



**МІНІСТЕРСТВО ОСВІТИ, НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КІЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ**

Лабораторна робота №4

Аналіз даних

Підготував:

студент 4 курсу

групи ФІ-84

Коломієць Андрій Юрійович

E-mail: andrew.kolomiets.work@gmail.com

Київ – 2021

Лабораторна робота №4

Аналіз даних

Завдання на самостійну роботу

1. Забезпечити функціонування на комп'ютері (сервері) засобів інтерактивної взаємодії з користувачем (наприклад, **CGI**).
2. Ознайомитися із можливостями пошуку, фільтрації в **Elasticsearch**.
3. Ознайомитись з основами **HTML**-розмітки.

Виконання лабораторної роботи

Запити до Elasticsearch

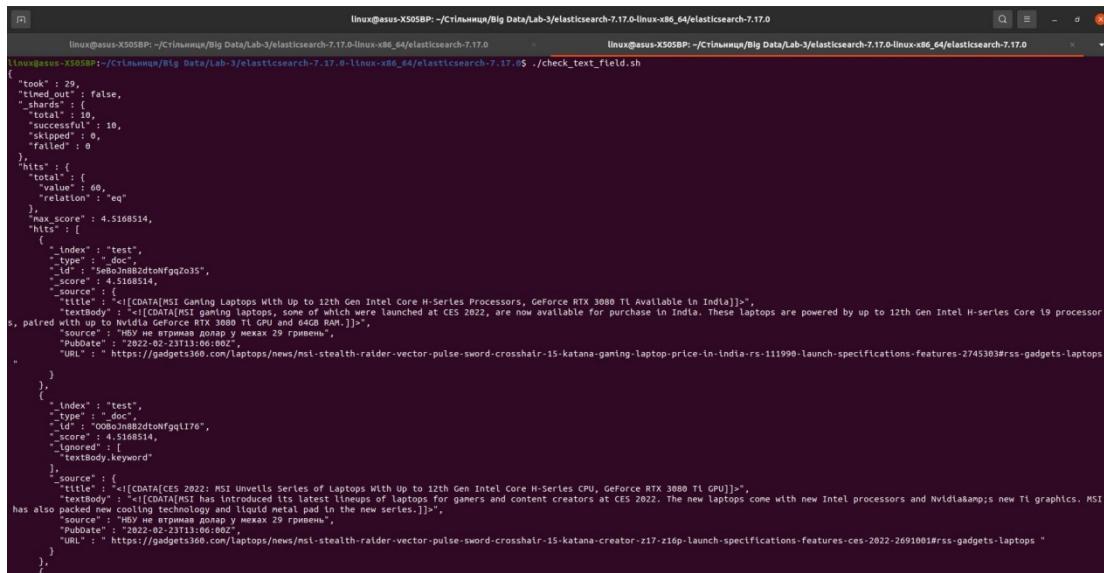
Зauważення: у минулій лабораторній роботі скрипт не відреагував на пошуковий запит, і вернув пусте поле, через те що даних таких не існувало за запитом. Нижче наводимо складнішу конструкцію пошукового запиту, для того, щоб упевнитися в правоті попередньо виконаної роботи.

Коли кінцевий користувач шукає будь-які терміни, іноді потрібно виконувати запити по двох полях "**textBody**" (тіло документа) та "**title**" (назва). Це можливо за допомогою запиту на пошук у декількох полях. Наступний запит знайде всі документи, які мають слово **Laptops** у полях назви або опису:

```
linux@asus-X505BP:~/Стільниця/Big Data/Lab-3/elasticsearch-7.17.0-linux-x86_64/elasticsearch-7.17.0$ cat check_text_field.sh ; echo
#!/bin/bash

curl -XGET 'http://localhost:9200/_all/_search?pretty=true&size=100' -H 'Content-Type: application/json' -d '
{
    "query": {
        "bool": {
            "must": [
                {
                    "multi_match": { "query": "Laptops",
                        "fields": ["textBody", "title"]
                    }
                }
            ]
        }
    }
}'
```

