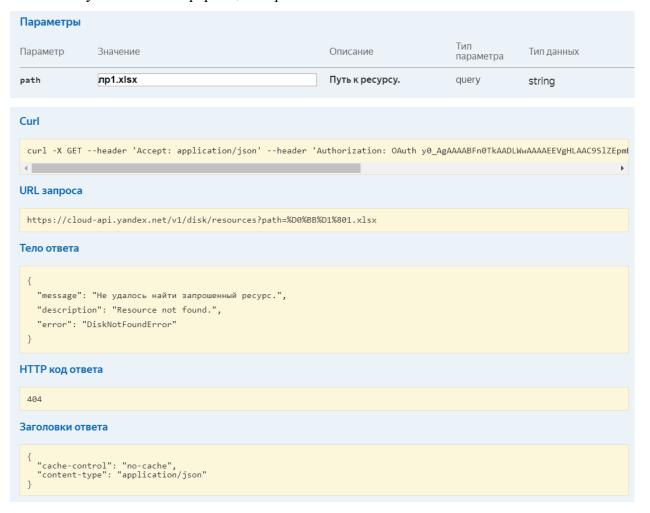
https://rias.yandex.ru/#access_token=y0_AgAAAABFn0TkAAvBbgAAAAEEVbvmAA
AwnorxD0hD56IYh0IbbUD9Yr51nQ&token_type=bearer&expires_in=31535603&cid=4q0cyx
n7zcza6r7f2fzhtmv4vm
, где красным выделен access.token.

Тестировать запросы можно на так называемом полигоне Яндекс. Диска, который находится по ссылке: https://yandex.ru/dev/disk/poligon/

1. Получение метаинформацию о диске

```
Curl
  curl -X GET --header 'Accept: application/json' --header 'Authorization: OAuth y0_AgAAAABFn0TkAADLWwAAAAEEVgHLAAC9S1ZEpmf
 URL запроса
  https://cloud-api.yandex.net/v1/disk
 Тело ответа
    "max_file_size": 1073741824,
    "paid_max_file_size": 53687091200,
     "total_space": 32212254720,
     "reg_time": "2020-09-13T04:07:12+00:00",
     "trash_size": 0,
     "is_paid": false,
     "used_space": 6350572107,
     "system_folders": {
      "odnoklassniki": "disk:/Социальные сети/Одноклассники",
      "google": "disk:/Социальные сети/Google+",
      "instagram": "disk:/Социальные сети/Instagram",
      "vkontakte": "disk:/Социальные сети/ВКонтакте",
      "attach": "disk:/Почтовые вложения",
      "mailru": "disk:/Социальные сети/Мой Мир",
      "downloads": "disk:/Загрузки/",
      "applications": "disk:/Приложения",
       "facebook": "disk:/Социальные сети/Facebook",
       "social": "disk:/Социальные сети/",
      "messenger": "disk:/Файлы Мессенджера".
НТТР код ответа
Заголовки ответа
  "cache-control": "no-cache",
"content-type": "application/json; charset=utf-8"
```

2. Получение метаинформации о файле или каталоге



3. Создание папки

Параметрь	ı			
Параметр	Значение	Описание	Тип параметра	Тип данных
path	C:\Users\User\Downloads	Путь к создаваемой папке.	query	string
fields		Список возвращаемых атрибутов.	query	string

```
Curl -X PUT --header 'Content-Type: application/json' --header 'Accept: application/json' --header 'Authorization: OAuth |

URL sanpoca

https://cloud-api.yandex.net/v1/disk/resources?path=C%3A%SCUser%SCUser%SCUser%SCDownloads

Teno orseta

{
    "message": "Указанный формат ресурса Диска \"\\Users\\User\\Downloads\" не корректен. Должен начинаться с /.",
    "description": "Specified path \"\\Users\\User\\Downloads\" has incorrect format",
    "error": "DiskPathFormatError"
    }

HTTP код ответа

400

Заголовки ответа

{
    "cache-control": "no-cache",
    "content-length": "276",
    "content-type": "application/json"
}
```

