



МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
КАФЕДРА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

**Теорія та методи соціальної
інженерії в кібербезпеці
Лабораторний практикум №3**

**Налаштування роботи веб-ресурсу
для упередження індексації деяких розділів
сайту пошуковими сервісами**

Перевірив:
Кіфорчук К. О.

Виконав:
студент I курсу
групи ФБ-41мп
Сахній Н. Р.

Київ 2025

Мета роботи: Ознайомитись із засобами, які мають використовувати веб-розробники для уникнення непередбачених ефектів індексації, і, як наслідок, доступності даних сайту широкому колу осіб (соцінженерам).

Завдання до виконання:

1. Сформувати сайт з декількох сторінок, заповнити інформацією,

а) яка буде використовуватись при пошуку сторінки

і

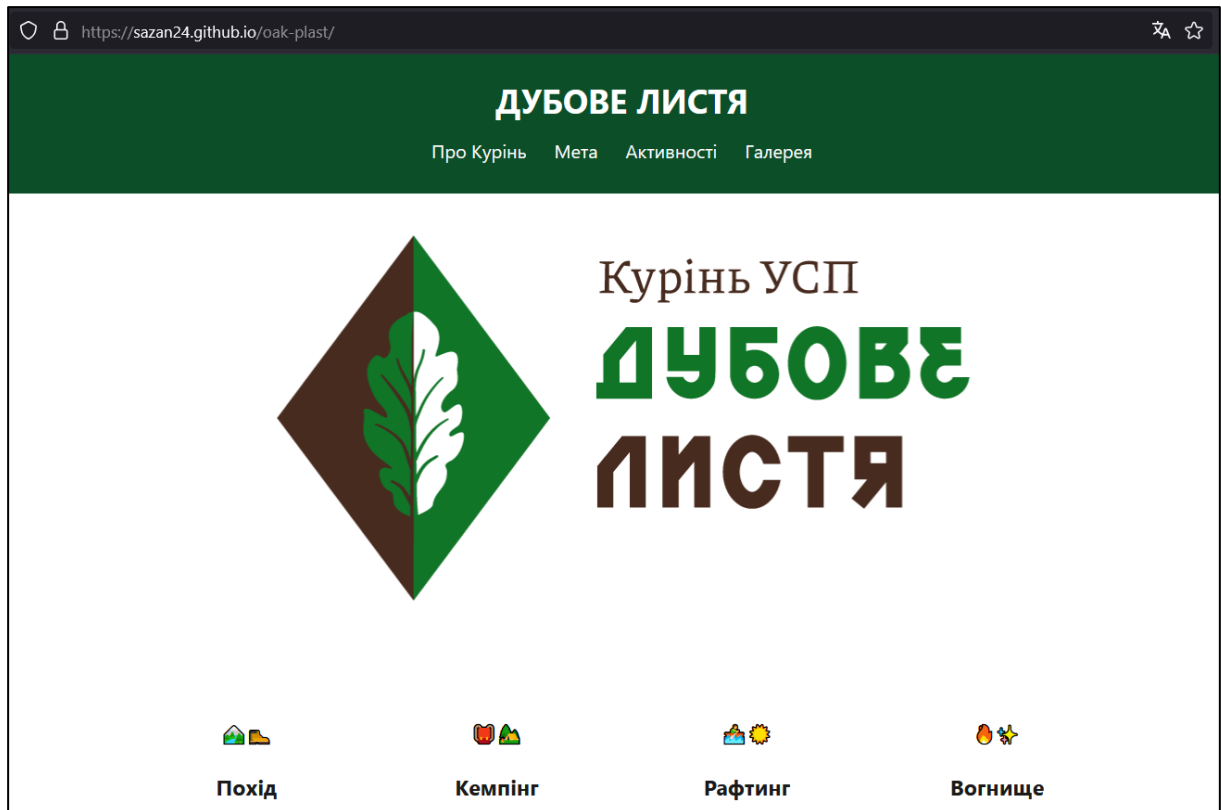
б) яка не передбачена для індексування пошуковиками.

