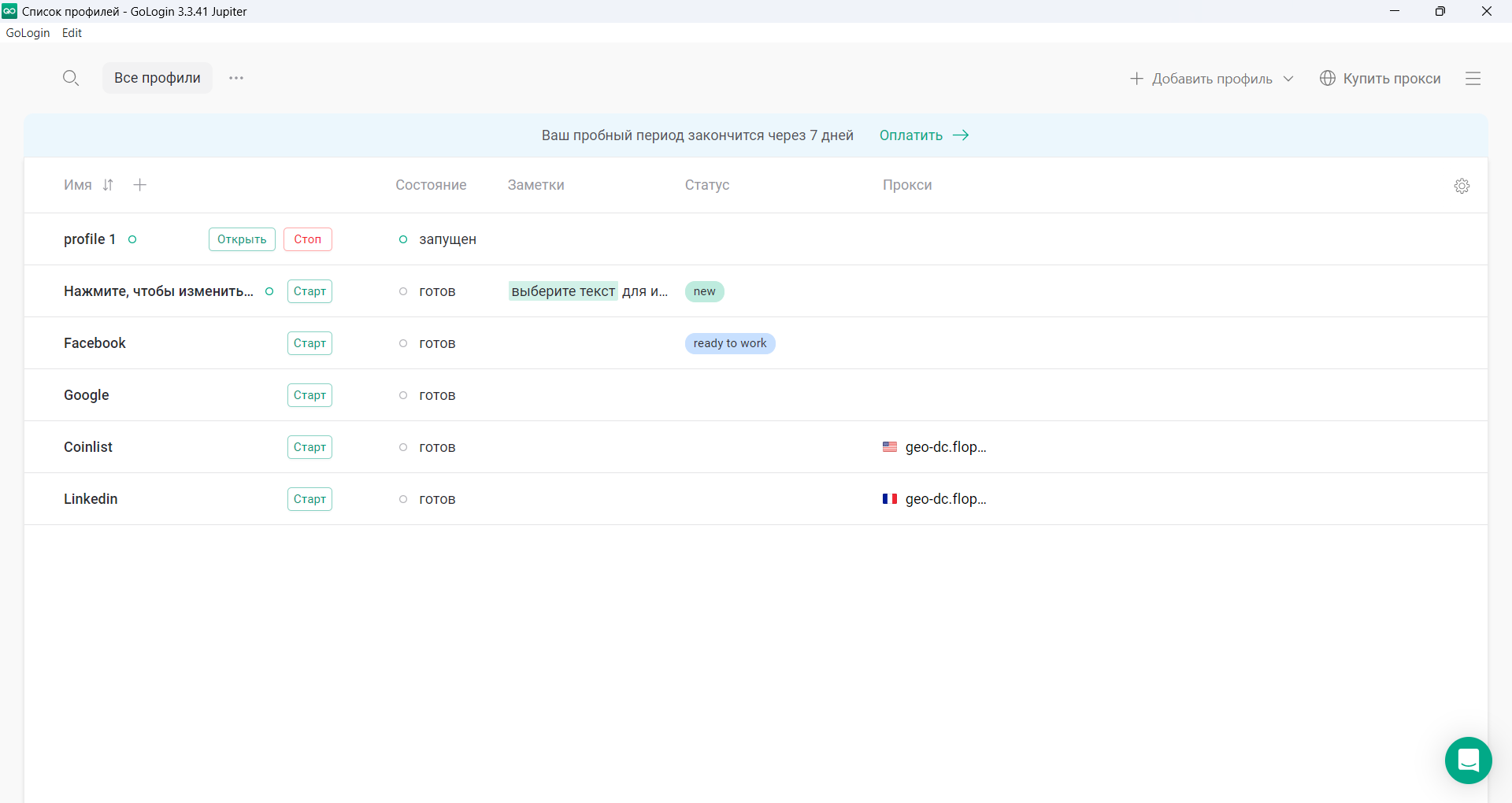
Попытка номер два по парсеру Facebook.

Шаг №1

По парсеру Facebook, попробуйте использовать GoLogin.

