УДК 004.056.53 **DOI:10.30837/rt.2024.4.219.03**

І.В. ОЛЕШКО, канд. техн. наук, К.О. ПАПАЗОВ

ЗАБЕЗПЕЧЕННЯ АНОНІМНОСТІ КОРИСТУВАЧІВ ПРИ ПРОВЕДЕННІ ОНЛАЙН-ОПИТУВАНЬ

Вступ

В епоху цифрових технологій, коли обмін інформацією відбувається миттєво, а дані стали цінним ресурсом, питання анонімності користувачів в онлайн-середовищі набуває особливої ваги. Зокрема, це стосується онлайн-опитувань, які є потужним інструментом для збору даних. Анонімність в онлайн-опитуваннях є важливою складовою для забезпечення правдивості відповідей, захисту особистих даних респондента та для запобігання небажаним наслідкам, таким як дискримінація або переслідування.

Проблема забезпечення анонімності користувачів у веб-середовищі є предметом численних досліджень. Багато з них підтверджують широке використання JavaScript для різноманітних цілей, включаючи відстеження поведінки користувачів, зокрема за допомогою трекерів [1]. Трекери можуть збирати різноманітну інформацію, від IP-адрес та даних про браузер до історії відвідувань та поведінки на сайті [2]. Рис. 1 демонструє загальний сценарій веб-відстеження, реалізований JavaScript-програмою. На рисунку показано, як дані про відвідувачів без їх явної згоди можна надсилати через заголовки HTTP на сервери третіх сторін.

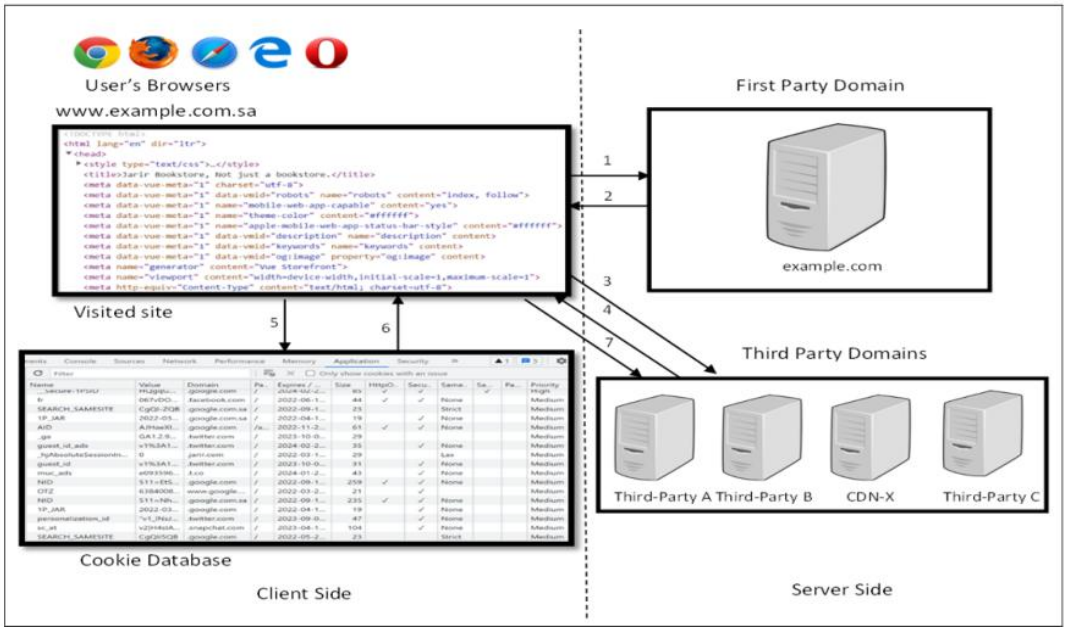

Рис. 1. Процес відстеження веб-сторінки та її відтворення за допомогою виконання JavaScript-програм. Examples.com використовує сторонні сервери (A, B, CDN-X і C) для отримання додаткового контенту та аналітичних досліджень

Останнім часом розроблено декілька інструментів для захисту від відстеження веб-користувачів, включаючи плагіни для браузерів. Ці інструменти засновані на блокуванні програм JavaScript та інших компонентів веб-сайту, які впливають на порушення анонімності. Наразі бракує ефективного підходу, який би надав оцінювання щодо анонімності респондентів при проведенні опитувань онлайн-додатками.

