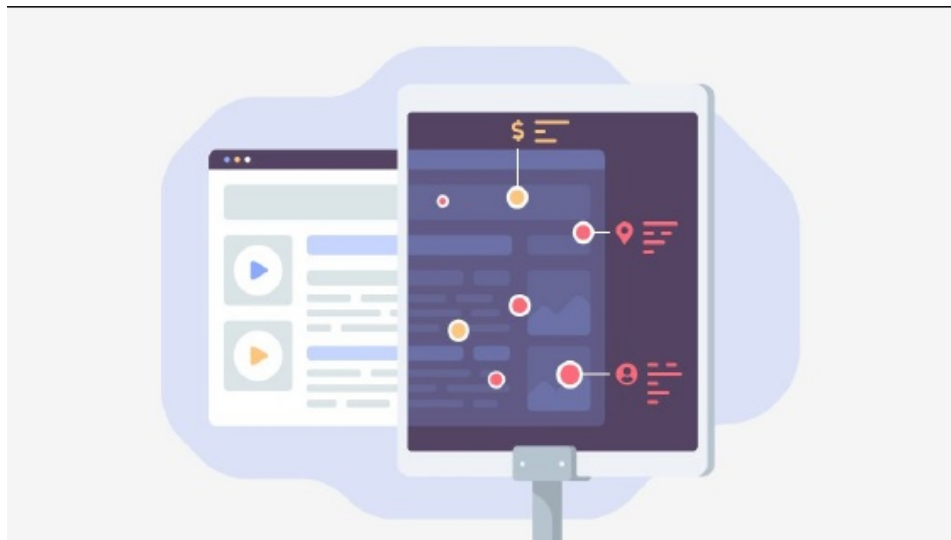


DuckDuckGo Tracker Radar a rejtett követések felfedezéséhez

PCLinuxOS Magazine – 2020. április

A DuckDuckGo [Blog](#)-jából kiemelve
2020. március 5.



- **A DuckDuckGo Tracker Radar a maga kategóriájában legjobb összeállítás az automatikusan generált és a folytonos pásztázások, illetve elemzések segítségével karbantartott követőkről.**
- **Ez az adatbázis most nyilvánosan elérhető kutatás és követők blokkolási listájának létrehozása céljából. A mögötte álló kód pedig nyílt forrású.**
- **Mi ezt a DuckDuckGo Privacy Browser mobil applikációs böngésző követés elleni védelmére és a DuckDuckGo Privacy Essentials asztali böngésző-kiterjesztésben használjuk.**

Manapság, amikor böngészünk, olyan az érzés, mintha a böngészési előzményeink kísértetei vadásznának ránk. A tegnap vett cipő, vagy fejhallgató ma is kitartóan követnek. Ezeket a bosszantó hirdetéseket a legtöbb weblapon leskelődő elrejtett követők táplálják. Sajnos, a vásárlási szokásaid ismerete csak a jéghegy csúcs annak, amit rólad tudnak és kideríthetnek.

A vásárlási előzmények mellett a követők képesek kideríteni a tartózkodási helyeidet, a keresési előzményeket és mást, ami az életkorodból, etnikumodból,

nemedből, érdeklődési körödből és szokásaidból következik. Cégek személyes adataidat részletes profilba rakják és [kiárusítják](#) a legtöbbet ígérőknek.

Az egyik legjobb, amit tehetsz a magad védelmében, hogy minőségi követőblokkolót használj. Miközben az emberek nagy többségének fontos a személyiség védelme, a [privát viselkedéssel foglalkozó kutatásaink](#) azt mutatják, hogy csak az emberek 19%-a használ követés elleni védelmet és nem feltétlenül a legkifinomultabbat. Ezt változtatjuk meg!

Ezért építettünk a DuckDuckGo Privacy Browser mobil (iOS/Android) applikációba és az asztali böngészőkhöz (Chrome/Firefox/Safari) készített DuckDuckGo Privacy Essentials kiterjesztésbe teljes követés elleni védelmet. Ez lehetővé teszi a folyamatos keresést és böngészést privát módban az összes eszközön. Az általunk „alapvető személyesség” dolgait tartalmazzák – követők blokkolását, [privát keresést](#) és [továbbfejlesztett weblap titkosítást](#) – egy csomagban.

