Лабораторна робота № 8(3.1.19)

Продвинуті пошуки

Хід роботи:

Частина 1: Google Advanced Searches (Dorking)

Таблиця операторів Google Advanced Search

ОПЕРАТОР	ОПИС
allintext:	Шукає сторінки, де всі ключові слова є в тексті
filetype:	Фільтрує за типом файлу (.pdf, .ppt, .doc)
intitle:	Шукає слова в заголовку сторінки
inurl:	Шукає слова в URL адресі
site:	Обмежує пошук конкретним доменом

Крок 1: Дослідження Google dorking

Базові команди:

Звичайний пошук (без операторів)

ethical hacker

Пошук на конкретному сайті

ethical hacker site:pearson.com

Пошук конкретного типу файлів на сайті

ethical hacker site:pearson.com filetype:pdf

Пошук за словом у заголовку

ethical hacker intitle:certification

Пошук за словом в URL

ethical hacker inurl:free

Пошук всіх слів у тексті сторінки

allintext:free ethical hacker practice test questions

Пошук точної фрази (в лапках)

"**ethical hacker certification**"

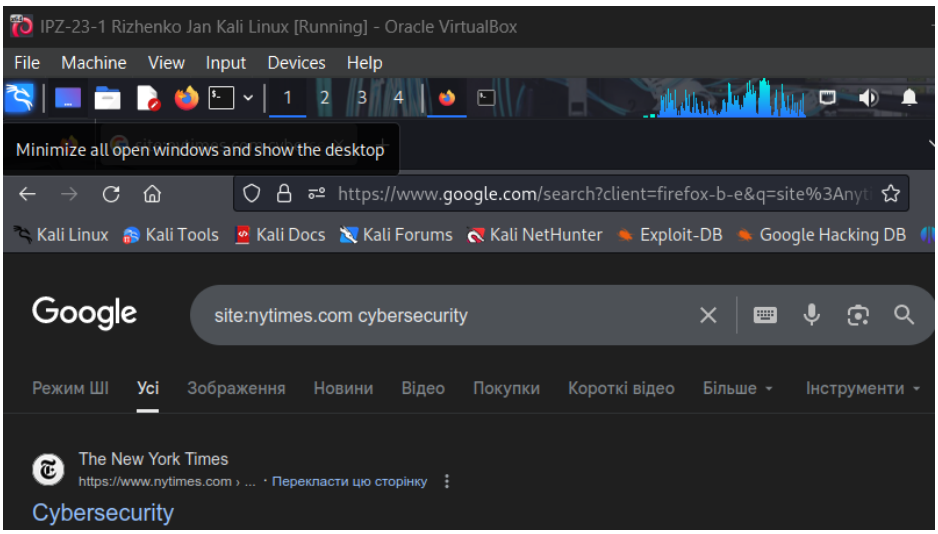


Рис. 1. Приклади пошукових запитів Google dorking.

Питання: Що спільного у всіх результатах пошуку ethical hacker site:pearson.com?
Відповідь:
Всі знайдені сторінки розташовані на сайті pearson.com і містять слова "ethical" та "hacker". Pearson є видавництвом, тому результати включають книги, курси та навчальні матеріали про етичний хакінг.
Питання: Який тип файлу відкривається кожним з результатів ethical hacker site:pearson.com filetype:pdf?
Відповідь:
Відкриваються PDF файли - це можуть бути фрагменти книг, навчальні посібники, презентації або технічна документація про етичний хакінг.

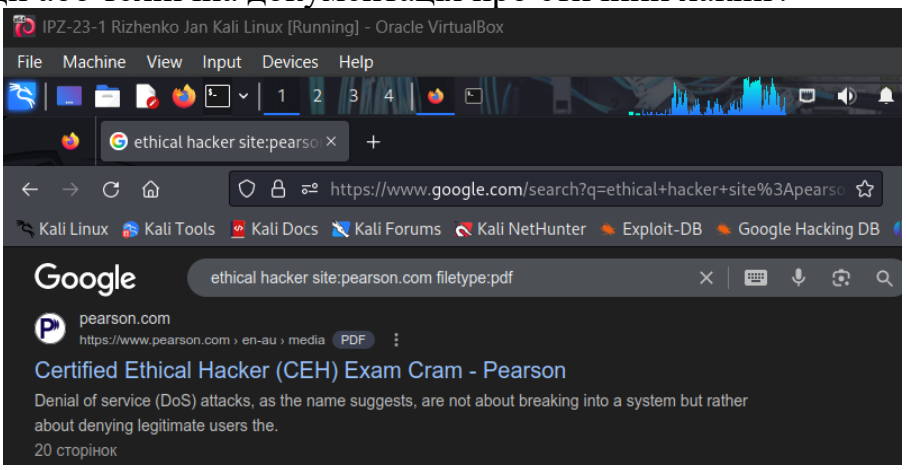


Рис. 2. Результат пошуку з різними операторами.

Крок 2: Використання форми Google Advanced Search
Доступ до форми:
В Google пошуку ввести:
advanced search
Або прямий URL:
https://www.google.com/advanced_search

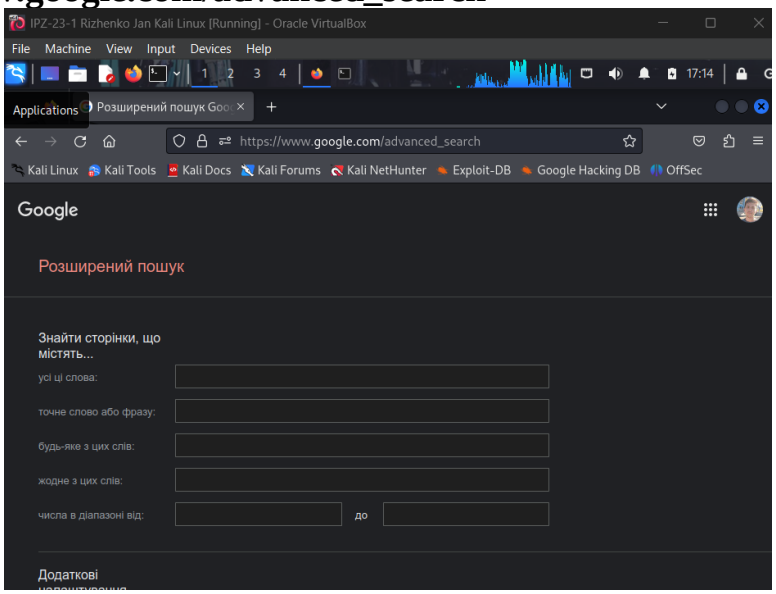


Рис. 3. Інтерфейс Google Advanced Search.

Еквіваленти між формою та операторами:

ФОРМА ADVANCED SEARCH	ОПЕРАТОР
all these words	(звичайний пошук)
this exact word or phrase	"речення в лапках"
any of these words	OR
none of these words	- (мінус)
site or domain	site:
file type	filetype:
terms appearing	allintext:, intitle:, inurl:

Крок 3: Пасивна розвідка з advanced search

Основні категорії пошуку:

1. Адміністративні сторінки та панелі входу:

site:examplecompany.com inurl:admin

site:examplecompany.com inurl:cpanel

site:examplecompany.com intitle:login

site:examplecompany.com intitle:"sign in"