Метою статті є порівняльний аналіз розповсюджених онлайн-сервісів для проведення опитувань за критерієм анонімності. В статті також розглядаються основні інструменти для дослідження анонімності. Дослідники та розробники можуть використовувати наші висновки для покращення анонімності користувачів онлайн-додатків, а респонденти – для вибору найбільш безпечного додатку для проходження онлайн-опитувань.

1. Веб-додатки анонімних опитувань

На сьогодні найбільш розповсюдженими веб-додатками для проведення анонімних опитувань є Google Forms, SurveyMonkey, Typeform, Survio, Xoyondo та Mentimeter (Menti) [3]. Проведемо порівняльний аналіз веб-додатків для анонімного опитування за наступними критеріями: безпека та конфіденційність, функціональні можливості, вартість використання та аналіз зворотних посилань. Аналіз зворотних посилань (Total Sites Linking In) є важливим аспектом для оцінки авторитетності та популярності веб-сайтів. Хоча цей показник не впливає безпосередньо на анонімність користувачів під час проходження опитувань, він відображає загальну довіру та розповсюдженість платформи. Дані отримано з інструменту Ahrefs [4]. Результати порівняльного аналізу розповсюджених веб-додатками для проведення анонімних опитувань наведені у табл. 1.

Таблиця 1

Аналіз веб-додатків анонімних опитувань

Критерії	Google Forms	SurveyMonkey	Typeform	Survio	Mentimeter	Xoyondo
Безпека та конфіденційність	Захист від XSS, CSRF, SQL-ін'єкцій. Використання TLS	Захист від XSS, CSRF, SQL-ін'єкцій. Використання TLS	Захист від XSS, CSRF, SQL-ін'єкцій. Використання TLS	Захист від XSS, CSRF, SQL-ін'єкцій. Використання TLS	Захист від XSS, CSRF, SQL-ін'єкцій. Використання TLS	Захист від XSS, CSRF, SQL-ін'єкцій. Використання TLS
Функціональні можливості	20+ типів питань, інтеграції з Google Sheets	25+ типів питань, аналітика, API інтеграції	30+ типів питань, інтерактивні форми, API інтеграції	15+ типів питань, експорт даних	10+ типів питань, інтерактив в реальному часі	5+ типів питань, планування зустрічей, голосування
Вартість використання	Безкоштовно	Від $25/міс	Від $35/міс	Безкоштовно	Від $10/міс	Безкоштовно
Кількість зворотних посилань	1,000,000+	500,000+	100,000+	50,000+	50,000+	10,000+

Порівнюючи Google Forms, SurveyMonkey, Typeform, Survio, Mentimeter та Xoyondo, можна зробити висновок, що всі розглянуті платформи забезпечують високий рівень безпеки та конфіденційності, мають дружній інтерфейс користувача та пропонують різні вартості використання. З розглянутих додатків варто виділити Google Forms. Google Forms надає безкоштовний доступ до платформи з веб-інтерфейсом та має найбільшу кількість зворотних посилань.

Розглянуті платформи значною мірою покладаються на JavaScript для забезпечення інтерактивності, збору даних та обробки результатів. JavaScript-програми, що використовуються в цих додатках, можуть автоматично збирати різноманітні метадані, які можуть бути використані для ідентифікації користувачів, навіть якщо вони не надають свої особисті дані безпосередньо в опитуванні. Наприклад, IP-адреса може бути використана для приблизного визначення місцезнаходження користувача, а комбінація даних про браузер та операційну систему може створити унікальний "відбиток" пристрою, що дозволяє відстежувати користувача навіть без використання cookies. Крім того, багато платформ використовують сторонні скрипти для аналітики, реклами та інтеграції з соціальними мережами, що ще більше збільшує ризик трекінгу. Наприклад, інтеграція з Facebook Pixel дозволяє відстежувати дії користувачів на сайті опитування та використовувати цю інформацію для таргетованої реклами в Facebook.