Вводимо команду та бачимо результат запиту

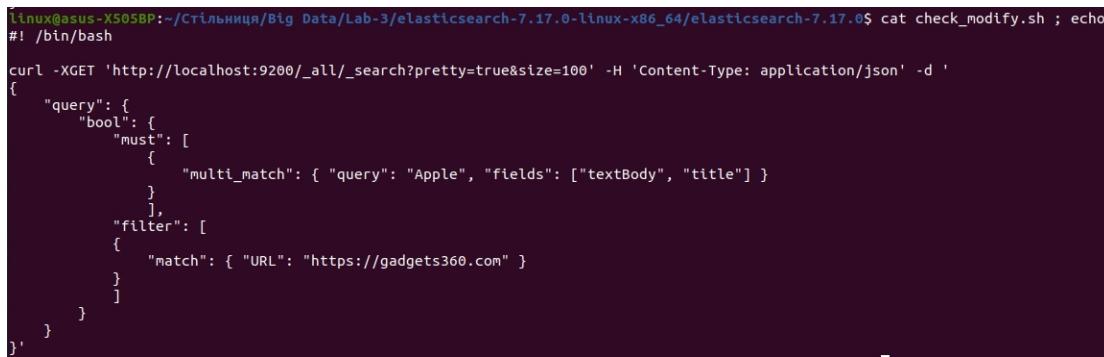


The screenshot shows two terminal windows side-by-side. Both windows have the title 'linux@asus-X505BP:~/Стільнича/Big Data/Lab-3/elasticsearch-7.17.0-linux-x86_64/elasticsearch-7.17.0'. The left window contains the command 'curl -XGET 'http://localhost:9200/_all/_search?pretty=true&size=100' -H 'Content-Type: application/json' -d @check_text_field.sh'. The right window displays the JSON response from the Elasticsearch search query. The response includes metrics like 'took', 'timed_out', 'shards', 'total', 'successful', 'skipped', and 'failed'. It also shows the 'hits' section with 'total' documents (60), 'max_score' (4.5168514), and individual hits. Each hit object contains '_index', '_type', '_id', '_score', '_source', and '_textBody'. The '_source' field for the first hit is a long string of XML-like data about MSI gaming laptops.

```
linux@asus-X505BP:~/Стільнича/Big Data/Lab-3/elasticsearch-7.17.0-linux-x86_64/elasticsearch-7.17.0
linux@asus-X505BP:~/Стільнича/Big Data/Lab-3/elasticsearch-7.17.0-linux-x86_64/elasticsearch-7.17.0$ curl -XGET 'http://localhost:9200/_all/_search?pretty=true&size=100' -H 'Content-Type: application/json' -d @check_text_field.sh
{
  "took": 29,
  "timed_out": false,
  "shards": 1,
  "total": 10,
  "successful": 10,
  "skipped": 0,
  "failed": 0
},
{
  "hits": [
    {
      "total": {
        "value": 60,
        "relation": "eq"
      },
      "max_score": 4.5168514,
      "hits": [
        {
          "_index": "test",
          "_type": "doc",
          "_id": "98e082d0aNfgqZo35",
          "_score": 4.5168514,
          "_source": {
            "title": "<![CDATA[MSI Gaming Laptops With Up to 12th Gen Intel Core H-Series Processors, GeForce RTX 3080 Ti Available In India]]>",
            "textBody": "<![CDATA[MSI gaming laptops, some of which were launched at CES 2022, are now available for purchase in India. These laptops are powered by up to 12th Gen Intel H-series Core i9 processor s, paired with up to Nvidia GeForce RTX 3080 Ti GPU and 64GB RAM.]]>",
            "source": "НВУ не втримав долар у межах 29 гривень",
            "PubDate": "2022-02-23T13:06:00Z",
            "URL": "https://www.gadgets360.com/laptops/news/msi-stealth-raider-vector-pulse-sword-crosshair-i5-katana-gaming-laptop-price-in-india-rs-111990-launch-specifications-features-274503#rss-gadgets-laptops"
          }
        },
        {
          "_index": "test",
          "_type": "doc",
          "_id": "98e082d0aNfgql176",
          "_score": 4.5168514,
          "_ignored": [
            "textBody.keyword"
          ],
          "_source": {
            "title": "<![CDATA[CES 2022: MSI Unveils Series of Laptops With Up to 12th Gen Intel Core H-Series CPU, GeForce RTX 3080 Ti GPU]]>",
            "textBody": "<![CDATA[MSI has introduced its latest lineups of laptops for gamers and content creators at CES 2022. The new laptops come with new Intel processors and Nvidia's new Ti graphics. MSI has also added new cooling technology with a liquid metal pad in the new series.]]>",
            "source": "НВУ не втримав долар у межах 29 гривень",
            "PubDate": "2022-02-23T13:06:00Z",
            "URL": "https://www.gadgets360.com/laptops/news/msi-stealth-raider-vector-pulse-sword-crosshair-i5-katana-creator-z17-z16p-launch-specifications-features-ces-2022-269100#rss-gadgets-laptops"
          }
        }
      ]
    }
  ]
}
```

Вивід занадто довгий тому приводиться лише його частина.

Нижче наведено ще один запит, який звертається до декількох полів за словом **Apple**. Крім того у запиті на пошук здійснюється фільтрація за джерелом **www.gadgets360.com**:



The screenshot shows a single terminal window with the title 'linux@asus-X505BP:~/Стільнича/Big Data/Lab-3/elasticsearch-7.17.0-linux-x86_64/elasticsearch-7.17.0\$ cat check_modify.sh ; echo "#! /bin/bash"'. The command executed is 'curl -XGET 'http://localhost:9200/_all/_search?pretty=true&size=100' -H 'Content-Type: application/json' -d @check_modify.sh'. The 'check_modify.sh' file contains a complex Elasticsearch search query. It uses a 'query' block with a 'bool' query, which includes a 'must' clause with a 'multi_match' query for the 'Apple' term across 'textBody' and 'title' fields, and a 'filter' clause with a 'match' query for the URL 'https://www.gadgets360.com'.

```
linux@asus-X505BP:~/Стільнича/Big Data/Lab-3/elasticsearch-7.17.0-linux-x86_64/elasticsearch-7.17.0$ cat check_modify.sh ; echo "#! /bin/bash"
curl -XGET 'http://localhost:9200/_all/_search?pretty=true&size=100' -H 'Content-Type: application/json' -d @check_modify.sh
{
  "query": {
    "bool": {
      "must": [
        {
          "multi_match": { "query": "Apple", "fields": [ "textBody", "title" ] }
        }
      ],
      "filter": [
        {
          "match": { "URL": "https://www.gadgets360.com" }
        }
      ]
    }
  }
}'
```