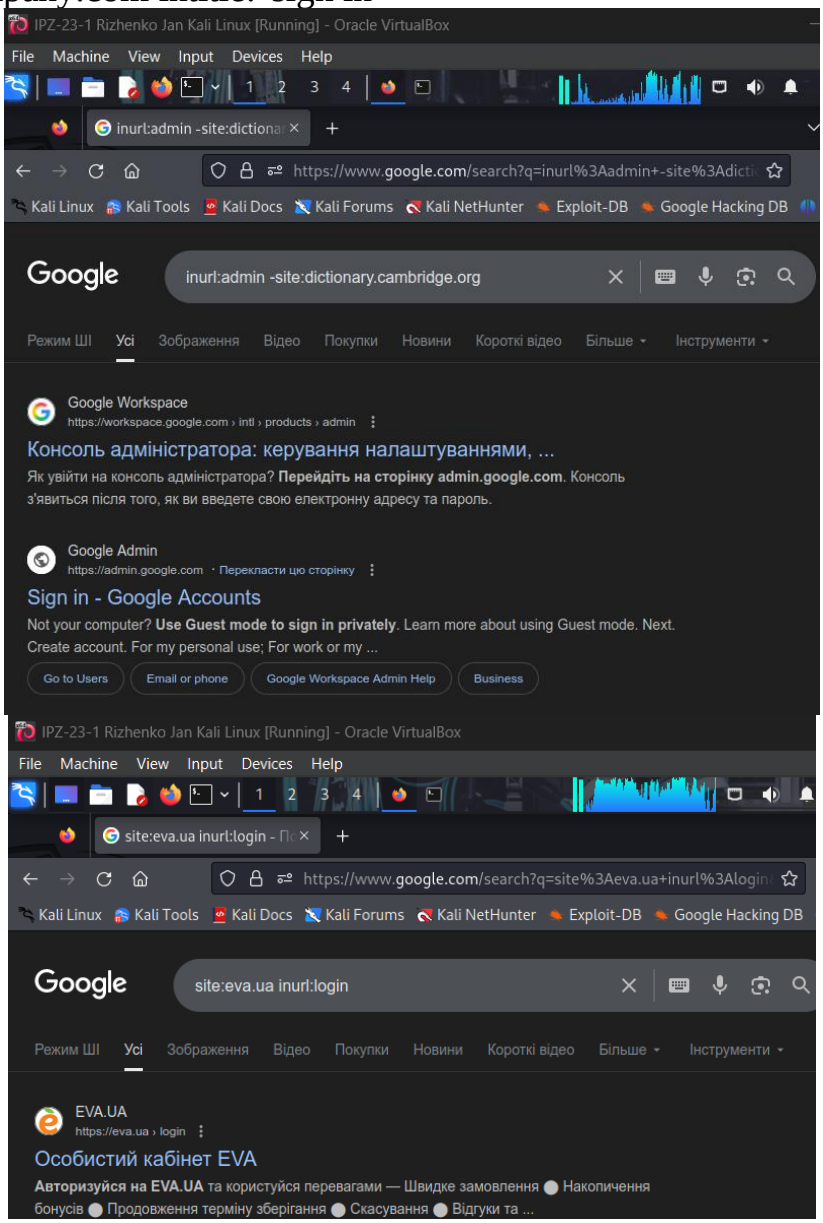


Рис. 4-5. Знайдені адмін-панелі та сторінки входу.

2. Документи та файли:
site:examplecompany.com filetype:pdf
site:examplecompany.com filetype:doc
site:examplecompany.com filetype:xlsx
site:examplecompany.com intext:"confidential" filetype:pdf
site:examplecompany.com intext:password filetype:xls

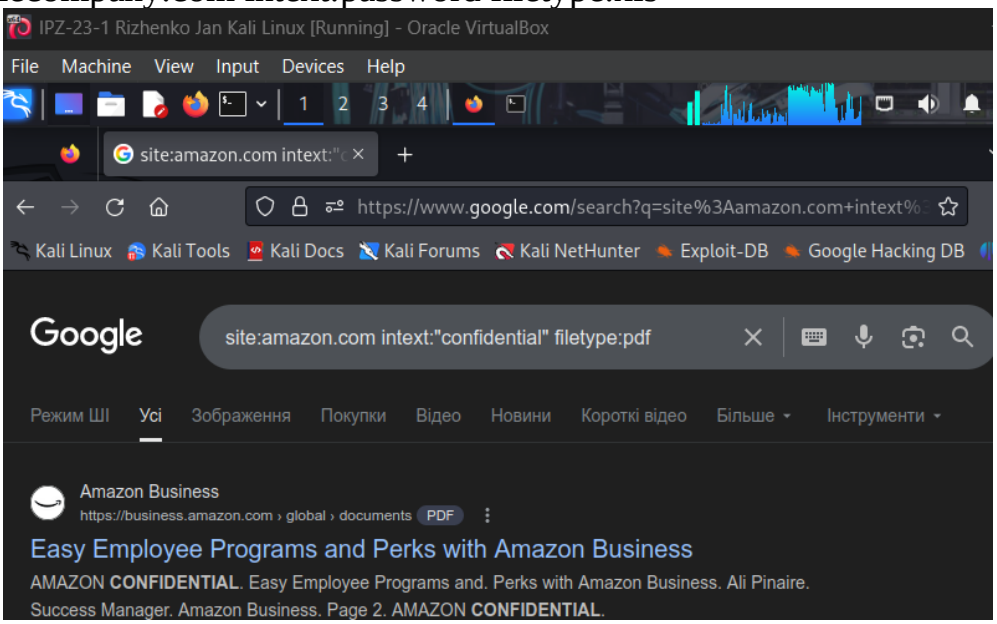


Рис. 6. Документи з конфіденційною інформацією

3. Конфігураційні та backup файли:
site:examplecompany.com filetype:xml inurl:config
site:examplecompany.com filetype:conf
site:examplecompany.com filetype:env
site:examplecompany.com filetype:bak
site:examplecompany.com ext:sql

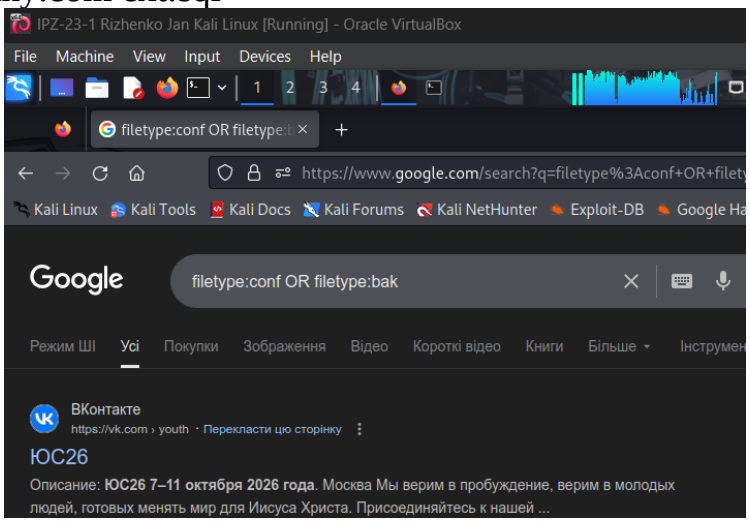


Рис. 7. Конфігураційні та резервні файли.

4. Пошук інформації про співробітників (LinkedIn та соціальні мережі):
site:linkedin.com intitle:"example company"
site:linkedin.com "example company" "CEO"
site:linkedin.com "example company" "IT Manager"
site:facebook.com "example company"

site:github.com "example company"

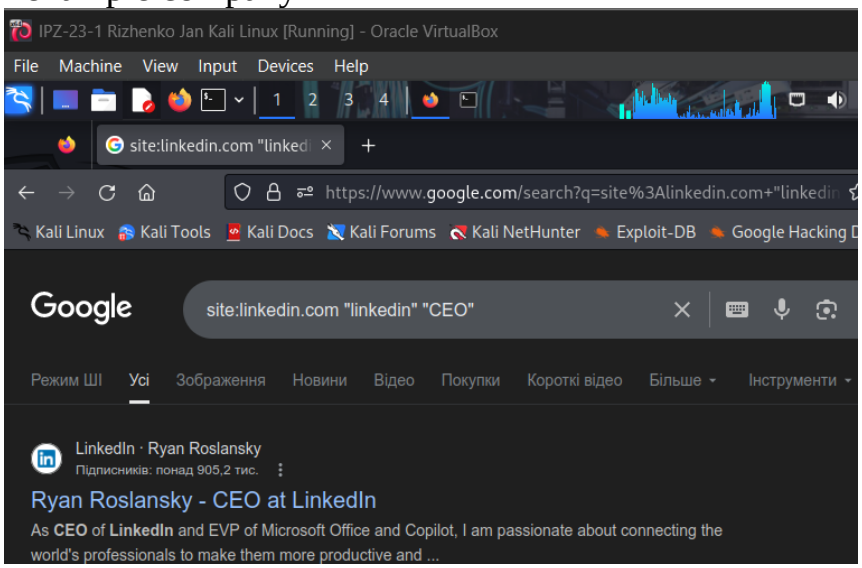


Рис. 8. Інформація про співробітників на LinkedIn.