Для покращення анонімності в веб-додатках для опитувань можна застосовувати наступні методи:

1. Псевдонімізація даних: заміна ідентифікуючих даних, таких як IP-адреси, імена користувачів або адреси електронної пошти, на псевдоніми або унікальні ідентифікатори.

Це дозволяє зберігати дані опитування без прямої прив'язки до особи користувача, забезпечуючи конфіденційність.

2. Мінімізація збору даних: збір лише тієї інформації, яка є абсолютно необхідною для проведення опитування. Уникання збору зайвих даних, таких як повна адреса проживання або номер телефону, якщо вони не є критично важливими для дослідження.

3. Обмеження виконання сторонніх скриптів (CSP): використання Content Security Policy (CSP) для контролю завантаження та виконання сторонніх ресурсів, таких як скрипти, стилі та зображення. CSP дозволяє визначити, з яких джерел дозволено завантажувати ці ресурси, що зменшує ризик використання шкідливих скриптів, які можуть відстежувати користувачів.

4. Шифрування даних (HTTPS): Забезпечення шифрованого з'єднання між браузером користувача та сервером, на якому зберігаються дані опитування. HTTPS гарантує, що дані, що передаються між браузером та сервером, не можуть бути перехоплені та прочитані третіми особами, забезпечуючи конфіденційність та цілісність даних.

5. Використання локального зберігання даних: за можливості, зберігання даних опитування на стороні користувача, наприклад за допомогою Local Storage або IndexedDB, та їх подальша передача на сервер лише за згодою користувача. Це надає користувачам більше контролю над своїми даними та підвищує рівень анонімності.

6. Використання Tor або VPN: використання мережі Tor або VPN може допомогти приховати IP-адресу користувача, що ускладнює його ідентифікацію та відстеження його дій в інтернеті. Tor перенаправляє трафік через кілька випадкових вузлів, а VPN створює захищений тунель між комп'ютером користувача та сервером.

2. Інструменти для дослідження анонімності

Для аналізу анонімності та виявлення потенційних вразливостей веб-додатків існують різноманітні інструменти, які можна класифікувати за їхнім призначенням та функціональністю:

- **Розширення для браузерів:**

1. Ghostery[5]: аналізує веб-сторінки та відображає список трекерів, рекламних мереж та інших сторонніх скриптів, дозволяючи користувачам блокувати їх вибірково або всі одразу. Ghostery також надає інформацію про призначення кожного трекера, що допомагає користувачам приймати обґрунтовані рішення.

2. Privacy Badger[6]: автоматично блокує трекери, які відстежують користувачів на різних веб-сайтах. Privacy Badger використовує евристичний підхід, аналізуючи поведінку скриптів, а не списки блокування, що дозволяє йому ефективно боротися з новими та невідомими трекерами.

3. uBlock Origin[7]: легкий та ефективний блокувальник реклами та трекерів, який використовує мінімум ресурсів браузера. uBlock Origin базується на списках блокування, але також підтримує користувацькі фільтри та правила.

4. AdBlock Plus[8]: популярний блокувальник реклами, який також блокує багато трекерів. AdBlock Plus дозволяє користувачам створювати власні фільтри та білі списки.

5. NoScript[9]: блокує виконання всіх JavaScript, Java та інших скриптів за замовчуванням, забезпечуючи максимальний рівень захисту, але може суттєво обмежувати функціональність веб-сайтів. NoScript вимагає від користувачів ручного налаштування дозволів для кожного сайту.