Розмістити сторінку (* на безкоштовному хостингу).

main	1 Branch	0 Tags	Go to file	Add file	Code
sazan24 Update hidden.html ✓ ed01b30 · 3 minutes ago 51 Commits					
skob	Update hidden.html			3 minutes ago	
dl-back.png	Add files via upload			50 minutes ago	
dl-logo.jpg	Add files via upload			2 hours ago	
index.html	Update index.html			22 minutes ago	
kali-trail.png	Add bg via upload			2 hours ago	
robots.txt	Update robots.txt			9 minutes ago	
styles.css	Update styles.css			24 minutes ago	

↑↑ Site Map ++ Meta Data ↓↓

```
1 <!DOCTYPE html>
2 <html lang="uk">
3 <head>
4   <meta charset="UTF-8">
5   <meta name="viewport" content="width=device-width, initial-scale=1.0">
6   <meta name="description" content="Офіційна сторінка куреня Дубове Листя (ДЛ) – інформація про мету, активності, галерею та контакти">
7   <meta name="keywords" content="Дубове Листя, курінь, пласт, УСП, Україна, кемпінг, похід, рафтинг, пластуни">
8   <meta name="robots" content="index, follow">
9   <title>Дубове Листя</title>
10  <link rel="stylesheet" href="styles.css">
11 </head>
```



2. Використати інструменти Google для індексування веб-сайту, зокрема Google Search Console. Перед цим застосувати методи запобігання індексації та роботизованого доступу до деяких елементів сайту із інформацією, яка не повинна бути проіндексована (виконавши конфігурування «.htaccess» і «robots.txt», реалізувати RECAPTCHA).

```
oak-plast / robots.txt

sazan24 Update robots.txt ✓ 992fe89 · 13 minutes ago History

Code Blame 2 lines (2 loc) · 31 Bytes Code 55% faster with GitHub Copilot Raw Copy Download Edit History

1 User-agent: *
2 Disallow: /skob/
```

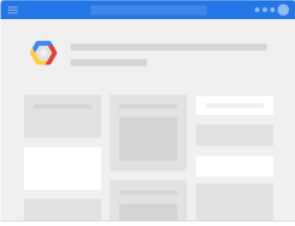
↑↑ «robots.txt» ++ «.htaccess» ↓↓

```
oak-plast / skob / .htaccess

sazan24 Update .htaccess 60bdefa · now History

Code Blame 3 lines (3 loc) · 55 Bytes Code 55% faster with GitHub Copilot Raw Copy Download Edit History


1 <Directory "/skob/">
2 Options -Indexes
3 </Directory>
```



Початок роботи з reCAPTCHA


Використовуйте додаткові функції, наприклад, багатофакторну автентифікацію, захист від спаму та шахрайства, а також інтеграцію з Google Cloud.

- ✓ До 10 000 безкоштовних перевірок на місяць
- ✓ Кредитна картка не потрібна

Ярлик 

sazan24.github.io

17/50

Тип reCAPTCHA 

☐ З використанням оцінок (v3) Виконувати перевірку запитів з урахуванням оцінок


☒ За допомогою завдань (v2)

Перевірка запитів за допомогою завдань

☒ Прапорець "Я не робот"


Виконувати перевірку запитів за допомогою прапорця "Я не робот".

☐ Невидимий значок reCAPTCHA Перевірка запитів у фоновому режимі

Домени 

+ sazan24.github.io

↑↑ [reCAPTCHA \(v2\)](#) ++ [Site Key](#) ↓↓



Налаштування reCAPTCHA завершено

- ✓ Керування налаштуваннями у проекті Google Cloud
- ✓ До 10 000 безкоштовних перевірок на місяць

Щоб увімкнути додаткові функції, відкрийте [проект Google Cloud Platform](#), в якому знаходяться ключі reCAPTCHA.

