

Michał "czesiek" Czyżewski
"Warszawski Hackerspace"

Przyjazna prywatności obserwacja ruchu na stronach WWW

Słowa kluczowe: Google Analytics, narzędzia informatyczne, ochrona danych osobowych, bezpieczeństwo informacji,

Jednym z nieodzownych narzędzi informacyjnych współczesnej biblioteki jest strona internetowa. Tworzona jest zwykle w formie prostego blogu na jednej z bezpłatnych platform lub w pełni „uszytego” na miarę instytucji rozwiązania. Stanowi ona przestrzeń informowania o organizowanych wydarzeniach, prezentacji oferty edukacyjnej i innych ważnych informacji. Często staje się również miejscem interakcji z pracownikami biblioteki czy dyskusji pomiędzy jej użytkownikami.

Szczególnie przydatnymi podczas prowadzenia strony internetowej (nie tylko bibliotecznej) są programy i usługi przeznaczone do analizy ruchu odwiedzających (ang. *web analytics*). Pozwalają one zorientować się w popularności treści zamieszczonych na stronie i dowiedzieć więcej o zainteresowaniach odwiedzających, co pomaga lepiej dopasować treści artykułów czy tematyki reklamowanych wydarzeń. Jednym z najpopularniejszych tego typu rozwiązań jest Google Analytics pozornie bezpłatna usługa oferowana przez internetowego giganta. Niestety, brak opłat z naszej strony jest spowodowany tym, że amerykańska korporacja postanowiła zarabiać na niej w inny sposób tworząc profile internautów i sprzedając dane o nich innym zainteresowanym.

Cechą wyróżniającą Google Analytics spośród podobnych rozwiązań jest jego wszechobecność było to pierwsze narzędzie o tak dużych możliwościach i prostocie obsługi, a jego darmowość przekonała wielu właścicieli stron do implementacji tego rozwiązania. Niewiele osób podejrzewało wtedy, że jest to celowy krok mający napędzać później interes korporacji. Stworzenie bazy stron korzystających z usługi służy zbieraniu dużych ilości informacji o aktywności użytkowniczek i użytkowników internetu, co pozwala z kolei na tworzenie dobrej jakości profili. Cztery lata temu szacowało się, że tylko ok. 1/3 treści w internecie nie ma żadnych oznaczeń firmy Google, które mogą zbierać dane do profili¹. W praktyce oznaczałoby to, że wchodząc na pięć losowo wybranych stron zostalibyśmy zarejestrowani z niemal stuprocentowym prawdopodobieństwem.

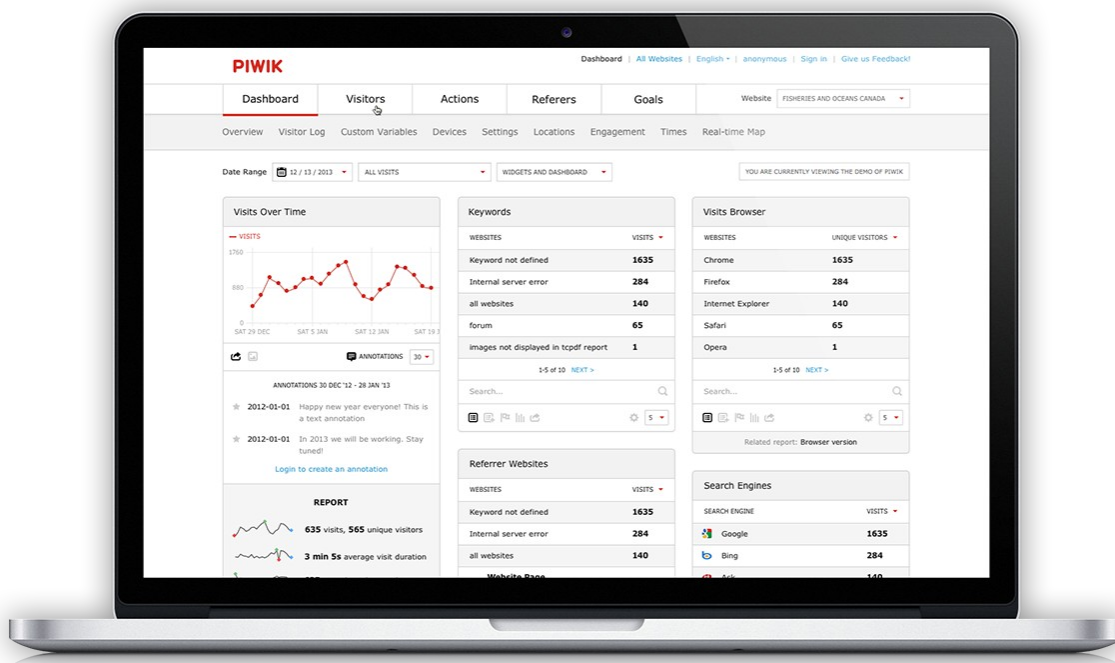
Szkodliwość profilowania do celów reklamowych łatwo sobie wyobrazić, obserwując coraz bardziej podstępne i lepiej dopasowane do naszych zainteresowań reklamy. Mniej oczywistymi są m.in. szkody wynikające z niezgodności naszego profilu z rzeczywistością, np. gdy nasz ubezpieczyciel na podstawie kupionych danych zacznie wyciągać złe wnioski.