6. Disconnect[10]: блокує широкий спектр трекерів (включаючи рекламні, аналітичні, соціальних мереж та контент-провайдерів), використовуючи бази даних відомих трекерів та евристичні методи. Візуалізує трекінг, відображаючи карту запитів з веб-сторінки, показуючи, які домени відстежують вашу активність. Забезпечує захист від деяких шкідливих програм та фішингових сайтів, блокуючи доступ до відомих шкідливих ресурсів. Табл. 2 описує кількість користувачів, правила фільтрації та статистику для інструментів захисту конфіденційності [11].

Таблиця 2

Правила фільтрації та статистика для інструментів захисту конфіденційності

Розширення для браузера	Користувачі	Правила фільтрування
Disconnect	400,000+	Чорний список
Ghostery	2,000,000+	Чорний список
Adblock Plus	43,000,000+	EasyList
uBlock	1,000,000+	Чорний список
NoScript	100,000+	Білий список
Privacy Badger	1,000,000+	Евристичний алгоритм

- **Інструменти розробника, вбудовані в браузер (Developer Tools)**:

1. Network (Мережа): Показує всі запити, які браузер відправляє на сервери, включаючи запити на завантаження скриптів, зображень та інших ресурсів. Це дозволяє ідентифікувати трекери та аналізувати їхню поведінку.
2. Storage (Сховище): Показує cookies, Local Storage та інші дані, які веб-сайти зберігають в браузері користувача. Це дозволяє виявити трекінгові cookies та інші механізми відстеження.
3. Debugger (Відладчик): Дозволяє відслідковувати виконання JavaScript-коду та аналізувати його роботу. Це може бути корисним для виявлення складних трекінгових скриптів.

- **Автоматизовані системи тестування**:

1. Selenium WebDriver: Потужний інструмент для автоматизації браузера, який дозволяє створювати скрипти для автоматичного аналізу веб-сторінок та збору даних про JavaScript-програми. Selenium можна використовувати для тестування веб-додатків (функціональне, регресійне тестування тощо), а також для аналізу їхньої безпеки та конфіденційності, зокрема для дослідження поведінки JavaScript та відстеження запитів до сторонніх серверів. Selenium WebDriver підтримує різні браузери, але Google Chrome є одним з найпопулярніших та найчастіше використовуваних браузерів для тестування з Selenium, завдяки своїй стабільності, швидкості та підтримці сучасних веб-технологій. Існує спеціальний драйвер (ChromeDriver), який дозволяє Selenium WebDriver ефективно взаємодіяти з Chrome.

3. Проведення експерименту

3.1. Налаштування експерименту

Інструменти тестування, які використовувалися для збору даних та аналізу веб-додатків для анонімних опитувань:

– набір із шести найпоширеніших веб-додатків для анонімних опитувань, а саме – Google Forms, SurveyMonkey, Typeform, Survio, Xoyondo та Mentimeter (Menti);

– шість найбільш популярних безкоштовних розширень для браузера (Disconnect, Ghostery, Adblock Plus, uBlock, NoScript та Privacy Badger);

– тести проводилися на операційній системі Windows 11 Pro x64 та IDE PyCharm 2024.3.1.1 (Community Edition) для написання Python-коду;

– відкритий веб-браузер із відповідним веб-драйвером для автоматизованого тестування веб-додатків. У нашому експерименті використовувався браузер Google Chrome з ChromeDriver[12] та Selenium WebDriver для автоматичного тестування сайтів.

3.2. Збір даних

Для сканування шести найпопулярніших веб-додатків анонімного опитування ми імітували процес перегляду додатків на пристрої за допомогою Selenium WebDriver [13] – набору інструментів, який дозволяє отримати всі елементи DOM-дерева HTML-сторінки. Використовуючи Python-скрипти та веб-браузер, ми автоматизували процес перегляду веб-додатків у Google Chrome. Під час перегляду веб-додатків анонімного опитування наш застосунок очікує 100 секунд. Після цього ми зберігаємо DOM-дерево кожного додатку та шукаємо