Використовуйте цей ключ у коді HTML, який ваш сайт передає на пристрої користувачів. [Інформація про інтеграцію на](#)

КОПІЮВАТИ КЛЮЧ САЙТУ

6Lc1jTkrAAAAAC9JqKacgCQ-rbUx7R_zC_bf92Dj

Використовуйте цей секретний ключ для обміну даними між сайтом та сервісом reCAPTCHA. [Інформація про інтеграц](#)

КОПІЮВАТИ СЕКРЕТНИЙ КЛЮЧ

6Lc1jTkrAAAAAEvOXrqZ_UgVa154hx9Sv9MAj60S

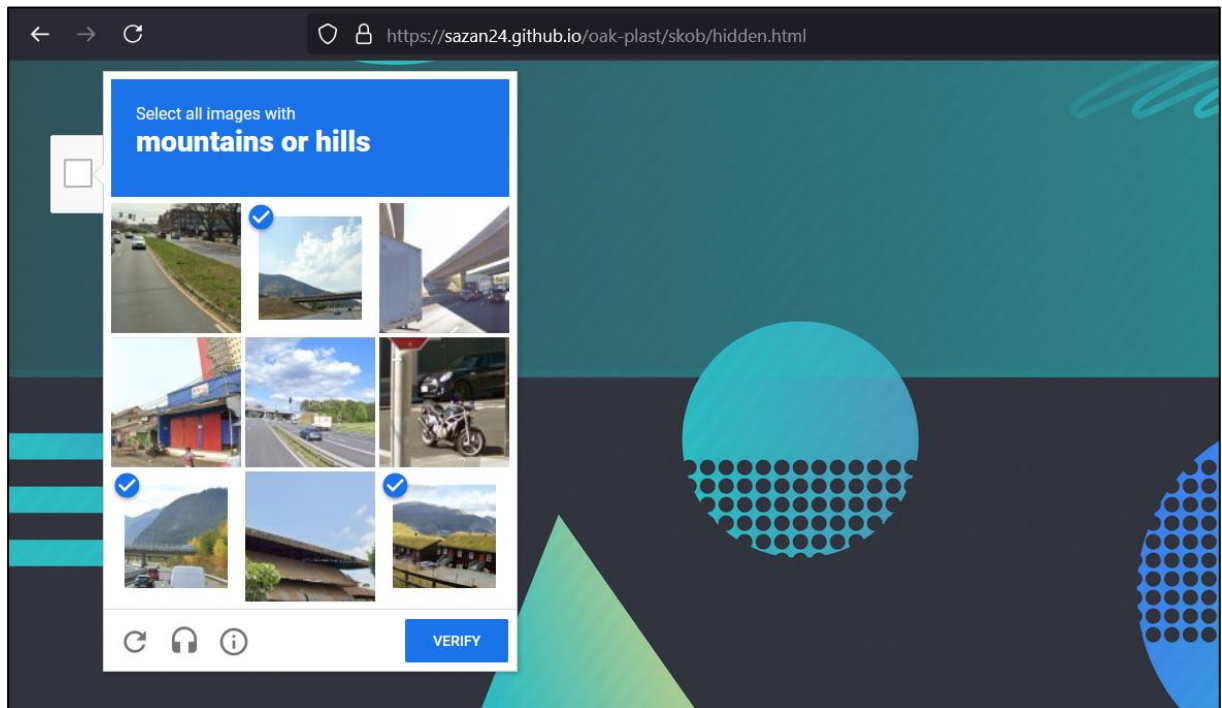
oak-plast / skob / hidden.html

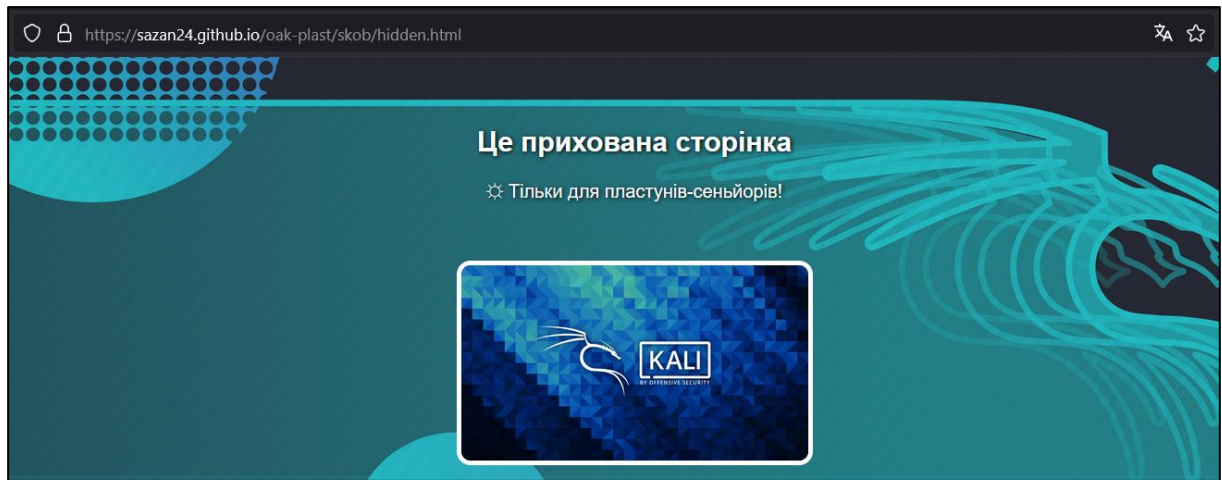
sazan24 Update hidden.html81bb849 · nowHistory