Маємо відповідний результат

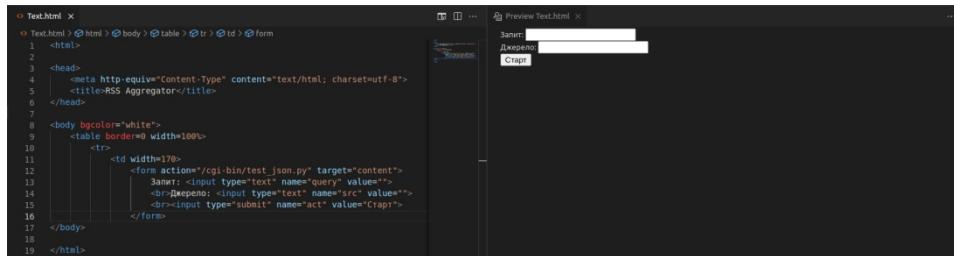
```
linuxstatus-X550SP:~/Cytinewp/Big Data/Lab-3/elasticsearch-7.17.0-linux-x64/elasticsearch-7.17.0$ ./check_modify.sh
{
  "took" : 89,
  "timed_out" : false,
  "_shards" : {
    "total" : 10,
    "successful" : 10,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 64,
      "relation" : "eq"
    },
    "max_score" : 4.8387365,
    "hits" : [
      {
        "_index" : "test",
        "_type" : "doc",
        "_id" : "G5t4J3B8BdwCOZSwNfLY",
        "_score" : 4.8387365,
        "_ignored" : false,
        "_textBody.keyword" : "textBody.keyword"
      },
      {
        "_source" : {
          "title" : "<![CDATA[Apple MacBook Pro 13-Inch Model With M2 Chip to Launch on March 8: Report]]>",
          "textBody" : "<![CDATA[Apple&#39;s upcoming entry-level MacBook Pro – said to be powered by the company&amp;s M2 silicon chipset – will reportedly launch on March 8 during Apple&amp;s Spring Launch event. Apple&amp;s supply chain partners reportedly continued production during the Lunar New Year holidays.]]>",
          "source" : "MSN News",
          "PubDate" : "2022-02-23T13:06:00Z",
          "URL" : "https://gadgets360.com/laptops/news/apple-macbook-pro-m2-launch-march-8-spring-event-iphone-se-ipad-report-2758002#rss-gadgets-laptops"
        }
      },
      {
        "_index" : "test",
        "_type" : "doc",
        "_id" : "G5t4J3B8BdwCOZSwNfLY09",
        "_score" : 4.6391187,
        "_ignored" : false,
        "_textBody.keyword" : "textBody.keyword"
      },
      {
        "_source" : {
          "title" : "<![CDATA[Apple Spotted Registering Three New Mac Models on EEC Ahead of Rumoured March 8 Event]]>",
          "textBody" : "<![CDATA[Three Apple Mac PCs have been spotted on Eurasian Economic Commission (EEC) website, and one of them is said to be a laptop, as per a report by a French publication. Apple is tipped to launch an updated version of its entry-level 13-inch MacBook Pro Apple M2 processor and a MacBook Air with M2 processor.]]>",
          "source" : "MSN News",
          "PubDate" : "2022-02-23T13:06:00Z",
          "URL" : "https://gadgets360.com/laptops/news/apple-macbook-pro-air-spotted-eec-launch-march-8-iphone-se-ipad-report-2771999#rss-gadgets-mobiles"
        }
      },
      {
        "_index" : "test",
        "_type" : "doc",
        "_id" : "G5t4J3B8BdwCOZSwNfLY09",
        "_score" : 4.5453606,
        "_ignored" : false,
        "_textBody.keyword" : "textBody.keyword"
      },
      {
        "_source" : {
          "title" : "<![CDATA[macos 12.2.1 With Fix for Exploited WebKit Flaw Out Now, Resolves Battery Drain on Intel-Based MacBook Models]]>",
          "textBody" : "<![CDATA[Apple has been pushing out updates to macOS 12.2.1 to users, bringing important security and bug fixes in tow. The latest update fixes an important security flaw that Apple says might have been actively exploited. An update to macOS 12.2.1 also fixes a Bluetooth bug affecting Intel-based MacBook models, while Apple Silicon Macbook owners may have to wait until macOS 12.3 for a...]]>",
          "source" : "MSN News",
          "PubDate" : "2022-02-23T13:06:00Z",
          "URL" : "https://gadgets360.com/laptops/news/macos-12-2-1-update-monterey-webkit-security-fix-battery-drain-intel-macbook-beta-emoji-apple-2762361#rss-gadgets-laptops"
        }
      }
    ]
  }
}
```

На жаль в програмі є дублікати та помітно поле **“source”** співпадає в багатьох місцях. На попередніх кроках виконання лабораторних робіт напевне допущено помилки але знаходяться данні котрі досить чітко відрізняються, потрібно реалізувати скрипт, що видаляє дублікати.

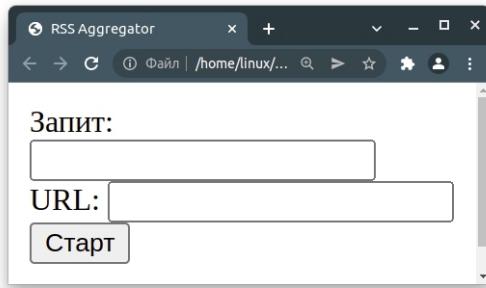
Модель пошукової системи на базі веб-інтерфейсу

Нижче наведено форму мовою **HTML (CGI-форму)** моделі пошукової системи:

Visual Studio Code



Chrome Browser



Наведена **CGI-форма** містить виклик **CGI-процедури** на мові **Python** (**test_json.py**) із приймання параметрів і виклику запиту до системи **Elasticsearch**:

```
#!/usr/local/bin/python2.7
import os
import re
import cgi
import subprocess

form = cgi.FieldStorage()
query = form.getFirst("query", "Query not defined")
src = form.getFirst("src", "Source not defined")
os.system('test_json.sh %s %s'
          % (query, src))
```

Shell-скрипт, що забезпечує передачу запиту системі **Elasticsearch** (**test_json.sh**) і подальший виклик програми форматування відповіді (**prog2_html**):

```
#!/bin/bash
query=$1
src=$2
/usr/local/bin/curl -XGET 'http://localhost:9200/_all/_search?pretty=true&size=20' -H 'Content-Type: application/json' -d '
{
  "query": {
    "bool": {
      "must": [
        {
          "multi_match": {
            "query": "$query",
            "fields": ["textBody", "title"]
          }
        }
      ],
      "filter": [
        {
          "match": {
            "URL": "$src"
          }
        }
      ]
    }
  }
}' | prog2_html
```