Питання: Яку інформацію може отримати зловмисник через пошук на LinkedIn?

Відповідь:

З LinkedIn можна дізнатися:

- Імена, посади та email адреси співробітників
- Структуру відділів та ієрархію компанії
- Використовувані технології (з навичок працівників)
- Зв'язки між співробітниками для соціальної інженерії
- Нових працівників (потенційно менш обізнаних у безпеці)
- Номери телефонів (інколи)
- Дати працевлаштування (скільки років працюють)
- Проекти та ініціативи
- Технологічні трансформації
- Зв'язки між співробітниками

Приклад використання:

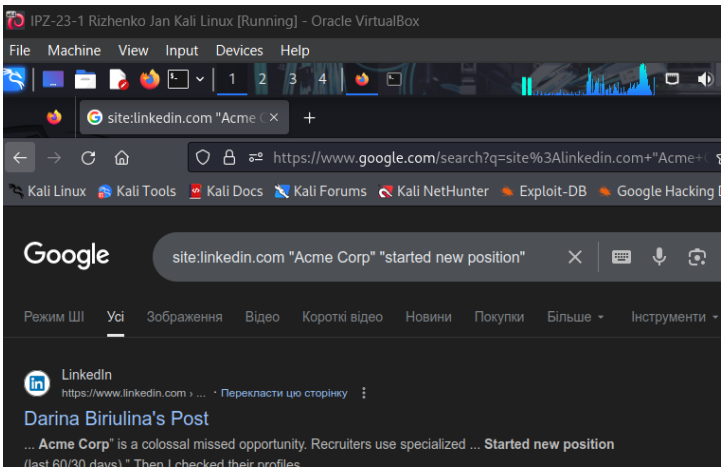


Рис. 9. Інформація про новий проект Acme Corp.

Частина 2: The Google Hacking Database (GHDB)

Крок 1: Огляд GHDBc

https://www.exploit-db.com/google-hacking-database

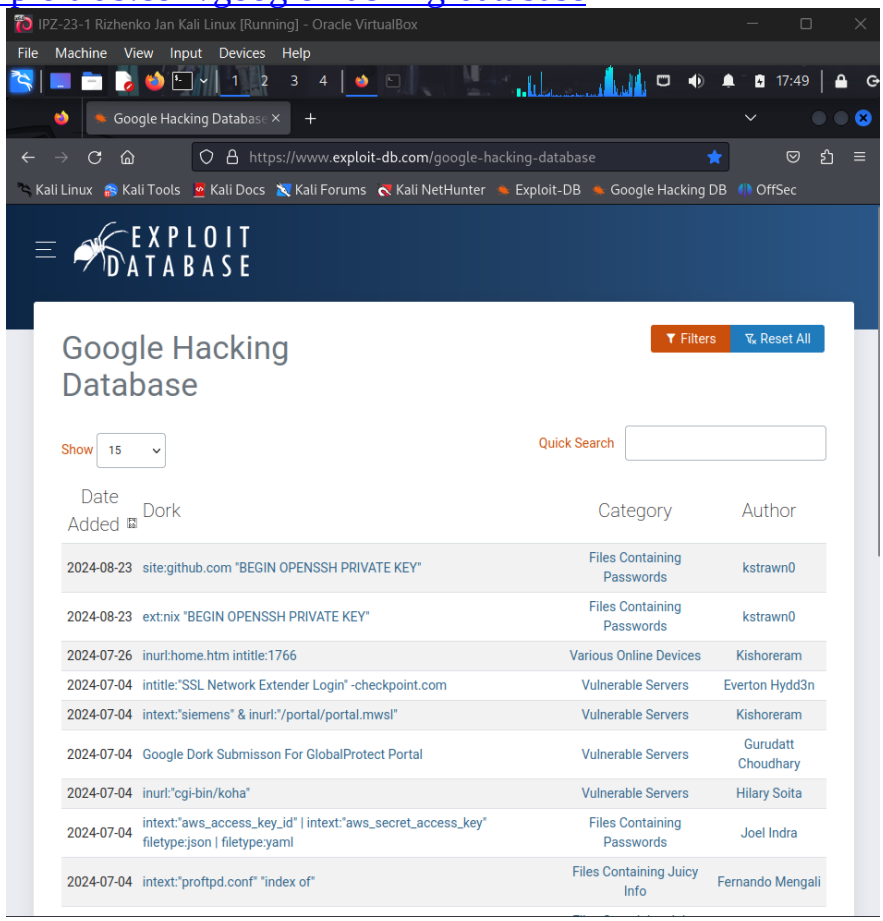


Рис. 10. Головна сторінка GHDB з категоріями.

Основні категорії:

- Footholds - точки входу
- Files Containing Usernames/Passwords - файли з обліковими даними
- Sensitive Directories - приховані директорії
- Vulnerable Files/Servers - вразливі файли та сервери
- Error Messages - повідомлення про помилки
- Login Portals - сторінки входу
- Network/Vulnerability Data - мережеві дані

Крок 2: Використання Quick Search

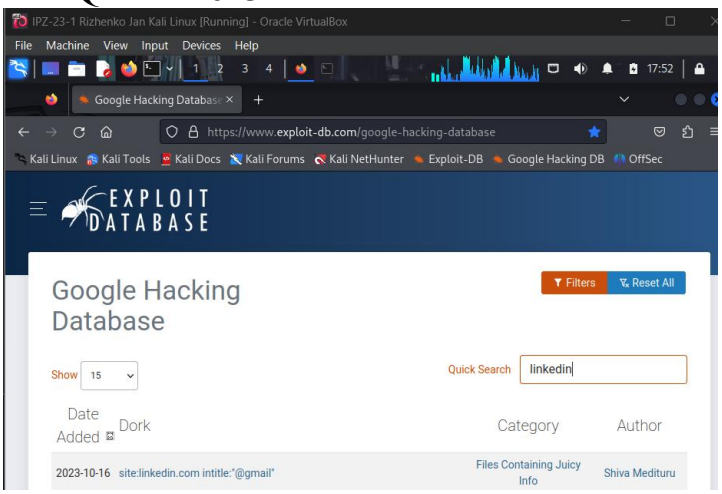


Рис. 11. Інтерфейс Quick Search та детальна інформація про Dork.