CodeBlame35 lines (32 loc) · 1.14 KBCode 55% faster with GitHub CopilotRawDownloadEditCopy

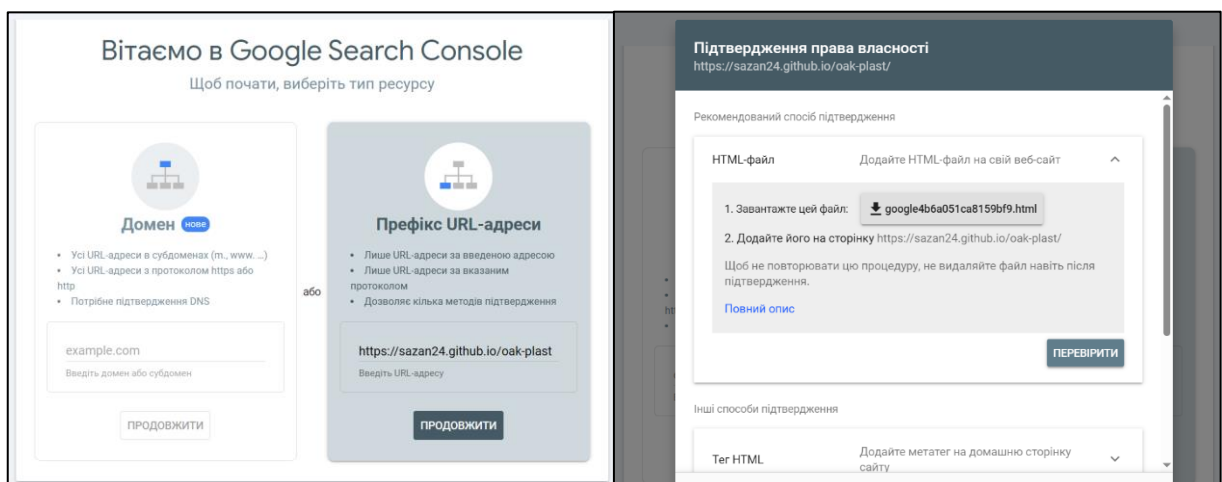
```
1  <!DOCTYPE html>
2  <html lang="uk">
3  <head>
4    <meta charset="UTF-8">
5    <meta name="robots" content="noindex, nofollow">
6    <title>Прихована сторінка</title>
7    <link rel="stylesheet" href="styles.css">
8    <script src="https://www.google.com/recaptcha/api.js" async defer></script>
9  </head>
10
11  <body>
12    <!-- reCAPTCHA -->
13    <div id="captcha" style="margin-top: 2rem;" class="g-recaptcha"
14      data-callback="recaptchaCallback"
15      data-sitekey="6Lc1jTkrAAAAAC9JqKacgCQ-rbUx7R_zC_bf92Dj">
16    </div>
17
18    <!-- Прихований контент -->
19    <div id="secret_content" style="display: none;">
20      <h2>Це прихована сторінка</h2>
21      <p>Тільки для пластунів-сеньйорів!</p>
22      <br>
23      
24    </div>
25
26    <script>
27      const content = document.getElementById('secret_content');
28      const captcha = document.getElementById('captcha');
29
30      function recaptchaCallback() {
31        content.style.display = 'block';
32        captcha.style.display = 'none';
33      }
34    </script>
35  </body>
36  </html>
```