Amikor elkezdtünk követés elleni védelmet beállítani, azt láttuk, hogy a követők listáját alapvetően kézzel kezelik, ami azt jelenti, hogy gyakran elavult és sosem teljes. És ami még rosszabb. Azok a listák néha megrongálják a weblapokat, ami akadályozza a széles körű felhasználásukat. Így az utóbbi években összeállítottuk a saját tracker adatbázisunkat egy olyan pásztázó segítségével, ami nem szenved ezektől a hibáktól. Ezt mi DuckDuckGo Tracker Radar-nak hívjuk. Automatikusan készül, folyamatosan frissített és állandóan ellenőrzött.

Ma büszkén adjuk közre a világnak a DuckDuckGo Tracker Radart és nyílt forrásúvá is tesszük az azt létrehozó kódot. Ezt azután tettük, hogy nemrég kiadtuk a [Smarter Encryption adat és pásztázó kódját](#) (az alkalmazásaink és kiterjesztéseink továbbfejlesztett weblaptitkosítási komponense erre épül).

A Tracker Radar ismeri a legelterjedtebb webhelyeken át követőket, részletes információkkal a követési viselkedésükről, benne elterjedtségük, tulajdonosuk, jelölő és süti használatuk, alapelveik, adott forrásokra vonatkozó szabályaik (az oldalfeltérési kivételekkel) és teljesítményadataikat.

- Az egyének a listára épülő alkalmazásaink és kiterjesztéseink használatával profitálhatnak belőle.
- Fejlesztők a saját egyéni követőlistáik összeállítására használhatják fel.
- Kutatók a követők világának vizsgálatára használhatják.

Túl sokan hiszik, hogy az Interneten nem lehet a személyes teret elérni. Nem értünk ezzel egyet és küldetésünknek tartjuk új mércét teremteni az online bizalomnak. A Tracker Radar és kódjának nyilvánossá tételével ezt a missziót fejlesztjük tovább. A továbbiakban sokkal részletesebb leírást adunk a generálása és karbantartása módjáról.

Mi van a DuckDuckGo Tracker Radarban?

A DuckDuckGo Tracker Radar valójában két információs csomagot tartalmaz:

1. egy fájlt az egyes harmadik fél (általában, de nem mindig követéshez kapcsolódó) doménekből részletes információkkal;

Az egyes [domén-fájlok](#) így néznek ki:

```
"domain": "doubleclick.net",
"owner": {
  "name": "Google LLC",
  "displayName": "Google",
  "privacyPolicy": "https://policies.google.com/privacy?hl=en&gl=us",
  "url": "http://google.com"
},
"source": [
  "DuckDuckGo"
],
"prevalence": 0.688,
"sites": 29758,
"categories": [
  "Ad Motivated Tracking",
  "Advertising"
],
"performance": {
  "time": 1,
  "size": 1,
  "cpu": 1,
  "cache": 1
},
"cookies": 0.506,
"subdomains": [
  "googleads.g",
  "stats.g",
  "securepubads.g",
  "pubads.g",
  "bid.g",
  "cm.g",
  "googleads4.g",
  "ad",
  "fls",
  "survey.g"
],
```

```
"fingerprinting": 2,
"resources": [
  {
    "rule": "doubleclick\\.net\\/dbm\\/ad",
    "cookies": 0.00728,
    "fingerprinting": 1,
    "subdomains": [
      "googleads.g"
    ],
    "apis": {
      "Date.prototype.getTime": 1407,
      "Navigator.prototype.userAgent": 1393
    },
    "sites": 3718,
    "prevalence": 0.086
  },
  {
    "rule": "doubleclick\\.net\\/gpt\\/pubads_impl_2020010202\\.js",
    "cookies": 0.0144,
    "fingerprinting": 2,
    "subdomains": [
      "securepubads.g"
    ],
    "apis": {
      "Navigator.prototype.userAgent": 689,
      "Date.prototype.getTime": 780,
      "Navigator.prototype.plugins": 689,
      "Navigator.prototype.mimeTypes": 689,
      "PerformanceTiming.prototype.navigationStart": 641,
      "Document.cookie getter": 763,
      "Document.cookie setter": 623,
      "Navigator.prototype.platform": 561,
      "Date.prototype.getTimezoneOffset": 637,
      "Navigator.prototype.javaEnabled": 615,
      "window.devicePixelRatio": 638,
      "Navigator.prototype.deviceMemory": 581,
      "Navigator.prototype.appName": 483,
      "Navigator.prototype.language": 403,
      "Screen.prototype.height": 496,
      "Screen.prototype.width": 496,
      "Screen.prototype.availHeight": 496,
      "Screen.prototype.availWidth": 496,
      "Screen.prototype.colorDepth": 496,
      "Element.prototype.getClientRects": 149,
      "Navigator.prototype.cookieEnabled": 3,
      "window.name": 69,
      "window.localStorage": 9,
      "Screen.prototype.orientation": 4
    },
    "sites": 791,
    "prevalence": 0.0183
  }
],
```