Питання: Яка інформація надається про Dorks?
Відповідь:
Про кожен Dork в GHDB надається наступна інформація:

- Сам пошуковий запит (можна скопіювати)
- Категорія, дата додавання та автор
- Опис того, що знаходить цей запит
- Кнопка для запуску пошуку в Google
- Іноді приклади та скріншоти

Крок 3: Приклад Dork'а
allinurl:tsweb/default.htm

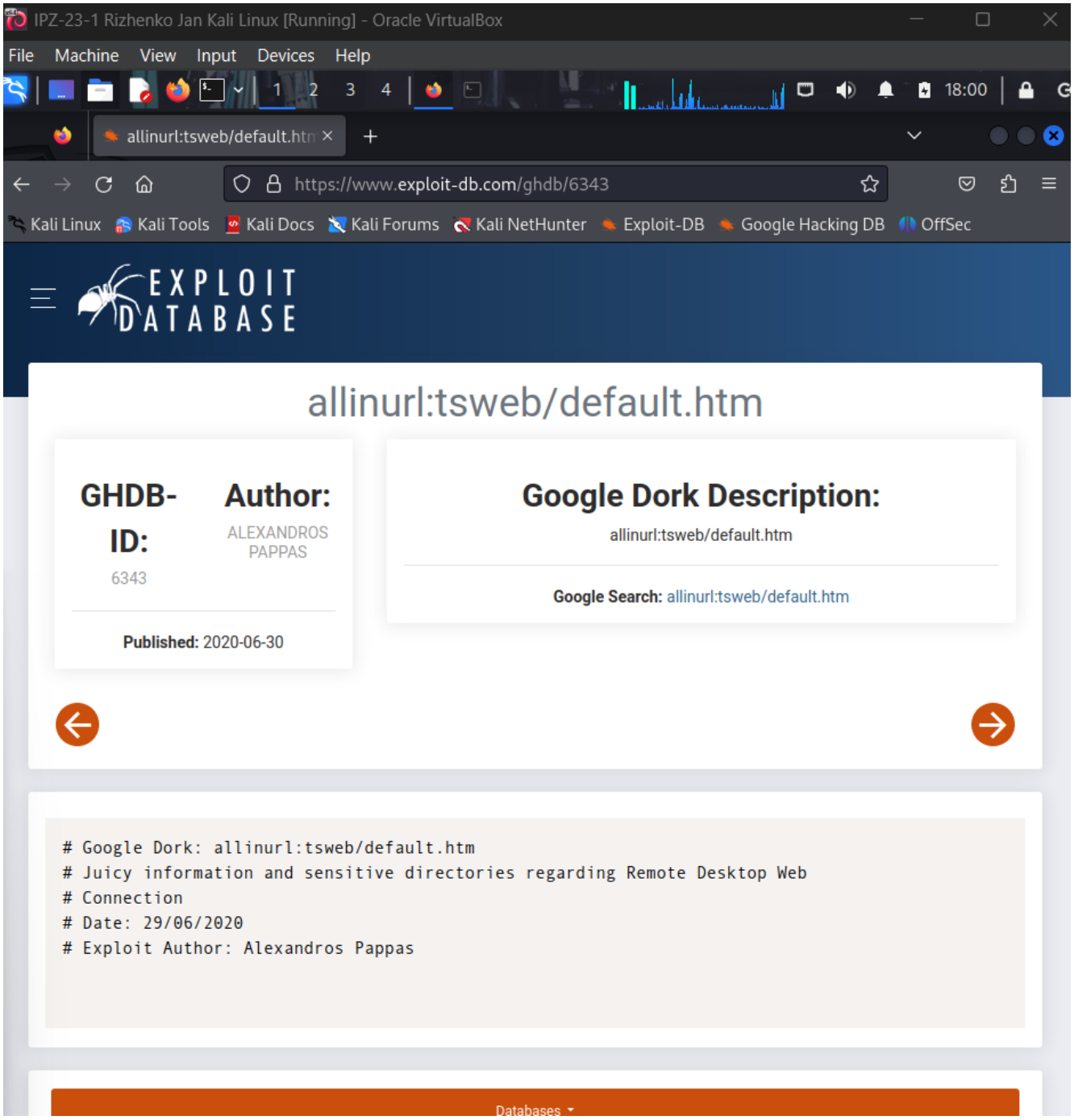


Рис. 12. Результати пошуку Terminal Services порталів.

Питання: Що повертає цей Dork?
Відповідь:
Цей запит знаходить сторінки входу Terminal Services Web Access - інтерфейс для віддаленого підключення до Windows серверів. Можна дізнатися версії Windows Server (часто застарілі), версію IIS, назви доменів та IP адреси. Це небезпечно, оскільки старі версії мають відомі вразливості, можна проводити brute-force атаки, і RDP відкритий в інтернет.
Крок 4: Комбінування категорій з пошуком
Корисні dorks за категоріями:

1. Файли з паролями:

filetype:env "DB_PASSWORD"
filetype:yml "password:"
filetype:config "dbpassword"

2. Повідомлення про помилки:

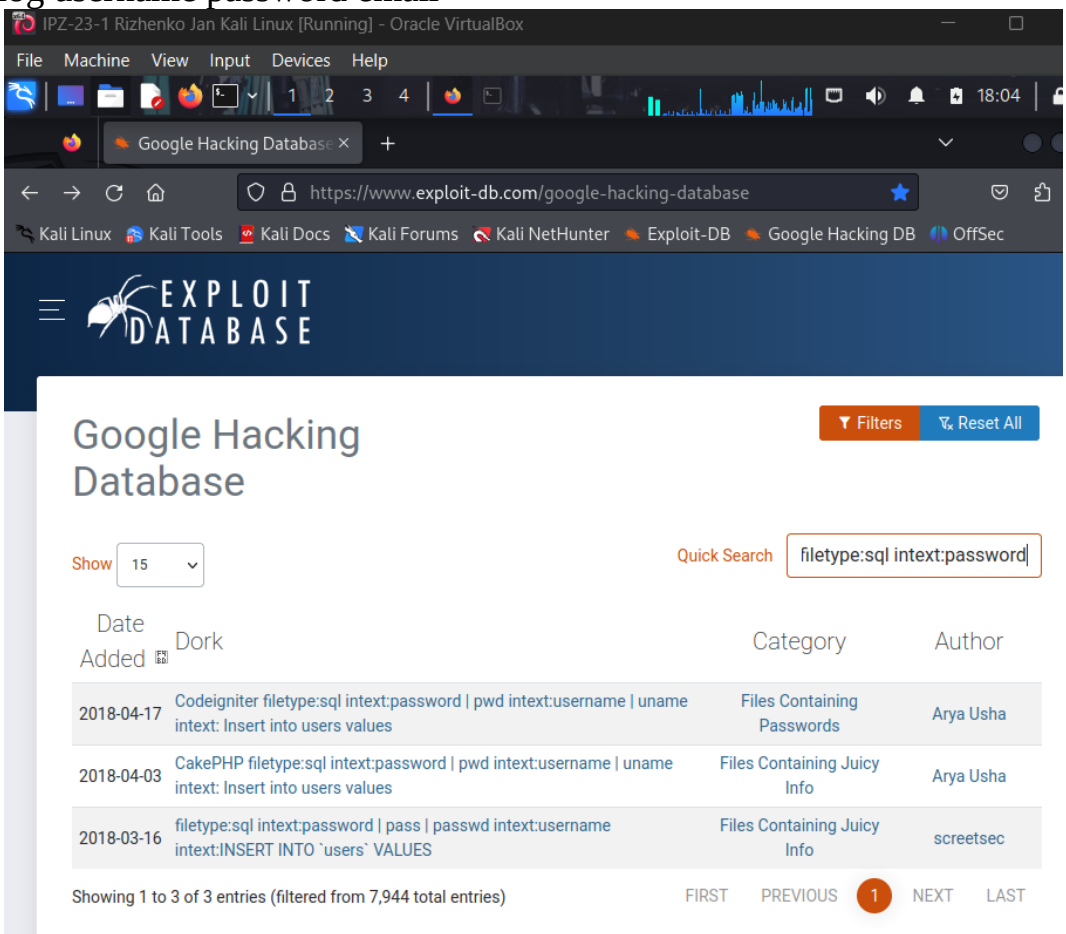
intext:"sql syntax near"
intext:"Warning: mysql_connect()"