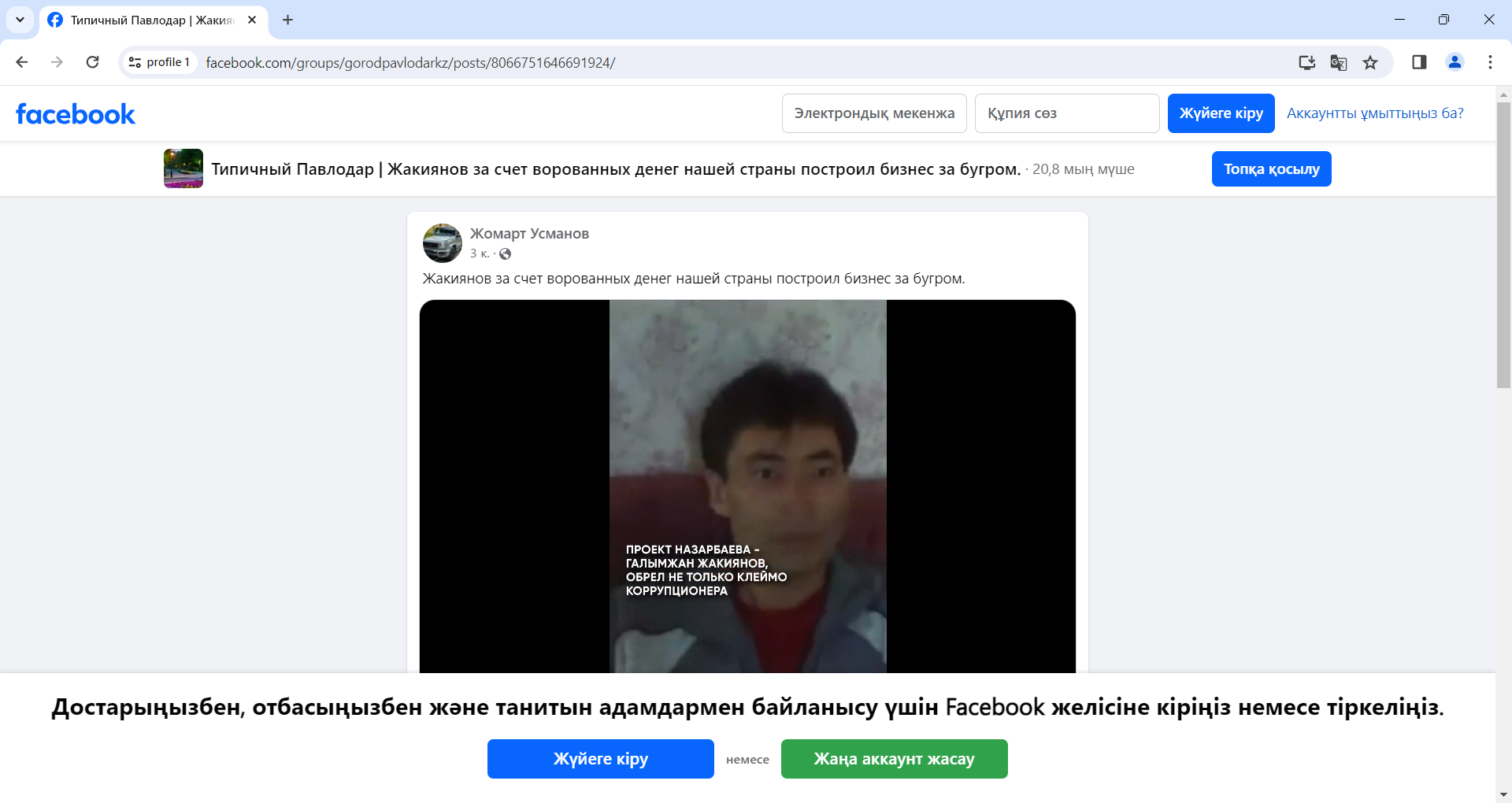
Попробую использовать GoLogin. Открыл GoLogin



Шаг №2

Добавил profile1 нажимаю на кнопку Старт. Открывается браузер. Вставляю ссылку

(https://www.facebook.com/groups/gorodpavlodarkz/posts/8066751646691924/)



