
Politecnico di Torino

Master's Degree in Computer Engineering (Cybersecurity)

# Development of a tool for the automated analysis and reporting of personal data transfers to non-EEA domains

Supervisor
**prof. Antonio Vetro'**

Candidate
**Lorenzo Laudadio**

in collaboration with:

MONITORA**PA**

*2023-10-27*

# This thesis is divided in **three main conceptual parts**

### 1. Legal Background

Understand the **GDPR** regulations which underlie the **personal data transfers** from the EU and EEA to **third countries**

### 2. Minos development

Develop a software which could help **users** to detect **personal data transfers** to third countries while **navigating the web**

### 3. Analysis on the Italian PA

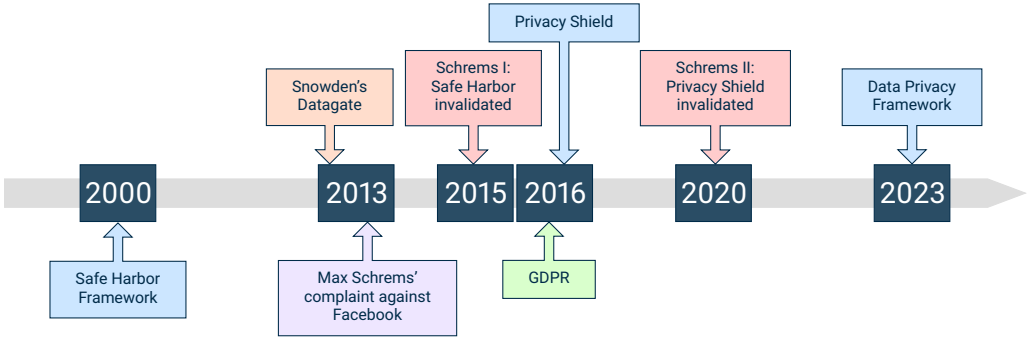
Run a **mass analysis** on data transfers from **Italian public administration websites**, in order to compute and visualize interesting **statistics**

# Personal data transfers from the EU and EEA to third countries are subject to **GDPR rules**

Personal data can only be transferred in one of the following cases:

- The European Commission issued an **adequacy decisions** about the destination Country (art. 45).
- The controller or processor provides **appropriate safeguards**, **enforceable data subject rights** and **effective legal remedies** for data subjects (art. 46).
- Some **specific derogations** are available (art. 49).

# EU-to-US data transfers - History



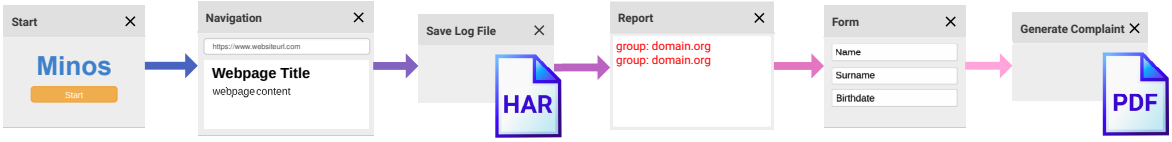
# There are still some **concerns** about the new **Data Privacy Framework**

The US law contains **FISA 702**, which **legitimates mass surveillance against non-US persons**.

FISA 702 has been considered **non compliant with the EU's Charter of Fundamental Rights** by the Court of Justice of the European Union (CJEU).

The US has refused to reform FISA 702, which will have to be **prolonged by the end of 2023**.

# **Minos** is a software which detects **requests to countries** for which an **adequacy decision** has not been issued



# To detect **bad requests**, Minos relies upon an internal **blacklist**

The **blacklist** was compiled by **MonitoraPA**, with the main contribution of **Federico Leva** (an Italian developer and activist).

Federico simply checked the list of **most popular services** in Italy according to **builtwith.com**.

The list is not complete and can change at any time. It includes providers from **US**, **Russia** and **China**.

# **Minos** - Specific requirements

- **Serverless**
- **Minimize** the **dependencies**
- **Cross-platform**
- **Proxyless**
- **Simple** and **minimal codebase**
- **Desktop** application

# **Minos** - Architecture

Minos **architecture** heavily depends upon the **Electron process model**, which involves two separate processes: the **main** process and the **renderer** process. They can communicate thanks to the **Electron IPC**.

### Main

- **Windows/view** management
- **Navigation**
- Network **traffic capture**
- **Log file** creation
- Requests **analysis**
- **Complaint** generation

### Renderer

- Display **web content**
- Display **GUI** (URL bar, form, etc.)

# The **log file** is saved in **HAR** format

**HAR** is a pretty old **JSON-based** format proposed by the W3C Web Performance Working Group. It was **abandoned** in 2012, therefore it is not a standard format.