3. Портали входу:

intitle:"login" inurl:admin
inurl:admin intitle:login

4. Вразливі файли:

filetype:sql intext:password
filetype:log username password email



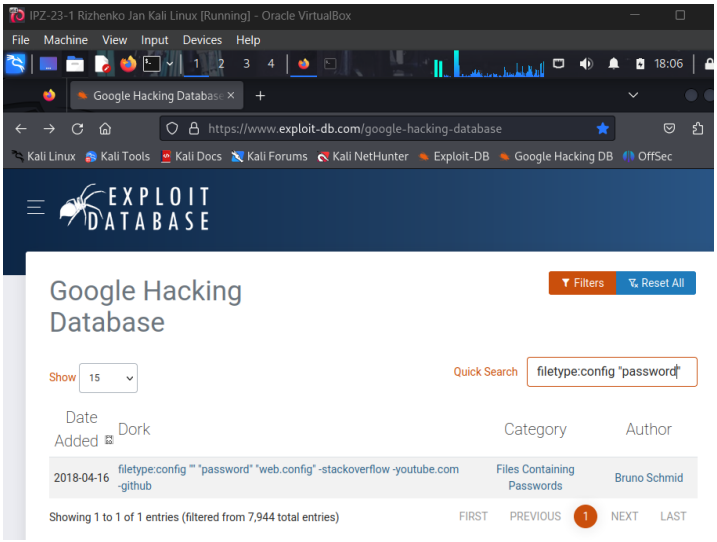


Рис. 13-14. (Приклади знахідок різних категорій dorks)

5. Backup файли та exposed директорії:

site:example.com ext:sql | ext:backup | ext:bak
intitle:"index of" "backup"
intitle:"Index of /" +.htpasswd
intitle:"Index of" .ssh

6. API ключі:

filetype:env "API_KEY"
site:github.com "AWS_ACCESS_KEY_ID"

Частина 3: The Wayback Machine

Крок 1: Огляд Wayback Machine

URL: https://web.archive.org

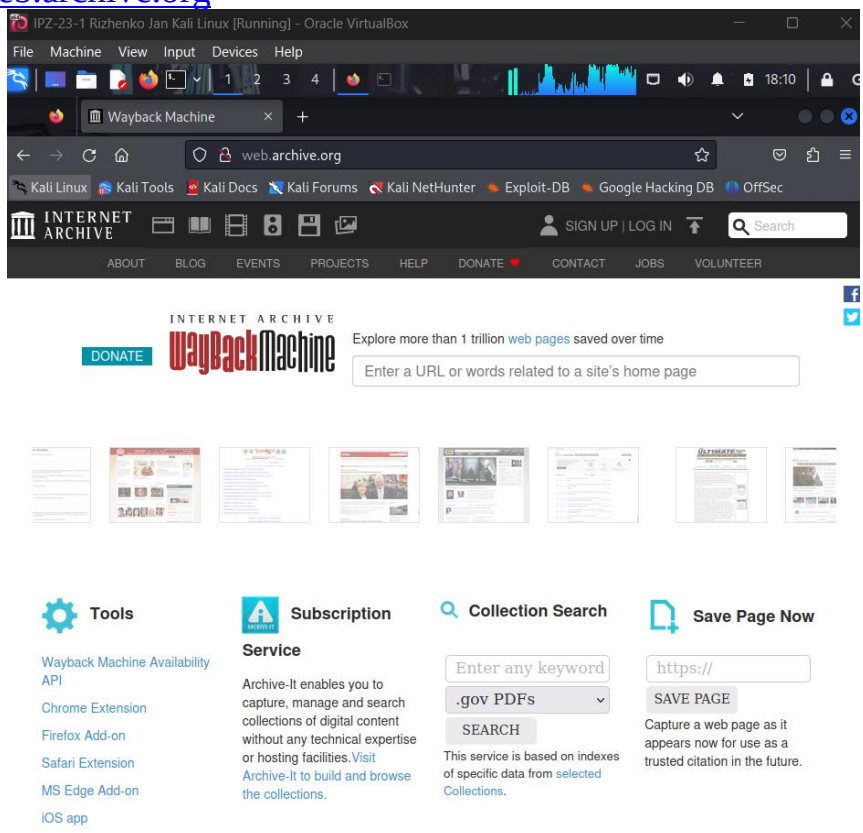


Рис. 15. Головна сторінка Wayback Machine.

Основні функції:
1. Пошук архівованих копій сторінок
2. Перегляд snapshots з різних часових періодів
3. Порівняння змін між версіями
4. Пошук конкретних файлів та URL

Крок 2: *Вкладка Calendar*

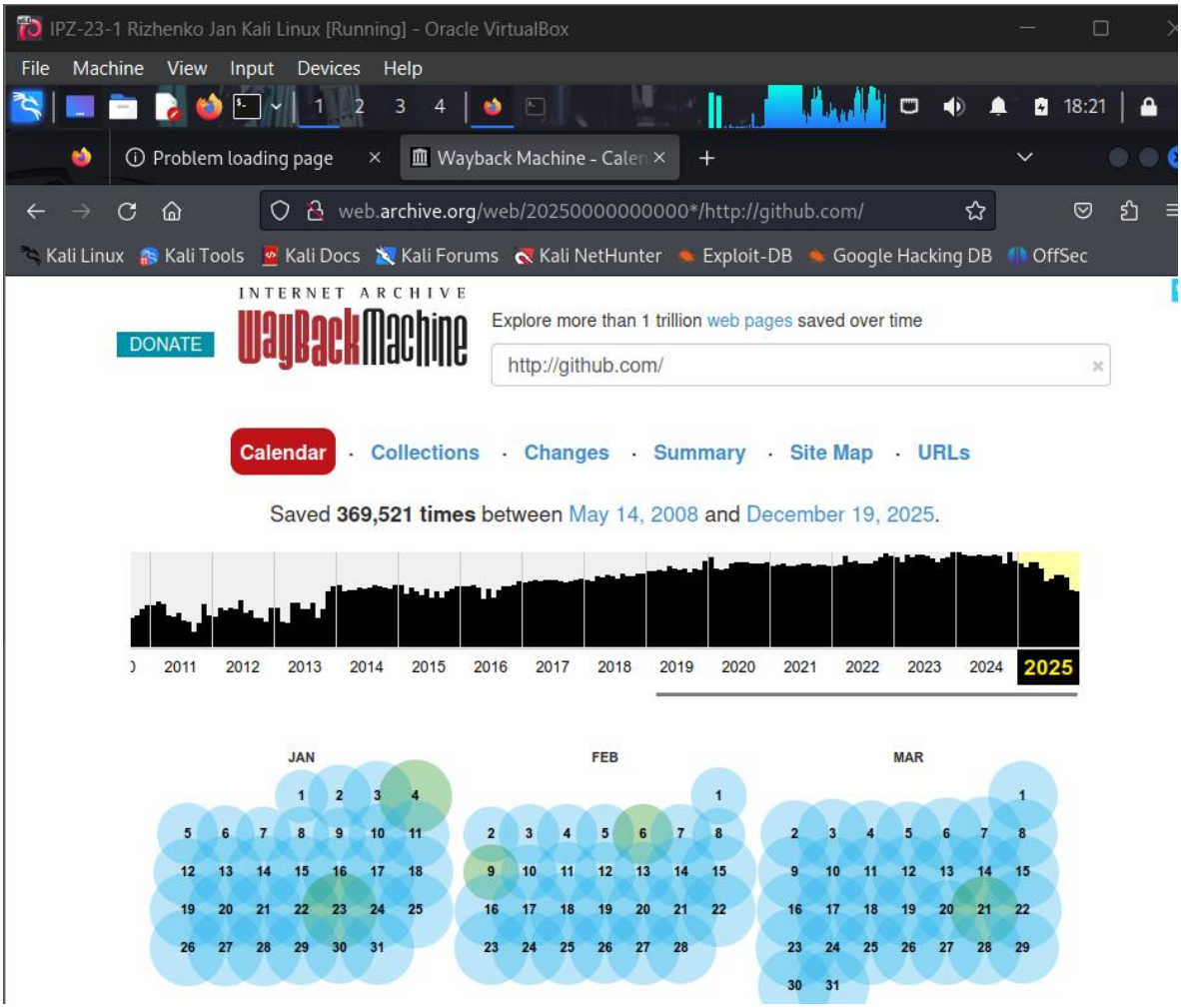


Рис. 16. Календар архівних копій з timeline графіком.

Календар показує:

- Сині кружечки - є збережені копії
- Розмір кружечка - кількість копій
- Графік частоти архівації

Питання: Як може бути вигідним для хакера збір інформації з архівованого сайту?