JavaScript-програми, які називаються JS-скриптами, коли засоби захисту конфіденційності (PPTs) вимкнені (PPTs Off). В процесі сканування ми змогли створити набір даних із 125 елементів JavaScript. Одночасно ми створили окремий профіль Google Chrome для кожного розширення, щоб паралельно отримати JS-програми, коли PPTs увімкнені. Розширення встановлювалися вручну в окремий профіль Google Chrome. Як результат, ми отримали шість різних профілів Google Chrome із офіційної сторінки розширень Google Chrome. Після цього до кожного з розширень застосовувалися налаштування за замовчуванням, за винятком Adblock Plus і Ghostery. У Adblock Plus ми активували рекомендовані фільтри (блокувати додаткове відстеження, блокувати cookie-попередження, блокувати push-сповіщення та блокувати відстеження іконок соціальних мереж). У Ghostery налаштування за замовчуванням не активують жодних функцій фільтрації для уникнення відстеження веб-додатків. Натомість користувач повинен вручну налаштувати параметри для захисту від веб-відстеження [5].

Усі ці кроки виконувалися для проведення експерименту та захисту веб-додатків анонімного опитування від відстеження. Для оцінки ефективності PPTs ми провели підрахунок кількості заблокованих і дозволених JS-програм під час обробки запитів із 6 найпоширеніших додатків анонімних опитувань. Враховуючи динамічний характер веб-додатків і рекламу, яка з'являлась в різний час при доступі до додатків, експеримент проводився одночасно, щоб забезпечити сталість кількості JS-програм. Рис. 2 узагальнює структуру процесу аналізу.

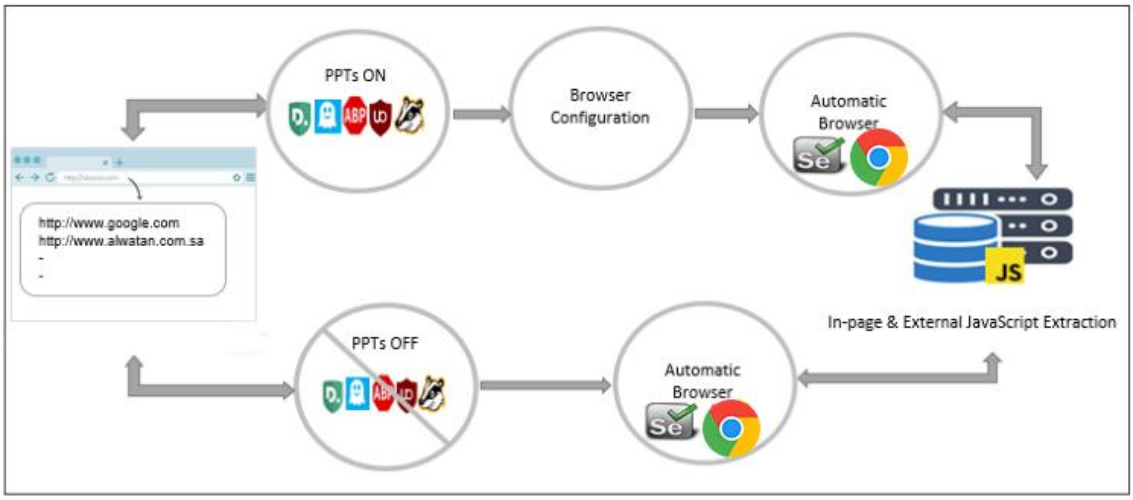

Рис. 2. Платформа вимірювання для інструментів захисту конфіденційності

3.3. Результати дослідження

У цьому розділі представлено результати аналізу веб-додатків онлайн-опитувань щодо забезпечення анонімності респондентів. Дослідження охоплює шість популярних платформ для анонімних опитувань: Google Forms, SurveyMonkey [14], Typeform [15], Survio, Xoyondo [16] та Mentimeter (Menti). Для оцінки ефективності інструментів дослідження анонімності ми отримали загальну кількість програм JavaScript, які були ідентифіковані на шести різних веб-додатках анонімних опитувань, коли засоби захисту конфіденційності вимкнено, а також, коли увімкнений кожен із засобів захисту окремо. На рис. 3 представлений результат дослідження.