↑↑ Raw View ++ Site Render ↓↓

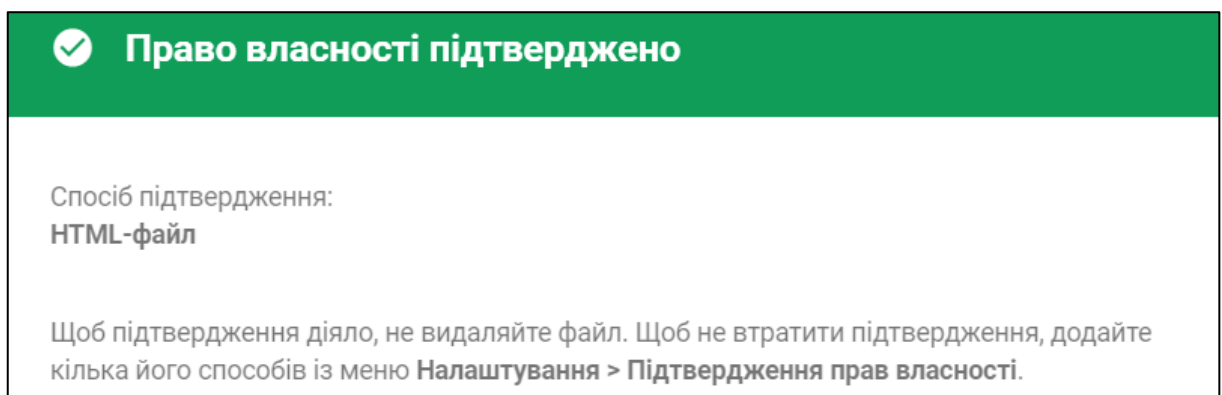


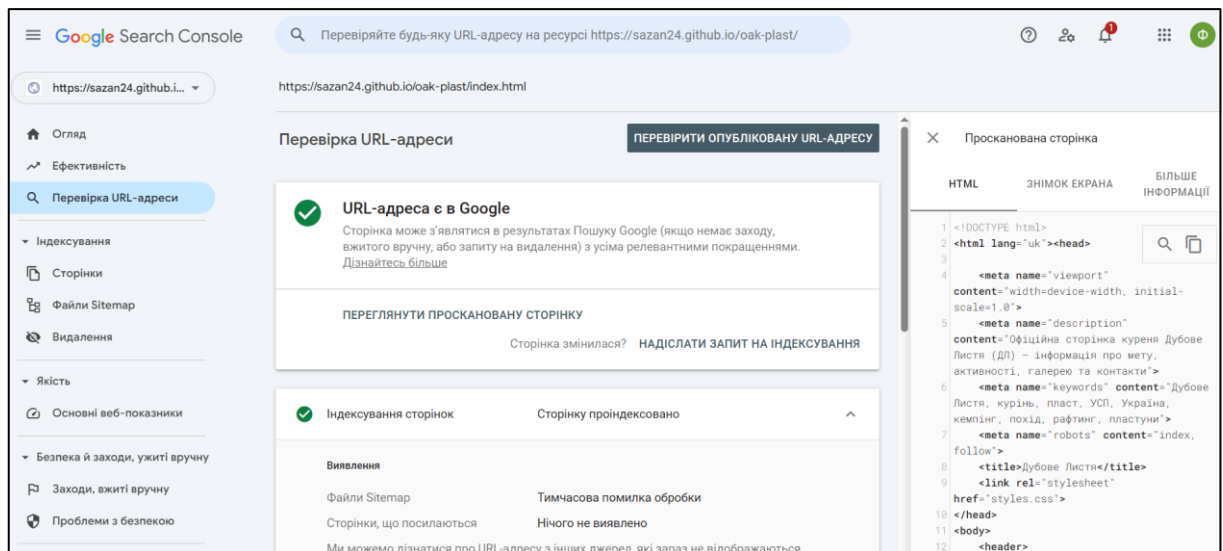


3. Пересвідчитись, що заборонені елементи не відображаються *
- * (див. розділ “Покриття” в інструменті Google Search Console).