Програма форматування JSON-файлу

```
#!/usr/bin/python3

import sys
import re
import datetime

t = sys.stdin.read()
json = t.split('\n')
t = ""

for i in range(len(json)):
    t += json[i]

t = re.sub('^\s+', '', t)
total = re.findall('"total" : (\d+)', t)

print("Content-type: text/html; charset=utf-8\n\n<html><body bgcolor=yellow>")
print("<b>Found: " + total[0] + "</b><hr><ol>")
title = re.findall('"title" : "(.*?)"', t)
text = re.findall('"textBody" : "(.*?)"', t)
url = re.findall('"URL" : "(.*?)"', t)

for i in range(len(title)):
    doc_id = i + 1
    print("<li><b>" + title[i] + "</b>")
    print("<br>" + text[i])
    print("<br><i>" + url[i] + "</i><br /><hr>")
print("</ol></body></html>")
```

Для запуску веб-інтерфейсу потрібно встановити apache за допомогою рекомендацій з довільних документацій поданих нижче:

<https://httpd.apache.org/docs/2.4/>

<https://ubuntu.com/tutorials/install-and-configure-apache#1-overview>

Особливо важливо, щоб виконувалися скрипти CGI:

<https://httpd.apache.org/docs/2.4/howto/cgi.html>

<https://www.frc1073.org/configure-cgi>

<https://code-maven.com/set-up-cgi-with-apache>

Коли встановимо **apache2**, та введемо в пошуковику **IP** адресу, ми отримаємо стартову сторінку **apache2** з коротким її описом. Даній сторінка матиме назву **index.html** та знаходитиметься в каталозі **/var/www/html**.

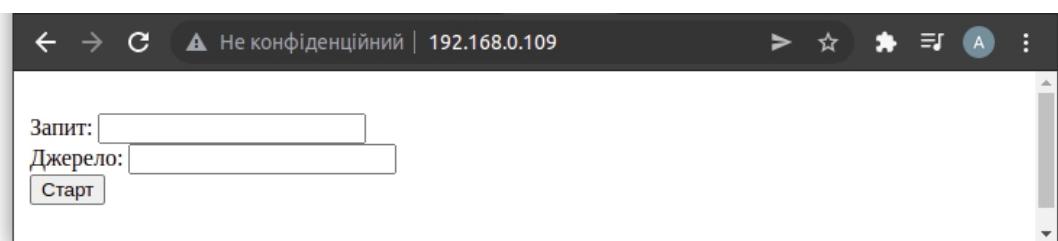
The screenshot shows the Apache2 Ubuntu Default Page. At the top, there's a logo for Ubuntu and the text "Apache2 Ubuntu Default Page". Below that, it says "It works!". The page contains several sections:

- Configuration Overview**: Describes the configuration layout for an Apache2 web server installation on Ubuntu systems, mentioning files like /etc/apache2/apache2.conf, ports.conf, mods-enabled, conf-enabled, and sites-enabled.
- Document Roots**: Explains that by default, Ubuntu does not allow access through the web browser to any file apart from those located in /var/www and /public_html. It also mentions /usr/share for web applications.
- Reporting Problems**: Instructions for reporting bugs using the ubuntu-bug tool, mentioning existing bug reports and specific modules.

Завантажимо до каталогу **apache2** власну форму.

```
linux@linux-X505BP:~/Стільниця/Big-data/Lab-4/cgi-bin$ sudo cp index.html /var/www/html
```

Перезавантажимо сторінку та побачимо, що під **IP** адресою, відображається інша веб-сторінка.



Тепер необхідно зробити програми виконуваними в **apache2**, для цього

необхідно сконфігурувати певний каталог для цього, але також зазначити конфігурації необхідні для цього, або знайти каталог який вже можна використовувати для того , щоб файли виконувалися в ньому.

Згідно конфігураційного файлу **apache2** таким каталогом є **/usr/lib/cgi-bin**

```
linux@linux-X505BP:/usr/lib/cgi-bin$ cat /etc/apache2/conf-enabled/serve-cgi-bin.conf
<IfModule mod_alias.c>
    <IfModule mod_cgi.c>
        Define ENABLE_USR_LIB_CGI_BIN
    </IfModule>

    <IfModule mod_cgid.c>
        Define ENABLE_USR_LIB_CGI_BIN
    </IfModule>

    <IfDefine ENABLE_USR_LIB_CGI_BIN>
        ScriptAlias /cgi-bin/ /usr/lib/cgi-bin/
        <Directory "/usr/lib/cgi-bin">
            AllowOverride None
            Options +ExecCGI -MultiViews +SymLinksIfOwnerMatch
            Require all granted
        </Directory>
    </IfDefine>
</IfModule>

# vim: syntax=apache ts=4 sw=4 sts=4 sr noet
```

Відповідно в зазначеній каталог переміщаємо всі наші фали котрі мають виконуватися, також відредактувати потрібно всі шляхи в файлах на виконання.

Файл **/etc/apache2/conf-available/serve-cgi-bin.conf** , який містить символічне посилання з **/etc/apache2/conf-enabled/serve-cgi-bin.conf**. Він має розділ, який зіставляє шлях **/cgi-bin** в URL-адресах до каталогу **/usr/lib/cgi-bin/** на жорсткому диску і дозволяє виконувати **CGI** в цьому каталозі.

```
lrwxrwxrwx 1 root root 36 бер 19 09:03 serve-cgi-bin.conf -> ../conf-available/serve-cgi-bin.conf
```

```
linux@linux-X505BP:/etc/apache2/conf-available$ ll
загалом 32
drwxr-xr-x 2 root root 4096 бер 19 09:02 .
drwxr-xr-x 8 root root 4096 бер 19 09:03 ..
-rw-r--r-- 1 root root 315 хов 1 2020 charset.conf
-rw-r--r-- 1 root root 127 сер 11 2013 javascript-common.conf
-rw-r--r-- 1 root root 3224 хов 1 2020 localized-error-pages.conf
-rw-r--r-- 1 root root 189 хов 1 2020 other-vhosts-access-log.conf
-rw-r--r-- 1 root root 2174 хов 1 2020 security.conf
-rw-r--r-- 1 root root 455 хов 1 2020 serve-cgi-bin.conf
```

Перевіримо згідно документації дозвіл на виконання в зазначеному каталозі.