4. Удаление файла или папки

Параметры				
Параметр	Значение	Описание	Тип параметра	Тип данных
path	C:\Users\User\Downloads	Путь к файлу или папке.	query	string
fields		Список возвращаемых атрибутов.	query	string

```
Curl -X DELETE --header 'Accept: application/json' --header 'Authorization: OAuth y0_AgAAAABFn0TkAADLWwAAAAEEVgHLAAC9S1Z(

URL запроса

https://cloud-api.yandex.net/v1/disk/resources?path=C%3A%5CUser%5CUser%5CDownloads

Teno otbeta

{
    "message": "Указанный формат ресурса Диска \"\Users\User\Downloads\" не корректен. Должен начинаться с /.",
    "description": "Specified path \"\Users\User\Downloads\" has incorrect format",
    "error": "DiskPathFormatError"
}

HTTP код ответа

400

Заголовки ответа

{
    "cache-control": "no-cache",
    "content-length": "276",
    "content-type": "application/json"
    "content-type": "application/json"
```

5. Создание копии папки или файла

Параметры				
Параметр	Значение	Описание	Тип параметра	Тип данных
from	лр1.xlsx	Путь к копируемому ресурсу.	query	string
path	лр1_копия.xlsx	Путь к создаваемому ресурсу.	query	string

6. Получение ссылки на скачивание файла

Параметры				
Параметр	Значение	Описание	Тип параметра	Тип данных
path	C:\Users\User\Downloads\Telegram Desktop\лр	Путь к ресурсу.	query	string
fields		Список возвращаемых атрибутов.	query	string

7. Получение списка упорядоченного по имени

```
Curl
 curl -X GET --header 'Accept: application/json' --header 'Authorization: OAuth y0_AgAAAABFn0TkAADLWwAAAAEEVgHLAAC9SlZEpmi
4
URL запроса
 https://cloud-api.yandex.net/v1/disk/resources/files?fields=C%3A%5CUsers%5CUser%5CDownloads%5CTelegram%20Desktop
Тело ответа
   "items": [
       "antivirus_status": "clean",
       "size": 443403,
         "private_resource": "1168065764:78b81634175fb6a5f9e8938a937b6cf086383036b513b196680a0d50ccb81f5a",
         "public_resource": "1168065764:78b81634175fb6a5f9e8938a937b6cf086383036b513b196680a0d50ccb81f5a"
       },
       "name": "C20-7xx (1сем)_ЭК3_ВОПРОСЫ.pdf",
       "exif": {},
       "created": "2021-01-22T09:44:37+00:00",
       "resource_id": "1168065764:78b81634175fb6a5f9e8938a937b6cf086383036b513b196680a0d50ccb81f5a",
       "modified": "2021-01-22T09:44:37+00:00",
       "mime_type": "application/pdf",
       "sizes": [
           "url": "https://downloader.disk.yandex.ru/preview/599ed0e9d7984fa678f73959e8fbe4e8430cb22cf2c111948664d33be6
          "name": "DEFAULT"
```

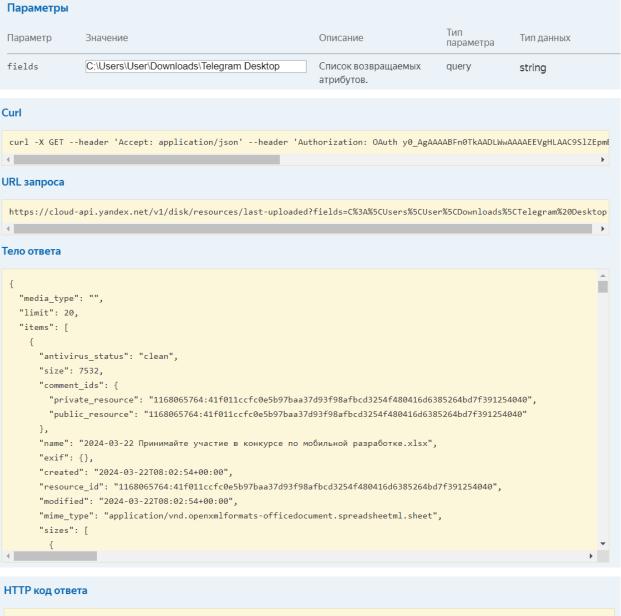
```
HTTP код ответа

200

Заголовки ответа

{ "cache-control": "no-cache", "content-type": "application/json; charset=utf-8" }
```