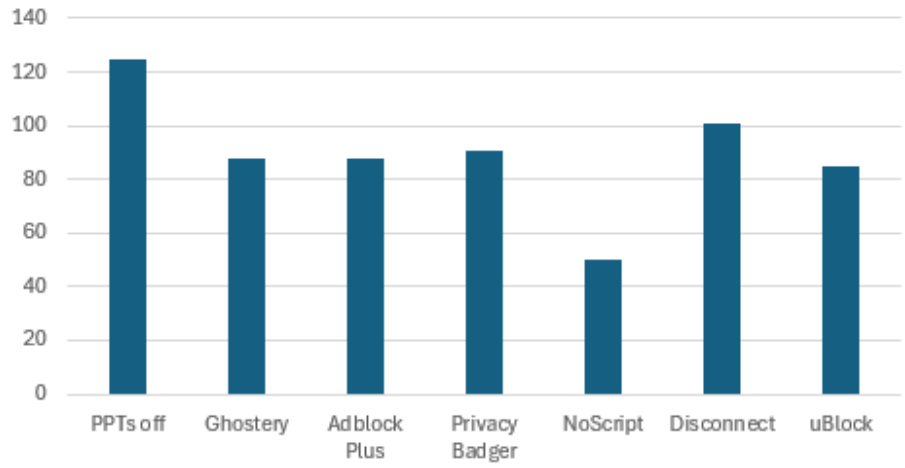

Рис. 3. Загальна кількість програм JavaScript, отриманих із 6 додатків

Табл. 3 відображає відсоткове співвідношення дозволених і заблокованих програм JavaScript для досліджуваних додатків анонімних опитувань загалом.

Таблиця 3

Відсоток усіх дозволених і заблокованих програм JavaScript

JS	PPTs off	Ghostery	Adblock Plus	Privacy Badger	NoScript	Disconnect	uBlock
На сторінці	125	88	88	91	50	101	85
Заблоковано	-	29,6 %	29,6 %	27,2 %	60 %	19,2 %	32 %
Не заблоковано	-	70,4 %	70,4 %	72,8 %	40 %	80,8 %	68 %

Проведені дослідження показують, що інструмент Ghostery здатний блокувати до 87 % [3] сторонніх скриптів, включаючи ті, що використовуються для реклами та веб-аналітики. Privacy Badger використовує евристичний підхід для автоматичного виявлення та блокування трекерів, що дозволяє йому ефективно боротися з новими та невідомими JavaScript. uBlock проявив себе, як ефективний інструмент з низьким споживанням ресурсів.

Нами було проведено середню оцінку кількість заблокованих програм JavaScript для кожного з шести додатків анонімних опитувань. Для розрахунку середньої кількості заблокованих Java-скриптів використовувалась формула

$$average = \frac{\sum_{i=1}^{n} X_i}{n}, \quad (1)$$

де n – кількість інструментів, використаних для забезпечення конфіденційності, X_i – кількість заблокованих Java-скриптів i-тим інструментом.

Ця формула дозволяє об'єктивно оцінити ефективність кожного додатку анонімних опитувань, що є важливим кроком у виборі найбільш оптимального рішення для забезпечення конфіденційності користувачів.

На рис. 4 представлено результати оцінки кількісті заблокованих програм JavaScript для кожного з шести додатків анонімних опитувань.

Табл. 4 показує чисельні значення щодо знайдених програм JavaScript та середньої кількість заблокованих JS-файлів для кожного додатку.

Результати дослідження показали, що на платформі Google Forms кількість заблокованих Java-скриптів дорівнює нулю, що свідчить про гарний рівень анонімності. У той же час на платформі Xoyondo в середньому блокується 22 Java-скрипти, що свідчить про значно більшу кількість JS-програм у веб-додатку та найгіршу анонімність його респондентів.