Configuring Apache to permit CGI

In order to get your CGI programs to work properly, you'll need to have Apache configured to permit CGI execution. There are several ways to do this.

Note: If Apache has been built with shared module support you need to ensure that the module is loaded; in your httpd.conf you need to make sure the `LoadModule` directive has not been commented out. A correctly configured directive may look like this:

```
LoadModule cgid_module modules/mod_cgid.so
```

On Windows, or using a non-threaded MPM like prefork, A correctly configured directive may look like this:

```
LoadModule cgi_module modules/mod_cgi.so
```

```
linux@linux-X505BP:/usr/lib/cgi-bin$ locate cgi.load
/etc/apache2/mods-available/authnz_fcgi.load
/etc/apache2/mods-available/cgi.load
/etc/apache2/mods-available/proxy_fcgi.load
/etc/apache2/mods-available/proxy_scgi.load
linux@linux-X505BP:/usr/lib/cgi-bin$ cat /etc/apache2/mods-available/cgi.load
LoadModule cgi_module /usr/lib/apache2/modules/mod_cgi.so
```

Модуль **cgi** називається `cgi.load`. Встановлені модулі **Apache** можна знайти в каталозі `/etc/apache2/mods-available`.

Увімкнені модулі мають символічні посилання в `/etc/apache2/mods-enabled`, але, як я з'ясував, модуль **CGI** не мав там символічного посилання: модуль **CGI** не був увімкнений за замовчуванням.

```
$ cd /etc/apache2/mods-enabled
$ sudo ln -s ../mods-available/cgi.load
```

Додано символічне посилання.

Перезавантажте конфігурацію **Apache**.

Оскільки конфігурація **Apache** змінилася, нам потрібно вказати **Apache** перезавантажити свої файли конфігурації.

```
sudo service apache2 reload
```

Переглянемо каталоги необхідні **apache2**.

```
linux@linux-X505BP:/var/www/html$ ll
загалом 12
drwxr-xr-x 2 root root 4096 бер 19 10:36 .
drwxr-xr-x 3 root root 4096 бер 19 09:02 ..
-rwxr-xr-x 1 root root 617 бер 19 10:36 index.html*
```

```
linux@linux-X505BP:/usr/lib/cgi-bin$ ll
загалом 24
drwxr-xr-x  2 root root 4096 бер 19 11:28 .
drwxr-xr-x 119 root root 4096 бер 19 09:49 ../
-rwxr-xr-x  1 root root  626 бер 19 11:28 index.html*
-rwxr-xr-x  1 root root  658 бер 19 11:28 prog2_html.py*
-rwxr-xr-x  1 root root 263 бер 19 11:28 test_json.py*
-rwxr-xr-x  1 root root 622 бер 19 11:28 test_json.sh*
linux@linux-X505BP:/usr/lib/cgi-bin$
```

Спробуємо запустити скрипт та отримаємо помилку.



Наш скрипт знаходить але він має синтаксичні помилки.

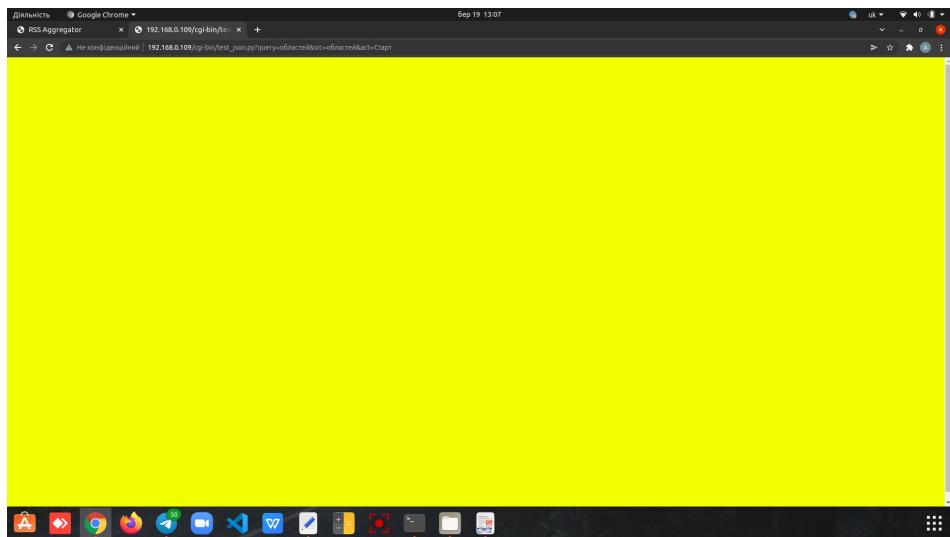
Якщо запити **curl** повертають **500** внутрішньої помилки сервера , можливо, ви забули зробити сценарій виконуваним за допомогою **sudo chmod +x**, або рядок **hashbang** (перший рядок, який має виглядати так **#!/bin/bash** , був введений неправильно.

Іншим поширеним джерелом помилок є різниця між новими рядками в **MS Windows** і **Linux** у текстових файлах . Ви повинні переконатися, що ваш сценарій має закінчення рядка **Unix/Linux**. Одним із найкращих способів може бути запуск **dos2unix test.sh** , який замінить **Windows**-нових рядків на **Unix**-нових рядків.