A teljes lebontáshoz lásd a [dokumentációt](#), itt a kiolvasott főbb összetevők rövid magyarázatát adjuk meg:

- Prevalence – a pásztázásban milyen gyakorisággal került elő a domén.
- Sites – a domén a pásztázás során aktuálisan hány oldalon volt megtalálható.
- Categories – a domén harmadik félként történő használatának oka.
- Fingerprinting – harmadik félként milyen valószínűséggel használ a domén fingerprinting eljárást,
- Cookies – harmadik félként milyen gyakran állított be sütitet.
- Performance – erről a doménről forrás betöltésének hatása
- Owner – tulajdonosi adat.
- Resources – szkriptek, pixelek és a domén által elterjedten használt források.
- Subdomains – al-domének, ahol ezek a források megtalálhatóak.

2. a doménekhez kapcsolódó minden egyes származtató entitásról egy fájl;

Az [entitás adatfájlja](#) így néz ki:

```
"name": "Google LLC",
"displayname": "Google",
"properties": [
  "googleapis.com",
  "googleapis.co",
  "google-analytics.com",
  "gstatic.com",
  "googletagmanager.com",
  "google.com",
  "googletagservices.com",
  "doubleclick.net",
  "googlesyndication.com",
  "googleweblight.com",
  "translate.goog",
  "unfiltered.news",
```

```
"admeld.com",
"adgoogle.net",
"yting.com",
"youtube.com",
"googleadservices.com",
"googleusercontent.com",
"2mdn.net",
"blogspot.com",
"googledrive.com",
"googlemail.com",
"googlecommerce.com",
"ggpht.com",
"doubleclickusercontent.com",
"blogger.com",
"blogblog.com",
"feedburner.com",
"ampproject.org",
"googlevideo.com",
"appspot.com",
"google.de",
...
],
"prevalence": {
  "tracking": 0.859,
  "nonTracking": 0.0735,
  "total": 0.932
}
```

Az entitás összes, birtokolt doménjéről tájlista. A doméneket a pásztázónk találja és a WHOIS-ban, valamint az SSL- tanúsítványokban kereszthivatkozással bírnak a doménekhez és entitásnevekhez. A lista annak megállapítására használható, hogy mikor használják harmadik fél kontextusban.

Például a doubleclick.net-et 29758 oldalon találtuk meg a vizsgálat során (68%). Mivel a Google tulajdona, amihez adatbázisunkban 479 domén tartozik, azt jelenti, hogy a doubleclick.net-et az esetek 98%-ában nem Google-oldalon használták.

Mindent összevetve a Google doménjeire a vizsgált oldalak 93,2%-án hivatkoztak, Ezt feljebb „prevalance”-ként láthatjuk.

Mit kezdhettek a Tracker Radarral?

Három fő dolgot csinálhatsz vele:

1. egyének jobb követőblokkolót szerezhetnek be.

A Tracker Radarból származtatott legjobb minőségű követővédelmet akarsz? Akkor ott a DuckDuckGo Privacy Browser mobil alkalmazás (iOS/Android) és a DuckDuckGo Privacy Essentials asztali böngészőkhöz (Chrome/Firefox/Safari).

2. fejlesztők egyéni követőblokkoló listát készíthetnek.

Az adatbázisban található információcsomaggal létrehozhatod a saját követőblokkoló listádat.

Miközben a jelentősebb böngészők fontos lépéseket tesznek a felhasználók megvédésére a követőtől, jelenleg arra koncentrálnak az erőfeszítéseiket, hogy a követőket megakadályozzák sütik elhelyezésében, és korlátozzák a fingerprinting-re használt böngésző-erőforrásokhoz való hozzáférést. Ugyanakkor nem blokkolják a jelenlegi legnagyobb követőket, amivel nyitva hagyják a kikaput sok más jellegű követőnek (pl. az IP-címedet kiolvasó és a böngészési előzményeidet tároló).

A Tracker Radar javíthat ezen, mivel beazonosítja a követőket, amiket teljesen ki lehet tiltani a betöltésből anélkül, hogy a weblap sérülne. Ez az, amit az alkalmazásunkban és kiterjesztésünkben a tracker blockerrel csinálunk és reméljük, mások is követni fognak. Ennek megfelelően készek vagyunk együttműködni a böngészőkkel a Tracker Radar minél jobb beépítésében.

