



Privacypolicy

DMAP

Datum

12 januari 2022

Classificatie

Publiek

Auteur

SIDN Labs

Blad

1/3

Contact

T 026 352 55 00

support@sidn.nl

www.sidn.nl

Bezoekadres

Meander 501

6825 MD Arnhem

Postadres

Postbus 5022

6802 EA Arnhem

Naam
onderzoek/applicatie

DMAP

Ingangsdatum policy

20 januari 2022

Doel van de applicatie of
het onderzoek

Analyseren/scannen van domeinnamen om informatie te verzamelen over het gebruik van een domeinnaam.

Een domeinnaam wordt gescand d.m.v. verschillende sub-crawlers, zoals:

- DNS
- HTTP
- TLS
- SMTP

De data worden opgeslagen in een database en wordt gebruikt voor onderzoek en interne SIDN-toepassingen zoals:

- Registrar Scorecard (RSC), een incentive programma om internetstandaarden bij registrars te stimuleren (met als doel het verbeteren van de technische en inhoudelijke kwaliteit van .nl).
- De statistiekenpagina van SIDN Labs (stats.sidnlabs.nl) om inzicht te geven in de ontwikkelingen van diverse internetstandaarden (geaggregeerd).
- SIDN Merkbewaking, een tool die gebruikers attendeert op (potentieel) malafide websites.
- Diverse analyses en onderzoeken van collega's van SIDN Labs die gericht zijn op het verhogen van de security en stabiliteit van .nl en het internet.



De crawler beschikt ook over een screenshotfunctionaliteit die potentieel persoonsgegevens kan vastleggen. De functionaliteit staat standaard uitgeschakeld, maar kan voor specifieke onderzoeken en toepassingen worden ingeschakeld.

Persoonsgegevens

Er worden IP-adressen opgeslagen die worden verkregen bij het uitvoeren van DNS-queries (A-, AAAA-records e.d.). Deze IP-adressen zijn van servers waarop bepaalde diensten draaien, zoals een website of mailserver. Deze servers bevinden zich bijna altijd in een datacentrum. Het is echter niet uit te sluiten dat een klein aantal IP-adressen aan DSL/kabel (thuis)gebruikers toebehoren.

DMAP maakt gebruik van een lijst van domeinnamen zoals de .nl-zonefile, alle domeinnamen op deze lijst worden geanalyseerd. Per domeinnaam wordt er een aantal HTML-pagina's gedownload. De ruwe HTML-data worden voor 1 maand opgeslagen voor mogelijke verdere analyse.

Om te bepalen of een website voor zakelijk gebruik is, wordt er geprobeerd om de volgende attributen te vinden op een website:

- Post-/bezoekadres
- KvK-nummer
- Telefoonnummer

De bovenstaande attributen worden ook opgeslagen.

De DMAP-crawler heeft ook de mogelijkheid om een screenshot van een webpagina te maken, hier kunnen mogelijk persoonsgegevens in worden opgenomen.

Grondslag

Er is een gerechtvaardigd belang, de resultaten van DMAP worden gebruikt om de kwaliteit, veiligheid en stabiliteit van .nl en het internet te onderzoeken en eventueel waar mogelijk te verbeteren.

Filters

Geen

Retentie

De data worden op 2 locaties opgeslagen en geanalyseerd:

- Postgresql-database, hier worden de data 6 maanden bewaard;
- Hadoop, de data worden hier voor lange termijn bewaard om historisch onderzoek naar de .nl-zone mogelijk te maken. De data worden voor 5 jaar bewaard.



Datum
12 januari 2022

Classificatie
Publiek

Blad
3/3

Niet alle data zijn relevant voor historisch onderzoek. Deze data verwijderen we na 2 jaar.

Toegang

De data zijn alleen beschikbaar voor geautoriseerde gebruikers van SIDN met een gebruikersnaam en wachtwoord. Het SIDN Labs-personeel heeft een uitgebreide instructie over het belang van privacy gekregen.

Publicatie/delen

De gegevens worden gebruikt voor onderzoek en toepassingen, de onderzoeksresultaten worden mogelijk gepubliceerd. Voor het RSC-programma wordt er een geaggregeerde dataset gedeeld met de afdeling NMS van SIDN. De complete ruwe dataset wordt nooit gedeeld door SIDN.

Type

R&D-onderzoek en SIDN-toepassingen

Andere beveiligingsmaatregelen

Geen