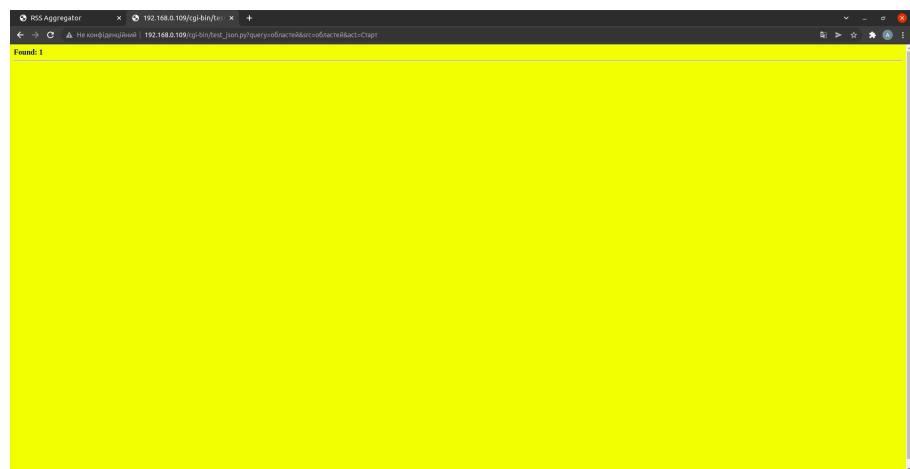
Нарешті, помилка **500** може виникнути, якщо перший рядок, надрукований сценарієм, не є **Content-type: text/html** , за яким слідують **2(!)** нові рядки.

Інколи є синтаксичні помилки котрі не відразу можна визначити, тому подальші запуски програми були пов'язані з визначенням цих помилок.

Наприклад можна було спіткнутися з такими наступними виводами.



Вказаний скріншот, говорить про те що невірний код пов'язаний з виводом **HTML** сторінки. Після деякий виправлень маємо.



В поданому скріншоті, вже все працює правильно, але код не реагує на пошуковий запит, тому не відбувається вивід інформації. Що діється саме тут можна глянути на скріншот.

```
{
  "took" : 3,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 0,
      "relation" : "eq"
    },
    "max_score" : null,
    "hits" : [ ]
  }
}
```

Це свідчить про те, що пошук відбувся але результату ми не отримали ніякого.

Перевірка файлів в **/cgi-bin** дозволяє віднайти помилки, котрі були пов'язані з передачею аргументів в код файлу.

```
| GNU nano 4.8                                     test_json.sh
#!/bin/sh

curl -XGET 'http://localhost:9200/test/_search?pretty=true' -H 'Content-Type: application/json' -d '
{
    "query": {
        "bool": {
            "must": [
                {
                    "multi_match": {
                        "query": "$1",
                        "fields": ["textBody", "title"]
                    }
                }
            ],
            "filter": [
                {
                    "match": {
                        "source": "$2"
                    }
                }
            ]
        }
    }
}' | /usr/lib/cgi-bin/prog2_html.py
```

Слід замінити було змінні на $\$1$ та $\$2$.

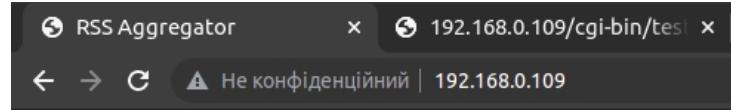
Здійснимо перевірку.

```
linuxlmx-UX505R:/usr/lib/cgi-bin$ ./test_jon.sh 'apple' 'gadget'
* Total % Received % Xfered Average Speed Time Time Current
          Dload Upload Total Spent Left Speed
108 7478 100 7063 100 415 200088 1179 -.-:--:--:--:--:--:--:--:--:--:21184
Content-type:text/html

<html>
<head>
<meta http-equiv="Content-type" charset="utf-8" />
<body bcolor="#yellow">
<br/>Realme GT Pro India Launch Tipped for March, Specifications Expected to Include Snapdragon 8 Gen 1 Soc, 1TB RAM </b>
<br/>Realme GT Pro India Launch Tipped to Launch in India in March. The smartphone was launched in China in January, and made its debut in global markets on February 28 at MWC 2022 alongside the Realme GT 2 smartphone. It is powered by a Snapdragon 8 Gen 1 Soc and comes with a 6.7-inch LTPO 2.0 AMOLED display. Realme is yet to officially announce the launch date for the Realme GT ... </b>
<br/><a href="https://www.gadgets360.com/mobiles/news/realme-gt-2-pro-india-launch-date-march-2022-leak-specifications-features-2887884#rss-gadgets-mobiles"></a><br /><br />
<br/><b> Asus ExpertBook B3 Detachable 2-in-1 Laptop With Snapdragon 7c Gen 2 SoC Launched at CES 2022 </b>
<br/>Asus has announced the ExpertBook B3 Detachable 2-in-1 laptop, which runs on Qualcomm's Snapdragon 7c Gen 2 SoC. Launched at CES 2022, the ExpertBook B3 is a 10.5-inch WUXGA LCD display with 120Hz peak brightness. Its price and availability have not yet been announced. </b>
<br/><a href="https://www.gadgets360.com/laptops/news/asus-expertbook-b3-detachable-2-in-1-laptop-specification-snapdragon-7c-gen-2-soc-cs-2022-2696451#rss-gadgets-laptops"></a><br /><br />
<br/><b> Netflix March 2022 Releases: Bridgerton Season 2, Venom: Let There Be Carnage, Outlander Season 6, and More </b>
<br/>Netflix has announced its March 2022 releases, which include the second season of Bridgerton, the third season of Outlander, the sixth season of the comic book movie adaptation of Eternally Confused and Eager for Love, Formula 1: Drive to Survive season 4, Taarak Mehta Ka Ooltah Chashmah season 2, Jai Lalla, Moon Knights in Woodland, the biggest movies and TV shows on Netflix in March 2022. </b>
<br/><a href="https://www.gadgets360.com/entertainment/features/netflix-march-2022-releases-india-bridgerton-2-venom-adam-project-outlander-6-formula-1-2789498#rss-gadgets-apps"></a><br /><br />
<br/><b> Bridgerton Season 2, Jai Lalla, Moon Knight, and More: March Guide to Netflix, Disney+ Hotstar, Prime Video </b>
<br/>Bridgerton season 2, Jai Lalla, Moon Knight, and Adam Project, Audra, Edie, and the rest of the Bridgerton family, Jai Lalla, Moon Knights in Woodland, the biggest movies, web series and TV shows on Netflix in March 2022. </b>
<br/><a href="https://www.gadgets360.com/entertainment/features/bridgerton-jai-lalla-moon-knight-march-2022-netflix-hotstar-prime-video-ott-movies-web-series-2796145#rss-gadgets-internet"></a><br /><br />
<br/><b> Bridgerton Season 2, Jai Lalla, Moon Knight, and More: March Guide to Netflix, Disney+ Hotstar, Prime Video </b>
<br/>Bridgerton season 2, Jai Lalla, Moon Knights in Woodland, the biggest movies, web series and TV shows on Netflix in March 2022. </b>
<br/><a href="https://www.gadgets360.com/entertainment/features/bridgerton-jai-lalla-moon-knight-march-2022-netflix-hotstar-prime-video-ott-movies-web-series-2796145#rss-gadgets-apps"></a><br /><br />
<br/><b> Samsung Galaxy Book 2 Pro, Galaxy Book 2 Pro 360 India Launch Tipped by Amazon, Official Website Listing </b>
<br/>Samsung Galaxy Book 2 Pro and Samsung Galaxy Book 2 Pro 360 are making their way to India, shortly after the company's first laptop, while the Galaxy Book 2 Pro 360 is a thin 2-in-1 laptop with P5 support. Samsung is yet to reveal when these laptops will be launched. </b>
<br/><a href="https://www.gadgets360.com/laptops/news/samsung-book-2-pro-360-india-launch-amazon-listing-website-official-specifications-galaxy-201145#rss-gadgets-laptops"></a><br /><br />
</ol></body></html>
```