Internauci mają niewielkie możliwości ochrony przed takim profilowaniem decyzja o tym,

¹ GELBMANN, M. Google can't track every single click of your web surfing. Only most of them. W: *W3Techs* [on-line] Posted by on 27 February 2012 [dostęp 18.01.2016]. Dostępny w: <http://goo.gl/r3VjLB>.

czy będą obserwowani, leży w gestii twórców i opiekunów stron internetowych, z których korzystają. Jeśli mamy zastrzeżenia do zasad, na jakich Google zbiera i przechowuje nasze dane, to w przypadku skargi, zdani jesteśmy na dobrą wolę wielkiej firmy. Nie ma ona polskiego przedstawicielstwa, więc nie obowiązują jej standardy ochrony danych odpowiednie dla polskich obywateli proponowane przez GIODO na podstawie Ustawy o ochronie danych osobowych (Dz. U. 2015. poz. 2135) http://www.giodo.gov.pl/144/id_art/386/j/pl/. W przypadku naruszenia naszych praw nie jesteśmy w stanie dochodzić roszczeń. Jedyną możliwością w tej chwili jest instalacja oprogramowania blokującego tego typu śledzenie, takiego jak na przykład „Disconnect.me” <https://disconnect.me/>. Decyzja leży tu w rękach opiekunów stron, którzy mogą takiej sytuacji przeciwdziałać.

Jest też inna możliwość. Istnieją alternatywne dla Google Analytics narzędzia pozwalające na obserwację popularności naszej strony. Jednym z nich jest Piwik <https://piwik.org/>, czyli wolne oprogramowanie rozwijane przez programistów z całego świata. Jedną z zalet Piwika jest to, że możemy zachować pełną kontrolę nad zbieranymi danymi, instalując go na własnym serwerze lub hostingu - jest to równie proste, jak założenie bloga. Dalszym krokiem jest podpięcie obserwowanych stron poprzez dodanie do nich naszego elementu śledzącego. Ten etap przebiega identycznie jak w przypadku Google Analytics. W rezultacie dostajemy stronę internetową, na której możemy przeglądać i analizować popularność poszczególnych treści bez szkodenia naszym użytkownikom.



Il. 1. Liberating Analytics. W: PIWIK [on-line] [dostęp 18.01.2016]. Dostępny w: <https://piwik.org/>.

Jeśli zdecydujemy się na Piwika możemy dodatkowo skorzystać z usług płatnego wsparcia w ramach abonamentu Piwik PRO <https://piwik.pro/>². Obejmuje on m.in.

² Wszystkie odesłania do stron internetowych przedstawiają wersję aktualną w dn. 18.01.2016 r.

instalację narzędzia, pomoc techniczną oraz szkolenia dla korzystających z niego pracowników. Co ciekawe, część z tych usług jest świadczona przez firmę Emea z Wrocławia³.

Należy również zastanowić się nad tym, jakie informacje o aktywności naszych użytkowników rzeczywiście wykorzystujemy. Wyłączenie zbierania niektórych danych nie tylko odciąży nasz serwer, ale też przyczyni się do ograniczenia strat w razie ewentualnego „wycieku”. Może się okazać, że interesujące nas informacje już zbieramy w inny sposób (np. są gromadzone przez silnik naszej strony internetowej i dostępne po zalogowaniu do analizatora treści The Webalizer (<http://www.webalizer.org/>) lub są łatwe do pozyskania z logów naszego serwera (przykład dla WordPressa. <https://wordpress.org/plugins/wp-statistics/>).

Innym rozwiązaniem pozwalającym na sondowanie popularności ogłoszenia lub akcji może być też skorzystanie ze skracania adresów linków. Większość z nich ma wbudowane proste statystyki, które pozwalają na zorientowanie się, ile osób skorzystało z danego adresu. Jak widać istnieją alternatywy dla narzędzi proponowanych przez wielkie korporacje. Warto pamiętać o nich wybierając technologie, z których korzystamy, te bowiem mogą ograniczać prawa albo wspierać prywatność innych ludzi.

Bibliografia:

1. GELBMANN, M. Google can't track every single click of your web surfing. Only most of them. W: *W3Techs* [on-line] Posted by on 27 February 2012 [dostęp 18.01.2016]. Dostępny w: <http://goo.gl/r3VjLB>.
2. Ustawa o ochronie danych osobowych (Dz. U. 2015. poz. 2135). W: *GIODO Generalny Inspektor Ochrony Danych Osobowych* [on-line] [dostęp 18.01.2016]. Dostępny w: http://www.giodo.gov.pl/144/id_art/386/j/pl/.

Czyżewski, M. Przyjazna prywatności obserwacja ruchu na stronach WWW. *Biuletyn EBIB* [on-line] 2015, nr 1 (163), *Prywatność w bibliotece*. [Dostęp 25.02.2016]. Dostępny w: <http://open.ebib.pl/ojs/index.php/ebib/article/view/413>. ISSN 1507-7187.

³ EMEA, Piwik PRO Sp. z o.o. ul. Św. Antoniego 2/4 50-073 Wrocław, Polska.