Відповідь:

Архівовані версії сайтів дають доступ до:

1. Видаленої інформації:

- Старі контакти співробітників
- API endpoints, які можуть ще працювати
- Сторінки з вразливостями
- Приховані директорії (/test/, /dev/)

2. Конфіденційні дані:

- Паролі в коментарях коду
- API ключі в JavaScript
- Рядки підключення до баз даних

3. Технічну інформацію:

- Версії ПЗ з відомими вразливостями
- Використовувані фреймворки
- Інформація про сервери

4. Для соціальної інженерії:

- Історію компанії
- Інформацію про співробітників
- Проекти та партнерства

Наприклад:

В архіві 2018 року знайдено API ключ у коді:

var apiKey = "sk_live_abc123xyz789";

Потім перевіряється, чи він ще працює.

Крок 3-6: Огляд інших вкладок

- **Collections** - різні джерела сканування (Internet Archive Bot, Alexa Crawls)

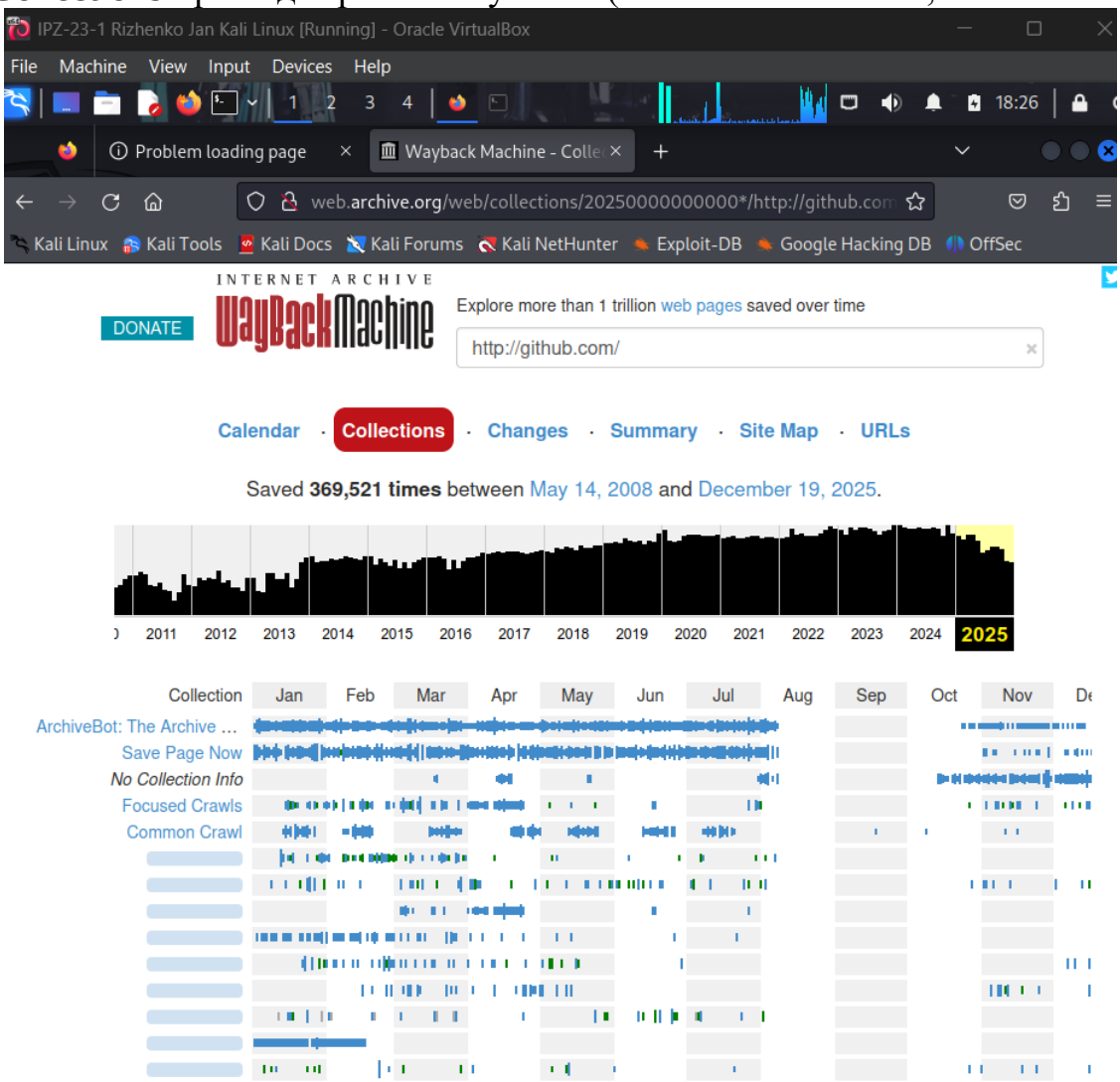


Рис. 17. Вкладка Collections.

- **Changes** - візуалізація змін між копіями, можна порівнювати версії

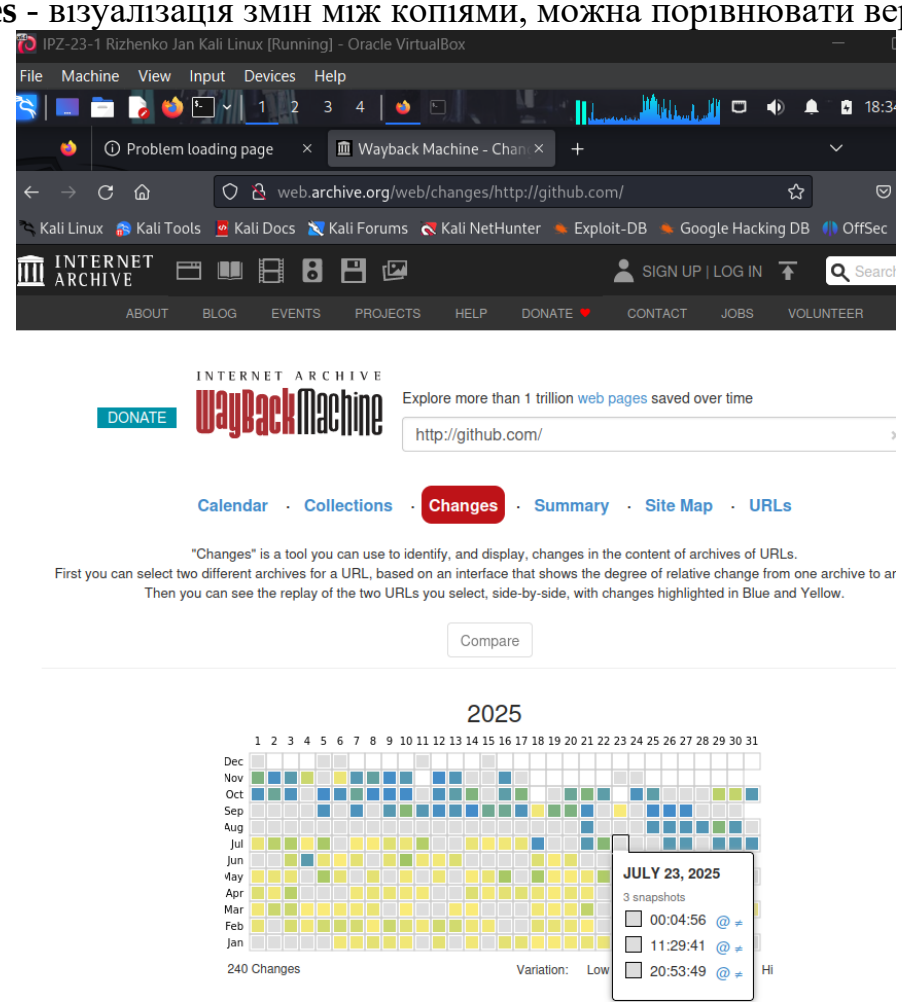


Рис. 18. Вкладка Changes.