Передача аргументів здійснюється успішно.

Спробуємо перезавантажити сторінку.



Запит: ['Asus']
Джерело: ['gadget']
Старт

← → C ▲ Не конфіденційний | 192.168.0.109

Found: 6

1. : Realme GT 2 Pro India Launch Tipped for March. Specifications Expected to Include Snapdragon 8 Gen 1 SoC, 12GB RAM
Realme GT 2 Pro is tipped to launch in India in March. The smartphone was launched in China in January, and made its debut in global markets on February 28 at MWC 2022 alongside the Realme GT 2 smartphone. It is powered by a Snapdragon 8 Gen 1 SoC and comes with a 6.7-inch LTPO 2.0 AMOLED display. Realme is yet to officially announce the launch date for the Realme GT.
<https://gadgets360.com/mobiles/news/realme-gt-2-pro-india-launch-date-march-2022-leak-specifications-features-2807894#rss-gadgets-mobiles>

2. : Asus ExpertBook B3 Detachable 2-in-1 Laptop With Snapdragon 7c Gen 2 SoC Launched at CES 2022
Asus ExpertBook B3 Detachable 2-in-1 laptop was silently launched on Friday during CES 2022. The 2-in-1 laptop from Asus is powered by Qualcomm Snapdragon 7c Gen 2 SoC, paired with up to 8GB of RAM and 128GB of eMMC storage. It sports a 10.5-inch WUXGA LCD display with 320 nits of peak brightness. Its price and availability have not yet been announced.
<https://gadgets360.com/laptops/news/asus-expertbook-b3-detachable-2-in-1-laptop-specifications-snapdragon-7c-gen-2-soc-ces-2022-269645#rss-gadgets-laptops>

3. : Netflix March 2022 Releases: Bridgerton Season 2, Venom: Let There Be Carnage, Outlander Season 6, and More
Bridgerton season 2, Venom: Let There Be Carnage / Venom 2, Outlander season 6, Ryan Reynolds' "The Adam Project, Eternally Confused and Eager for Love, Formula 1: Drive to Survive season 4, Taarak Mehta Kka Chhota Chashmah season 2, Lily Collins in Windfall #3" the biggest movies and TV series on Netflix in March 2022.
<https://gadgets360.com/entertainment/features/netflix-march-2022-releases-india-bridgerton-2-venom-adam-project-outlander-6-formula-1-278940#rss-gadgets-apps>

4. : Bridgerton Season 2, Jaha, Moon Knight, and More March Guide to Netflix, Disney+ Hotstar, Prime Video
Bridgerton season 2, Jaha movie, Moon Knight, The Adam Project, Rudra: The Edge of Darkness, Sullivan, Undekhi season 2, JaggaJigarthi, Eternally Confused and Eager for Love, Formula 1: Drive to Survive season 4, Halo TV series, Deep Water, WeCrashed, Star Trek: Picard season 2, The Dropout, Upload season 2, Pachinko #3" the biggest OTT movies, web series and TV shows.
<https://gadgets360.com/entertainment/features/param-jaha-movie-moon-knight-and-more-march-guide-to-netflix-hotstar-prime-video-ott-movies-web-series-2796145#rss-gadgets-internet>