8. Получение списка файлов упорядоченных по дате загрузки



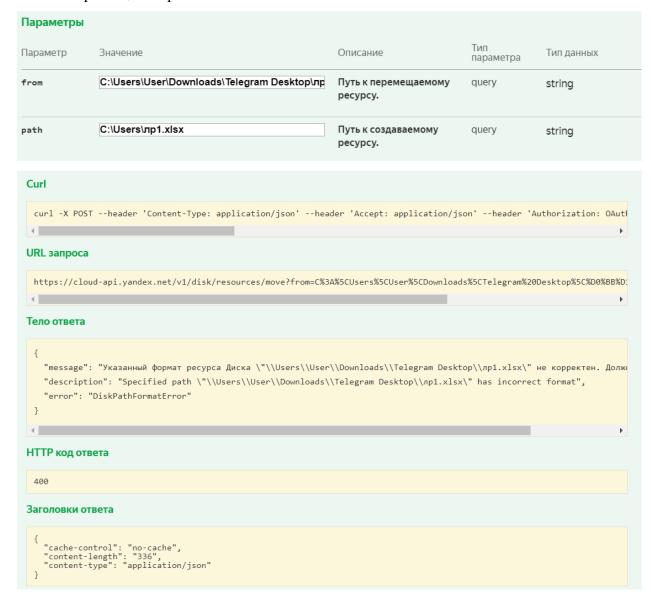
```
HTTP код ответа

200

Заголовки ответа

{ "cache-control": "no-cache", "content-type": "application/json; charset=utf-8" }
```

9. Перемещение файла или папки



10. Получение списка опубликованных ресурсов

```
Curl
 curl -X GET --header 'Accept: application/json' --header 'Authorization: OAuth y0_AgAAAABFn0TkAADLWwAAAAEEVgHLAAC9S1ZEpmf
URL запроса
 https://cloud-api.yandex.net/v1/disk/resources/public
Тело ответа
   "items": [
       "public_key": "St+ec50LQVaFZxgQNfQV4lWWsOggxYKw6G877zK6h8ou1CHvlxkyUlQHo8czumDnq/J6bpmRy0JonT3VoXnDag==",
      "name": "Адыгея 2022",
      "exif": {},
      "resource_id": "1168065764:fb8aac3dc60e8bcfd56f6dad460c2d6d3c9c89ce117dea3b32be7ae527b32a48",
      "public_url": "https://yadi.sk/d/qYKdCYSOiy3-DQ",
       "modified": "2022-09-10T17:23:39+00:00",
       "created": "2022-09-10T17:23:39+00:00",
       "path": "disk:/Адыгея 2022",
       "comment_ids": {
        "private_resource": "1168065764:fb8aac3dc60e8bcfd56f6dad460c2d6d3c9c89ce117dea3b32be7ae527b32a48",
        "public_resource": "1168065764:fb8aac3dc60e8bcfd56f6dad460c2d6d3c9c89ce117dea3b32be7ae527b32a48"
      },
       "type": "dir",
       "revision": 1662830673622230
   1.
   "limit": 20,
```

11. Публикация ресурса

12. Загрузка файла на диск по url

13. Получение ссылки для загрузки файла

Заключение

В ходе работы были изучены методы извлечения данных из социальной сети ВКонтакте и Яндекс. Диска с помощью АРІ запросов. На основе полученных знаний были осуществлены запросы по извлечению записей сообщества из социальной сети ВКонтакте, идентификаторов его участников и т.д., а также запросы по выводу информации о Яндекс. Диске и его содержимом, запросы по управлению файлами и папками.