- **Summary** - статистика типів файлів (HTML, PDF, зображення, відео)

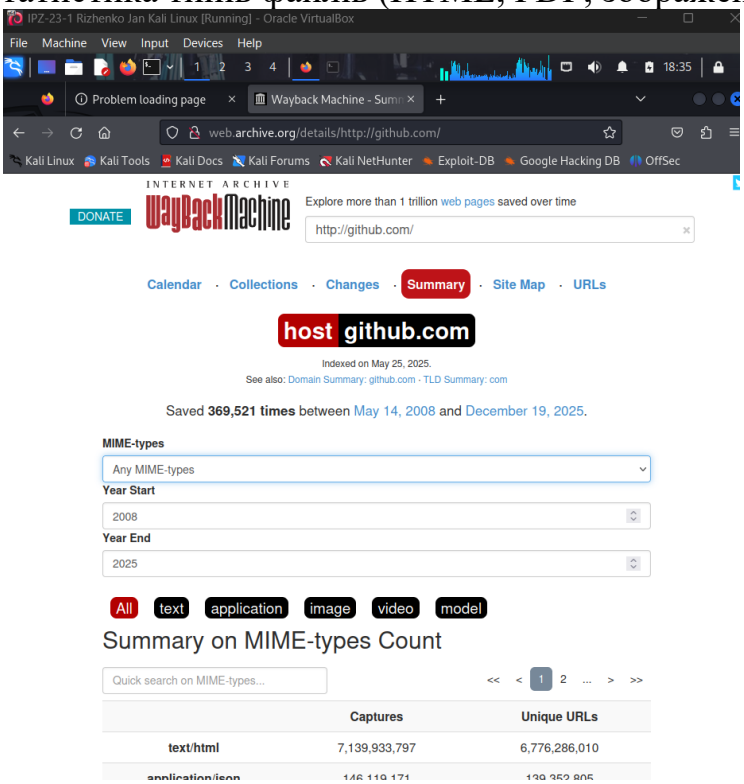


Рис. 19. Вкладка Summary.

- **Site Map** - візуальна карта структури сайту, показує глибину та складність

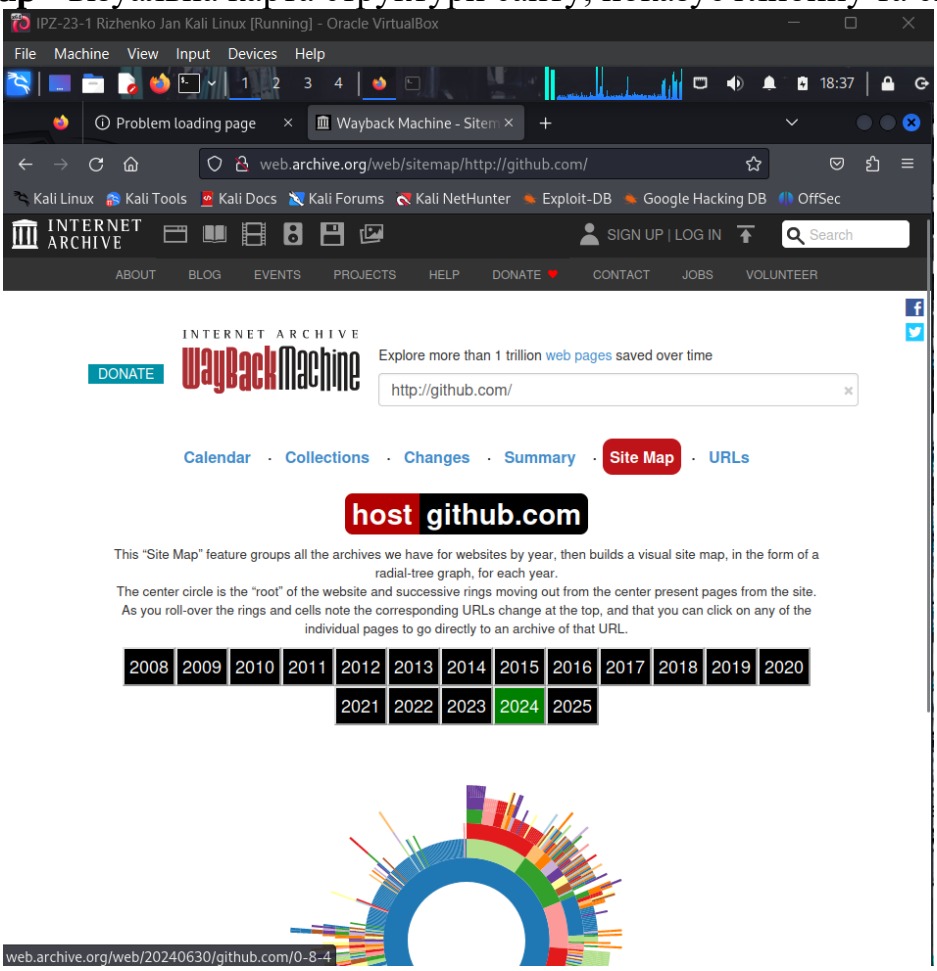


Рис. 19. Вкладка Site Map.

Крок 7: Вкладка URLs та пошук цікавих файлів

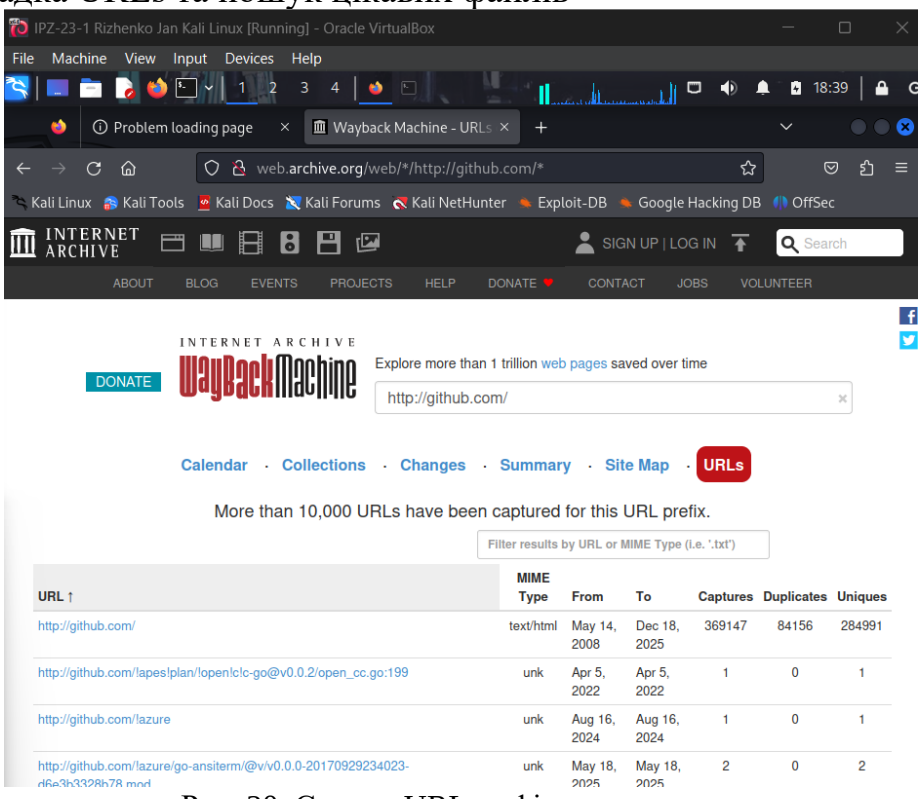


Рис. 20. Список URL та фільтри пошуку.