Шаг №3

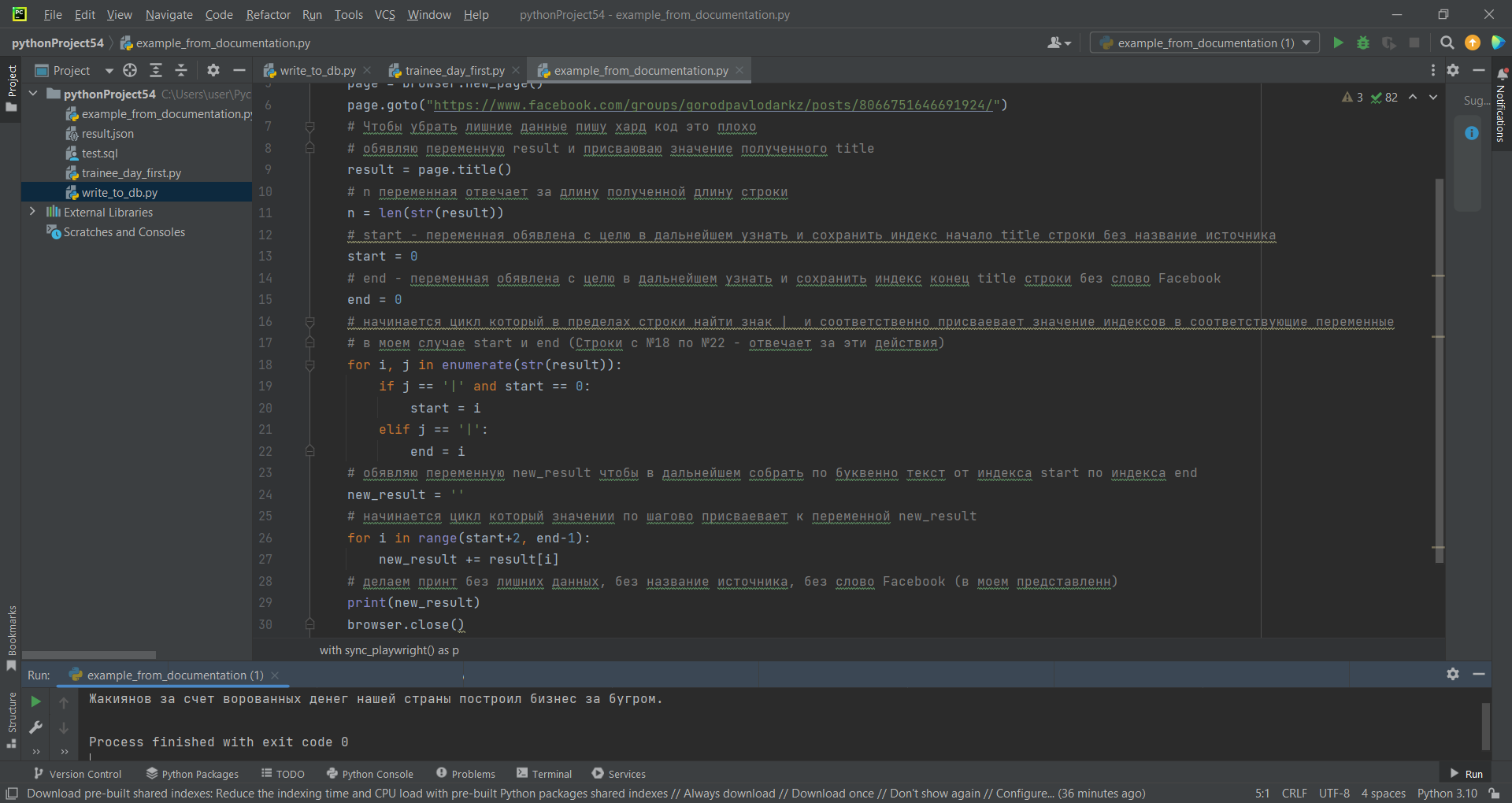
Иду ранее созданному парсеру, который до этого предоставляла. Запускаю. Этот созданный парсер в сторону GoLogin вообще не дышит. Код сам по себе. GoLogin браузер сам по себе. Никаких изменений. (Что я делаю не так)

Шаг №4

Для начала нужно авторизоваться. после чего уже парсить данные. (Вопрос - где я должна авторизоваться чтобы парсить данные)

Шаг №5

Еще, к примеру, тут выгружен не полный текст. И нужно убрать лишние данные, название источника, слово Facebook, т.д. и оставить только текст.



from playwright.sync\_api import sync\_playwright

with sync\_playwright() as p:

browser = p.chromium.launch()

page = browser.new\_page()

page.goto("https://www.facebook.com/groups/gorodpavlodarkz/posts/8066751646691924/")

# Чтобы убрать лишние данные пишу хард код это плохо

# обявляю переменную result и присваюваю значение полученного title

result = page.title()

# n переменная отвечает за длину полученной длину строки

n = len(str(result))

# start - переменная обявлена с целю в дальнейшем узнать и сохранить индекс начало title строки без название источника

start = 0

# end - переменная обявлена с целю в дальнейшем узнать и сохранить индекс конец title строки без слово Facebook

end = 0

# начинается цикл который в пределах строки найти знак | и соответственно присваевает значение индексов в соответствующие переменные

# в моем случае start и end (Строки с №18 по №22 - отвечает за эти действия)

for i, j in enumerate(str(result)):

if j == '|' and start == 0:

start = i

elif j == '|':

end = i

# обявляю переменную new\_result чтобы в дальнейшем собрать по буквенно текст от индекса start по индекса end

new\_result = ''

# начинается цикл который значении по шагово присваевает к переменной new\_result

for i in range(start+2, end-1):

new\_result += result[i]

# делаем принт без лишних данных, без название источника, без слово Facebook (в моем представленн)

print(new\_result)

browser.close()