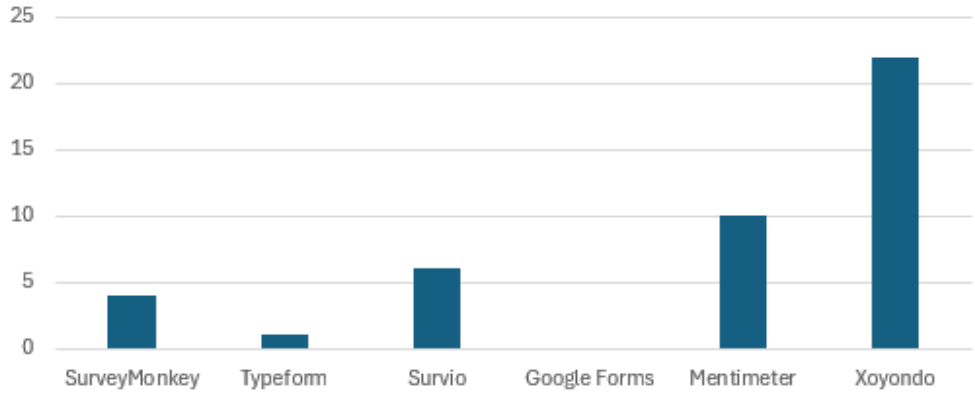

Рис. 4. Середня кількість заблокованих JavaScript для кожного веб-додатку

Таблиця 4

Статистика JS-файлів у веб-додатках для анонімних опитувань

JS	Google Forms	SurveyMonkey	Typeform	Survio	Xoyondo	Mentimeter
Всього	7	24	5	20	37	32
average	0	4	1	6	22	10

В табл. 5 представлено основні JS-файли, які було заблоковано в додатках анонімних опитувань, та надано їх опис.

Таблиця 5

Опис основних скриптів, що можуть впливати на конфіденційність респондентів

JS-файл	**Призначення**	**Джерело**	**Функція**
https://www.googletagmanager.com/gtm.js?id=GTM-WM39SZ5M	Google Tag Manager (GTM) – інструмент, який допомагає веб-майстрам додавати та керувати тегами (фрагментами коду) на своїх вебсторінках.	Google	Використовується для відстеження аналітики, налаштування подій або інтеграції сторонніх сервісів.
https://c.amazon-adsystem.com/aax2/apstag.js	Amazon Ads API	Amazon	Забезпечує завантаження реклами та її персоналізацію на основі користувальницьких даних.
https://securepubads.g.doubleclick.net/tag/js/gpt.js	Google Publisher Tags	Google DoubleClick	Відповідає за динамічне завантаження та управління рекламними оголошеннями.
https://cdn.rudderlabs.com/v2/rudder-analytics.min.js	Скрипт для аналітики від RudderStack.	RudderStack – платформа для потокової передачі даних	Збирає та надсилає дані про взаємодію користувачів з вебсайтом у різні аналітичні системи (наприклад, Google Analytics, HubSpot).
https://analytics.google.com/g/collect	Відстеження взаємодії з вебсайтом через Google Analytics.	Google Analytics	Збирає дані про поведінку користувачів: кількість відвідувачів, сторінки, що переглядаються, місцезнаходження, мову, пристрій тощо.

Висновки

1. Описано проблему забезпечення анонімності користувачів під час проведення онлайн-опитувань. Інтеграція JavaScript як основного інструмента для забезпечення інтерактивності веб-додатків, водночас створює ризики щодо порушення анонімності респондентів за рахунок збору метаданих.

2. Порівняльний аналіз розповсюджених додатків анонімних опитувань показав, що всі розглянуті платформи забезпечують високий рівень безпеки та конфіденційності інформації. З розглянутих додатків було виділено Google Forms, який надає безкоштовний доступ до платформи та має найбільшу кількість зворотних посилань.

3. Надано рекомендації щодо покращення рівня анонімності в веб-додатках для проведення опитувань.