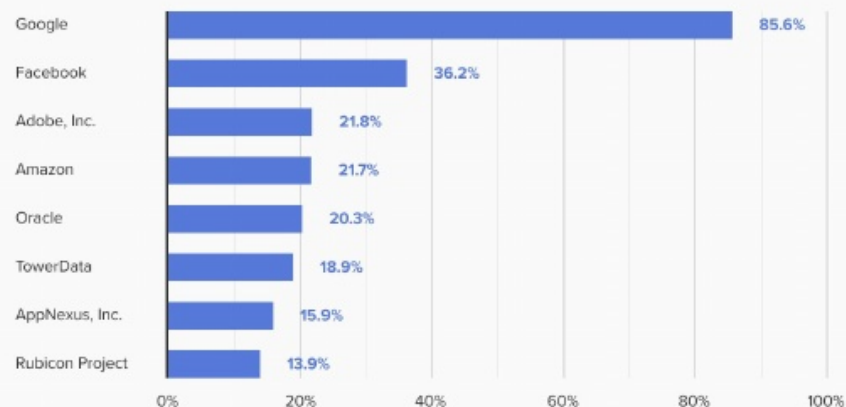
3. kutatóknak a követést tanulmányozhatják.

A Tracker Radar-ban szereplő részletes információkkal könnyebben vizsgálható a követés (és a jövő esetleges más típusú követése). Például az adatbázis mutatja, hogy a Google tracker-ei a top 50 ezer oldal 85%-ában, a Facebook 36%-ában megtalálhatóak.

A legtöbb adatbázis két kategóriába sorolható:

1. blokkolási listák közé, amik majdnem mindegyikét közösségek állították össze és tartják karban;
2. böngészők heurisztikus és öntanuló követő-azonosítói.

Source of the Most Common Trackers Found on the Top 50K Sites



DuckDuckGo

Survey of trackers according to the DuckDuckGo Tracker Radar (Dec 2019)

A közösségi forrású adatok az együttműködők prioritása és beállítottsága függvénye. Ennek megfelelően nem mindig tiszta, hogy miért kerül be valami, a fontossága, aktuálisan érvényessége miatt, hogyan ellenőrizzük és szétverik-e az oldalt? Brave 2018-ban vizsgálta az EasyList-et és úgy találta, hogy az „EasyList [több mint] 90%-a a szokásos böngészés során csekély mértékben használható a nagy mérete és elavult (ritkán használt, vagy lejárt) szabályai miatt.”

Emellett, a közösségi eredetű listák kockázatosak a részletes tesztelés hiányában, ami számos problémát okozhat, mint a honlap változásakor elavuló új szabály, vagy szabályok hozzáadása. Ezzel szemben mi rendszeresen végignézzük a top weblapokat, így mind teljesekek, mind frissek vagyunk. A teljes pásztázás során vizsgáljuk, hogy egy forrást milyen gyakran használnak harmadik fél kontextusban, milyen gyakran helyez el sütit, hogy használja a böngésző API-ját és azokat az API-kat mekkora valószínűséggel használják egyének azonosítására (fingerprinting). A teljes adatállomány esetén vizsgálata könnyebbé teszi az új technikák felfedezését is.

A böngészőbe épített követő-azonosítás a közösségi forrású adatokéhoz hasonlóktól szenvedhet – nehéz a nagy léptékű tesztelés és nem teljes. Emellett, ha nem jól csinálják, megvan a veszélye, hogy azt használják követésre, mivel a

viselkedésed alapján állítja össze a listát és a viselkedésed egyedi, a készített lista használható a te egyfajta lekövetésedre. Mindazonáltal, a böngészőn követő azonosítója kiegészíthet egy jól összeállított blokkolási listát, mint ami a Tracker Radarra épült.

Természetesen tudjuk, hogy mi sem vagyunk tökéletesek és a Tracker Radar fejlesztés alatt áll. Reméljük, hogy bővül és tovább fejlődik a jövőben.

Hogyan juthatok hozzá a Tracker Radar adataihoz?

Az adatok [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International](#) licenc alapján nyilvánosan elérhetőek. A Tracker Radar összeállítására használt kód nyílt forráskódú és a [GitHub-on](#) [Apache 2.0 licenc](#) alatt elérhető.

Örömmel fogadjuk a visszajelzéseket és reméljük, hasznosan találjátok a forrásokat. Amennyiben a Tracker Radar adatait kereskedelmi használatra licencelnéd, kérjük [keress meg](#) minket.