The HAR format has several disadvantages: a big **space occupance**, **privacy issues**, **inconsistent implementation choices**, etc.

However, we found it being our best option thanks to its **diffusion**: all the **major web browser** implement some version of HAR, even if with some restrictions.

# Comparison with similar software

| Name | Third countries data transfers checking | Anti-tracking | Server-based | Open source |
|---|---|---|---|---|
| Minos | Yes | No | No | Yes |
| Privacy Badger | No | Yes | No | Yes |
| Blacklight | No | Yes | Yes | Yes |
| ImmuniWeb | No | No | Yes | No |
| 2gdpr | Yes* | No | Yes | No |
| webXray | No | Yes | No | Yes |
| OpenWPM | No | Yes | No | Yes |

# **Mass analysis** on the **Italian PA** entities

The goal of this step was to **automatically detect data transfers** to **third countries** within the **Italian public administration** websites.

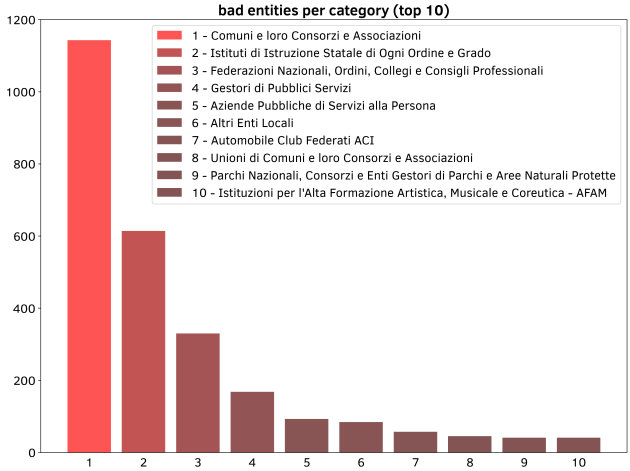
Data is taken from the **OpenDataIPA** portal, a public database which contains all the info about the Italian public entities.

In particular, we were interested in two datasets: **enti** and **categorie-enti**. Thanks to these datasets we can retrieve the **name**, **category** and **website** of all the entities.

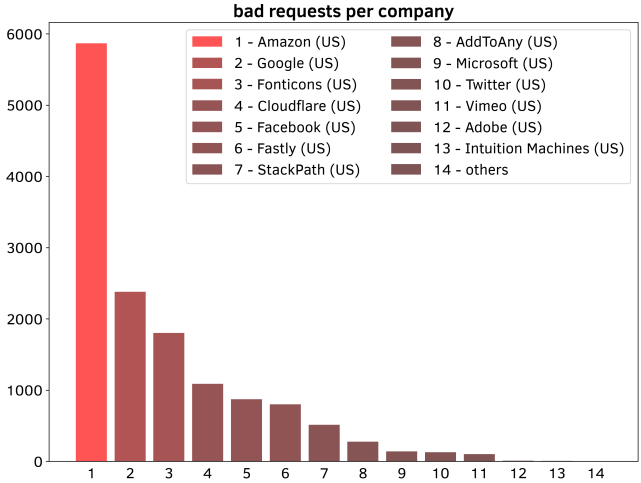
We found a total of **22890** PA entities. Only **19496** ($\sim$85%) websites were actually reachable.

# Bad entities

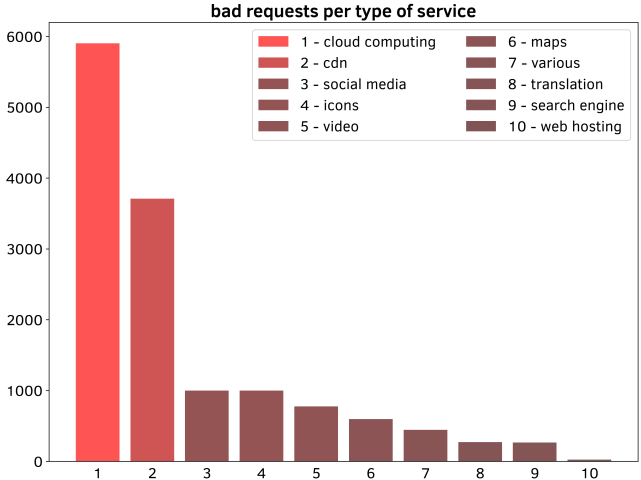
We found **3008** (∼15%) entities making requests to third-countries domains, over a total of 19496 entities analyzed.



# Bad requests per company



# Bad requests per type of service



# Conclusions

After nearly **three years** the **Privacy Shield was declared illegal**, many Italian PA entities still have difficulties **complying with the GDPR**.

Big tech companies offer a **wide range of services** at **competitive prices**, and the alternatives are not so popular for the time being.

# License