4. Основну увагу приділено оцінці рівня анонімності додатків онлайн-опитувань. У ході дослідження акцентовано увагу на ефективності популярних інструментів захисту конфіденційності, таких як Ghostery, Privacy Badger, uBlock Origin та NoScript. Встановлено, що хоча ці інструменти значно зменшують кількість небажаних трекерів, їх використання може впливати на функціональність веб-додатків. Проведений аналіз найпоширеніших платформ для анонімних опитувань (Google Forms, SurveyMonkey, Typeform, Xoyondo тощо) продемонстрував суттєві відмінності у підходах до забезпечення анонімності. Так, виявлено, що на платформі Google Forms кількість заблокованих Java-скриптів дорівнює нулю, що свідчить про гарний рівень анонімності. У той же час на платформі Xoyondo в середньому заблоковано 22 Java-скрипти, що свідчить про найгіршу анонімність респондентів серед додатків, які розглядалися.

5. У цілому вважаємо, що актуальними та необхідними є подальші дослідження у сфері розробки алгоритмів для виявлення шкідливих JavaScript, які впливають на анонімність, із застосуванням методів машинного навчання. Це дозволить ефективніше ідентифікувати обфусковані коди, які складно виявити традиційними способами.

Список літератури:

1. Ikram M., Asghar H.J., Kaafar M.A., Krishnamurthy B., Mahanti A. Towards Seamless Tracking-Free Web: Improved Detection of Trackers via One-class Learning. arXiv 2017, arXiv:1603.06289.
2. Kalavri V., Blackburn J., Varvello M., Papagiannaki K. Like a Pack of Wolves: Community Structure of Web Trackers // Proceedings of the International Conference on Passive and Active Network Measurement (PAM), Heraklion, Greece, 31 March–1 April 2016; Springer : Berlin/Heidelberg, Germany, 2016, pp. 42–54.
3. 18 сервісів для проведення опитувань [Електронний ресурс]. Режим доступу: https://www.plerdy.com/ua/blog/18-servisov-dlja-provedenija-oprosov/.
4. Your digital marketing strategy backed by real, actionable data [Електронний ресурс]. Режим доступу: https://ahrefs.com/.
5. Bouhnik D., Carmi G. Interface Application Comprehensive Analysis of Ghostery // J. Comput. Syst. 2018. No 5. P. 4–10.
6. Privacy badger. Electronic Frontier Foundation [Електронний ресурс]. Режим доступу: https://privacybadger.org/.
7. UBlock origin – free, open-source ad content blocker. uBlock Origin [Електронний ресурс]. Режим доступу: https://ublockorigin.com/
8. Adblock Plus | The world's #1 free ad blocker [Електронний ресурс]. Режим доступу: https://adblockplus.org/.
9. What is it? – noscript: own your browser! [Електронний ресурс]. Режим доступу: https://www.noscript.net/.
10. Disconnect. [Електронний ресурс]. Режим доступу: https://disconnect.me/.
11. Bubukayr M., Frikha M. Effective techniques for protecting the privacy of web users // Applied sciences. 2023. V. 13, No 5. C. 3191.
12. Chrome DevTools Protocol [Електронний ресурс]. Режим доступу: https://chromedevtools.github.io/devtools-protocol/.
13. Selenium. [Електронний ресурс]. Режим доступу: https://www.selenium.dev/.
14. SurveyMonkey: the world’s most popular survey platform [Електронний ресурс]. Режим доступу: https://www.surveymonkey.com/.
15. Typeform: people-friendly forms and surveys [Електронний ресурс]. Режим доступу: https://www.typeform.com/
16. Xoyondo. Легкі планування та опитування [Електронний ресурс]. Режим доступу: https://xoyondo.com/.

Надійшла до редколегії 09.10.2024

Відомості про авторів:

Олешко Інна Вікторівна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри безпеки інформаційних технологій; Україна; e-mail: inna.oleshko@nure.ua; ORCID: https://orcid.org/0000-0002-8021-0467

Папазов Кирило Олексійович – Харківський національний університет радіоелектроніки, бакалавр; Україна; e-mail: kyrylo.papazov@gmail.com; ORCID: https://orcid.org/0009-0007-0102-461X