5. : Bridgerton Season 2, Jaha, Moon Knight, and More March Guide to Netflix, Disney+ Hotstar, Prime Video
Bridgerton season 2, Jaha movie, Moon Knight, The Adam Project, Rudra: The Edge of Darkness, Sullivan, Undekhi season 2, JaggaJigarthi, Eternally Confused and Eager for Love, Formula 1: Drive to Survive season 4, Halo TV series, Deep Water, WeCrashed, Star Trek: Picard season 2, The Dropout, Upload season 2, Pachinko #3" the biggest OTT movies, web series and TV shows.
<https://gadgets360.com/entertainment/features/bridgerton-jaha-moon-knight-march-2022-netflix-hotstar-prime-video-ott-movies-web-series-2796145#rss-gadgets-apps>

6. : Samsung Galaxy Book 2 Pro, Galaxy Book 2 Pro 360 India Launch Tipped by Amazon, Official Website Listing
Samsung Galaxy Book 2 Pro and Samsung Galaxy Book 2 Pro 360 are making their way to India. Shortly after the company announced its new Galaxy Book 2 Series laptops at MWC 2022, the Galaxy Book 2 Pro is the company's thinnest laptop, while the Galaxy Book 2 Pro 360 is a thin 2-in-1 laptop with S Pen support. Samsung is yet to reveal when these laptops will be launched.
<https://gadgets360.com/laptops/news/samsung-book-2-pro-360-india-launch-amazon-listing-website-expected-specifications-galaxy-2813145#rss-gadgets-laptops>

Отже вивід здійснюється коректно.

Код программы

```
index.html  
<html>  
  
    <head>  
        <meta http-equiv="Content-Type" content="text/html; charset=utf-8">  
        <title>RSS Aggregator</title>  
    </head>  
  
    <body bgcolor="white">  
        <form action="/cgi-bin/test_json.py" target="content">  
            <br> Запит: <input type="text" name="query" value="" />  
            <br> Джерело: <input type="text" name="src" value="" />  
            <br><input type="submit" name="act" value="Старт"/>  
        </form>  
    </body>  
  
</html>
```

```
prog2_html.py  
#!/usr/bin/python3  
  
import sys  
import re  
import datetime  
  
t = sys.stdin.read()  
json = t.split('\n')  
t=""  
  
for i in range(len(json)):  
    t=t+" "+json[i]  
  
t=re.sub('^\s+', '', t)  
total = re.findall('"took" : (\d+), ', t)  
title = re.findall('"title" : "(.*?)" , ', t)  
text = re.findall('"textBody" : "(.*?)" , ', t)  
url = re.findall('"URL" : "(.*?)" , ', t)  
  
print("Content-type:text/html\r\n\r\n")  
print("<html>")  
print("<head>")  
print('<meta http-equiv="Content-Type" charset=utf-8" />')
```

```

print("</head>")
print("<body bgcolor=yellow>")
print("<b>Found: "+str(len(title))+ "</b><hr><ol>")

for i in range(len(title)):
    doc_ind=i+1
    print ("<li><b>"+": "+title[i]+"</b>")
    print ("<br>"+text[i])
    print ("<br><i>"+url[i]+ "</i><br /><hr>")
print ("</ol></body></html>")

```

test_json.py

```

#!/usr/bin/python3

import os
import re
import cgi
import subprocess
form = cgi.FieldStorage()
query = form.getFirst("query", "Query not defined")
src = form.getFirst("src", "Source not defined")
os.system('/usr/lib/cgi-bin/test_json.sh' %s %s' % (query, src))

```

test_json.sh

```

#!/bin/sh

curl -XGET 'http://localhost:9200/test/_search?pretty=true' -H 'Content-Type: application/json' -d '
{
    "query": {
        "bool": {
            "must": [
                {
                    "multi_match": {
                        "query": "$1",
                        "fields": ["textBody", "title"]
                    }
                },
                {
                    "filter": [
                        {
                            "match": {
                                "source": "$2"
                            }
                        }
                    ]
                }
            ]
        }
    }
}'
```

```
        }
    }
}
'
| /usr/lib/cgi-bin/prog2_html.py
```

Питання до практичної роботи

Які типи запитів формуються в **Elasticsearch**?

У **Elasticsearch** існують такі типи запитів:

- запит за терміном, що входить до неструктурованого текстового поля;
- запит відповідності;
- запиту на пошук у декількох полях;
- запит на основі фільтрованих даних.

Як працює мова запитів **DSL (Domain Specific Language)** в системі **Elasticsearch**?

Найкращий опис роботи з **DSL**:

<https://www.elastic.co/guide/en/elasticsearch/reference/current/query-dsl.html>

Як сортуються результати пошуку в **Elasticsearch** згідно з релевантністю кожного документа до запиту?

Коли ми виконуємо запит на відповідність, очікуються наступні дії:

1. Пошук термінів по всіх документах у межах певного поля (наприклад, **"textBody"**).
2. Знаходження найкращих відповідностей, відсортованих за оцінкою в порядку зменшення частоти народження.

Коли кінцевий користувач шукає будь-які терміни, іноді потрібно виконувати запити по двох полях **"textBody"** (тіло документа) та **"title"** (назва). Це можливо за допомогою запиту на пошук у декількох полях.

Яка роль **JSON REST API** в **Elasticsearch**?

Elasticsearch надає широкий спектр **REST API** для інтеграції, запитів і керування даними.

Ми можемо робити багато речей, використовуючи **JSON API REST Elasticsearch**, наприклад:

- ❖ Перевірте стан кластера, вузла та індексу, стан, статистику тощо.
- ❖ Адмініструвати наші дані та метадані кластера, вузла, індексу.
- ❖ Виконуйте операції **CRUD** (створювати, читати, оновлювати та видаляти) та здійснювати пошук щодо наших індексів.
- ❖ Виконуйте розширені операції пошуку, такі як сторінка, сортування, фільтрація, сценарії, агрегації тощо.