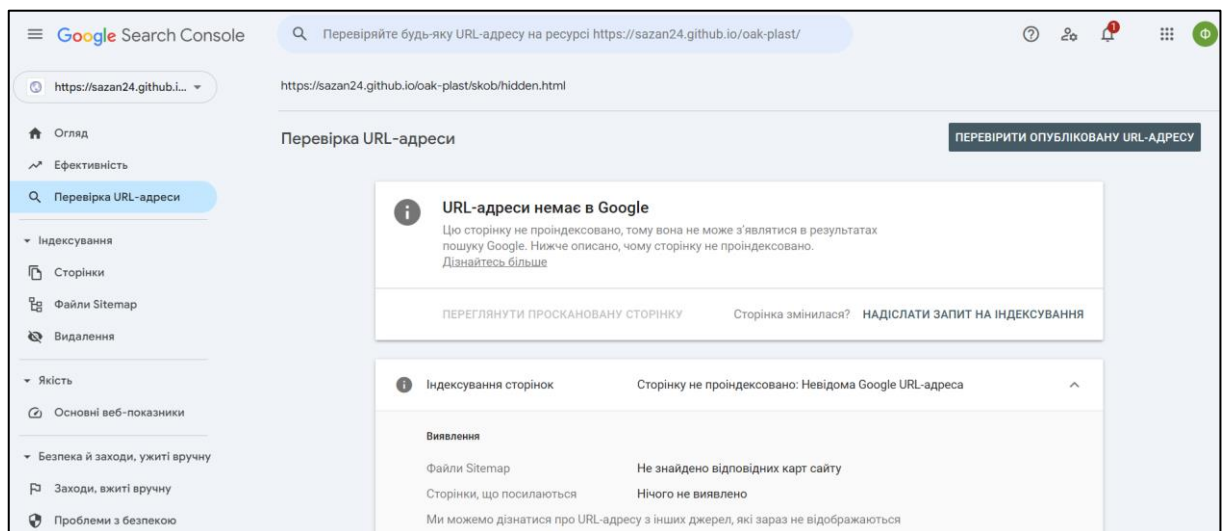


↑↑ [Google Search Console](#) (Ownership Approval) ↓↓





↑↑ «index.html»: OK ++ «hidden.html»: No ↓↓



4. Запропонувати перелік пентестів на основі Metasploit та засобів “Google Dorks” щодо пошуку інформації із заборонених елементів сайту у відкритому доступі.

❖ “Google Dorks” – пошук заборонених елементів сайту:

- `site:rep.github.io inurl:admin` – виявлення панелей адміністрування;
- `site:rep.github.io intitle:"index of"` – сторінки з відкритими каталогами;
- `site:rep.github.io filetype:sql` – витік SQL-дампів бази даних;

- `site:rep.github.io ext:env` – пошук файлів з конфігураціями;
- `site:rep.github.io inurl:backup` – резервні копії, які індексуються;
- `site:rep.github.io inurl:robots` – файл політики індексації сайту.

❖ Пентести з використанням Metasploit Framework:

- `auxiliary/scanner/http/dir_scanner` – виявлення відкритих розділів
- `exploit/unix/webapp/php_cgi_arg_injection` – впровадження коду ч/з CGI
- `exploit/unix/webapp/wp_admin_shell_upload` – завантаження шелкodu ч/з WP
- `auxiliary/scanner/http/http_login` – брутфорс сторінки логін-форм

Контрольні запитання:

1) Роль файлів «`.htaccess`» і «`robots.txt`» в пошуку інформації сторінки?

- «`.htaccess`» – конфігураційний файл веб-сервера Apache, який дозволяє обмежувати доступ до ресурсів, налаштовувати редиректи, а також автентифікацію і т.д.
- «`robots.txt`» – текстовий файл, який вказує пошуковим роботам, які частини сайту дозволено або заборонено індексувати (лише рекомендує не сканувати директорії).

2) Які інструменти пропонує Google для пошуку релевантної інформації?

Google пропонує такі інструменти для пошуку релевантної інформації, як Google Advanced Search (Google Dorks), Google Scholar, Google Image Search, Google Search Console, Google Trends, Google News, Google Books та Google Maps.

3) Які засоби допомагають уникнути витоку даних із ресурсу з відкритим доступом при його відлагодженні (debugging)?

Щоб уникнути витоку даних із ресурсу, зазвичай на практиці використовують тестове середовище з фіктивними даними, шифрування й обмеження доступу через «`.htaccess`», а також виключення з індексації завдяки використанню «`robots.txt`».