# Privacy Policy

Proactive recognition of domain abuse

| Date | Classification | Page |
|---|---|---|
| 16 June 2020 | Public | 1/2 |
| | Author | |
| | SIDN Labs | |

**Contact**
T +31 (0)26 352 5500
support@sidn.nl
www.sidn.nl

**Offices**
Meander 501
6825 MD Arnhem
The Netherlands

**Mailing address**
PO Box 5022
6802 EA Arnhem
The Netherlands

| | |
|---|---|
| Title of application/study | Proactive recognition of domain abuse |
| Policy start date | 19 March 2020 |

Purpose of application/study

The study has two purposes:

1. To increase the detection of incorrect registration data for the .nl domain
2. To increase the security and resilience of the .nl domain by using registration data to automatically identify suspect registrations

To summarise, the data to be used will come primarily from the DRS database. The data will be copied to a server within the LABS network for processing. Once the training and testing phases of the systems have been completed, the data will be deleted from the server. Other data sources that will be used are takedown notices (which include the domain name, the registrant's name and the reason for the takedown) and blacklists maintained by third parties, such as VirusTotal. The blacklists provide the URLs of websites linked to malicious activity and information about the timing. A system for the automatic detection of incorrect registration data will be created, making use of the DRS database. A system for the automatic detection of suspect registrations will be created, making use of the DRS database in combination with blacklist data. In both cases, the data will be used only to identify the characteristics of newly registered domain names that may indicate that a registration is erroneous or even malicious.

### Personal data

Registration data on all registered domains within the .nl domain, i.e. name, e-mail address, phone number, postcode and street address.

### Legitimate basis

Reasonable interest

### Filters

All the data from the DRS database and takedown notices (which include the domain name, the registrant's name and the reason for the takedown) will be used to train and test the systems. As soon as it is established that certain data in the dataset is no longer relevant to the research, use of the data in question will be terminated and the data will be deleted. Data will be deemed no longer relevant if it is established that inclusion of the data in question has no positive influence on the performance of the systems.

### Retention

Any new data sources (e.g. third-party blacklists) will be used only for training and testing the systems. Such data sources do not include personal data. At the end of the training and testing phase, the data obtained from these sources will be deleted. The training and testing phase is expected to last approximately four months. The systems to be created will not themselves contain any personal data. However, the systems will process incoming data about new registrations.

### Access

All SIDN Labs personnel. The data (DRS database, takedown notices and possibly blacklists) will be held on a server, access to which is controlled on the basis of personalised two-factor authentication (TOTP for SSH and client certificates for web, plus a user name-password combination).

### Publication/sharing

A master's thesis is to be written describing the research. The thesis will provide information about the performance of the systems and will not include any personal data.

### Type

R&D research and prototype

### Other security measures

None.