Основні фільтри для пошуку:

1. Backup та архіви:
   - *.bak, *.backup, *.old
   - *.zip, *.rar, *.tar
2. Конфігурації:
   - *.config, *.conf, *.env, *.ini
3. Бази даних та документи:
   - *.sql, *.db
   - *.csv, *.xls, *.doc, *.pdf
4. Адмін-панелі та API:
   - /admin/, /cpanel/, /manage/
   - /api/, /v1/, /rest/
5. Тестові середовища:
   - /dev/, /test/, /staging/

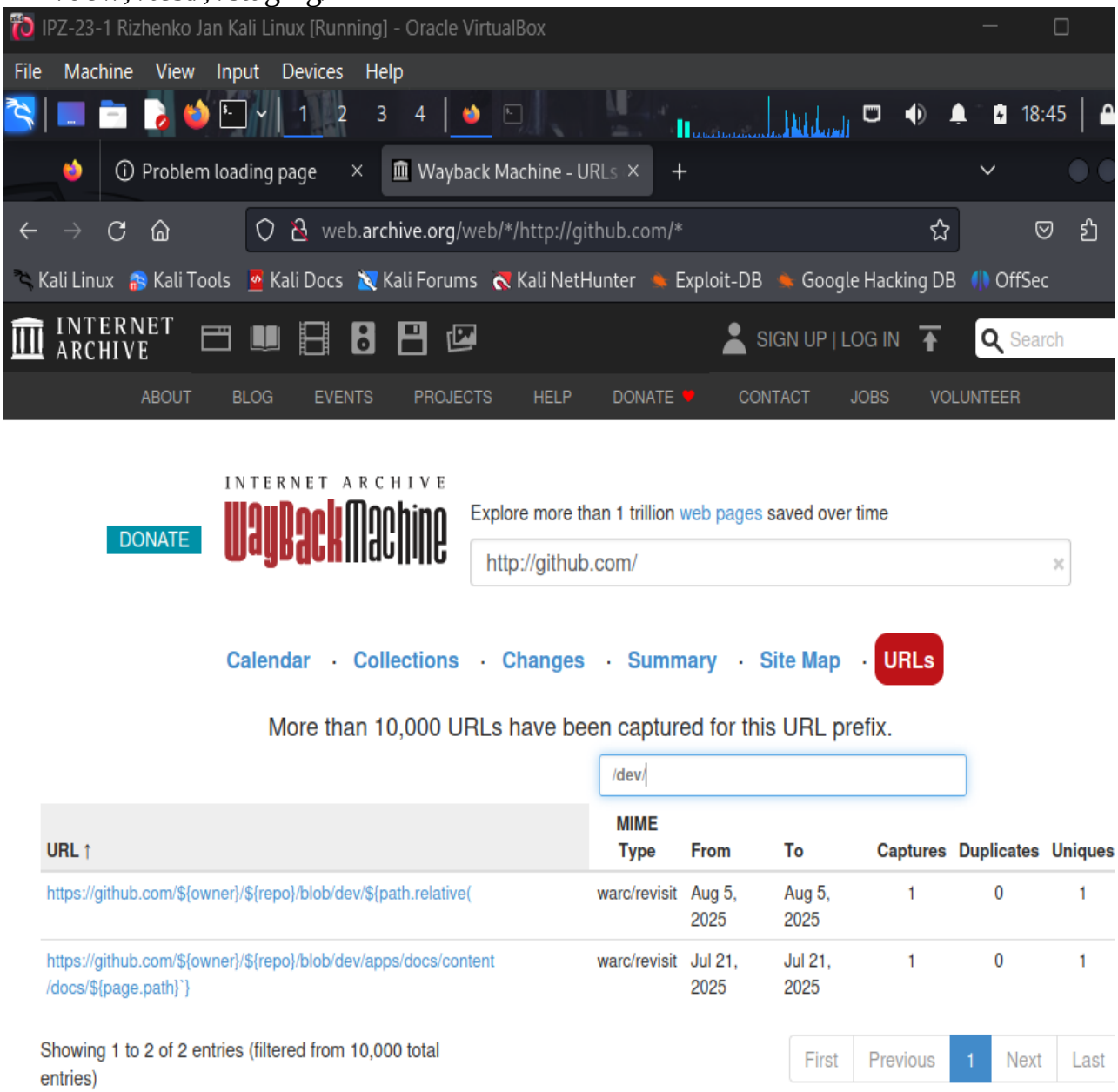


Рис. 21. Приклад пошуку URLs.

Автоматизація через CDX API:

- Отримати всі URL

curl "http://web.archive.org/cdx/search/cdx?url=example.com/*&output=json"

- Фільтрувати PDF

curl "http://web.archive.org/cdx/search/cdx?url=example.com/*&filter=mimetype:application/pdf"

- За датами

curl "http://web.archive.org/cdx/search/cdx?url=example.com/*&from=2015&to=2020"

Інструменти для автоматизації:

- Pagodo - автоматичний Google dorking
- GooDork - швидкий dorking
- waybackurls - витягування URL з Wayback
- Subfinder - пошук субдоменів

Reflection Question (Питання для рефлексії)
Питання: Чому пасивна розвідка настільки важлива для ефективного хакінгу та тестування на проникнення?
Відповідь:
Пасивна розвідка є критично важливою для ефективного хакінгу та пентестингу з наступних причин:

1. Непомітність:
   - Не залишає слідів у логах
   - Не викликає спрацювання IDS/IPS
   - Повністю легальна
2. Ефективність:
   - 70-80% потрібної інформації
   - Швидка автоматизація
   - Виявлення очевидних проблем
3. Планування атаки:
   - Визначає цілі для сканування
   - Зменшує "шумні" дії
   - Дає повну картину поверхні атаки
4. Виявлення слабких місць:
   - Забуті сервери та субдомени
   - Системи поза контролем IT
   - Витоки через треті сторони
5. Соціальна інженерія:
   - Інформація про співробітників
   - Структура організації
   - Використовувані технології

Робочий проце(зазвичай):

Етап 1: Пасивна розвідка (1-3 дні)
├── Google Dorking
├── LinkedIn/соціальні мережі
├── Wayback Machine
└── Перевірка витоків
Етап 2: Напів-пасивна (1-2 дні)
├── WHOIS, DNS
└── Shodan/Censys
Етап 3: Активна (цілеспрямована)
├── Сканування портів
└── Тестування додатків

Статистика ефективності:

Згідно з досліджень:

- **70-80%** успішних атак починаються з пасивної розвідки
- **60%** вразливостей можуть бути виявлені без active scanning
- **90%** phishing campaigns використовують інформацію з OSINT

Час: Пасивна розвідка = 20% часу, але надає 80% корисної інформації

Приклади успішних атак через пасивну розвідку(теоретично):

- Приклад 1: Google Dork знайшов .env файл з паролями до БД
- Приклад 2: Certificate Transparency показав забутий субдомен dev.company.com
- Приклад 3: LinkedIn - цільовий фішинг нового співробітника IT
- Приклад 4: GitHub зберіг репозиторій з валідним API ключем

Правило пентестингу: "Never rush into active reconnaissance without exhaustive passive reconnaissance first."

Висновок

У ході виконання лабораторної роботи було детально досліджено методи пасивної розвідки через Google Advanced Search (Google Dorking), Google Hacking Database (GHDB) та Wayback Machine. Google Dorking продемонстрував потужні можливості пошуку конфіденційної інформації, яка ненавмисно стала публічною, через використання спеціалізованих операторів (site:, filetype:, intitle:, inurl:, allintext:). GHDB надав структуризовану базу готових dorks для виявлення вразливостей, exposed credentials, admin panels та sensitive documents. Wayback Machine показав цінність історичних архівів сайтів для знаходження видаленої інформації, старих endpoints, forgotten subdomains та технічних деталей. Практична робота підкреслила критичну важливість пасивної розвідки як фундаменту для будь-якого пентесту - вона дозволяє зібрати до 80% необхідної інформації без прямої взаємодії з цільовою системою, залишаючись невидимою для систем моніторингу та захисту. Організації повинні регулярно проводити self-dorking та перевіряти свої архівні дані, щоб виявити та усунути потенційні витоки інформації до того, як їх знайдуть зловмисники.