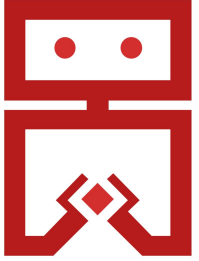# GDPR.observer

Hermes Center is an Italian NGO active since 2012 on the topics of **digital privacy**, **secure whistle-blowing** and **human rights in the digital sphere**. It produces software and develops campaigns and outreach.

The 2022 experience of **Monitora-PA** (In Italy)

It start as a pipeline with these specifications:

1. **Download a database of public agencies** names along with website address and contact email address of the DPO.
2. **Analyze each website and search for google-analytics**. Produce a short report and open data.
3. **Filter the dataset** from point 1, keep only the organizations/websites/emails that match the analysis of point 2.
4. **Send a notification email** to the DPO with a predefined template.
5. Measure and publish removal rate % each month.
6. Repeat from point 2.
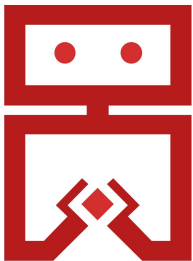
The 2022 experience of **Monitora-PA** (In Italy)

```
(.venv) v :~/D/monitorapa main $ python3 cli/data/enti/download.py

out/enti/2023-02-03/enti.tsv

(.venv) v :~/D/monitorapa main $ cd out/enti/2023-02-03/

(.venv) v :~/D/monitorapa/out/enti/2023-02-03 main $ ls -l

total 7652

-rw-rw-r-- 1 v v 7822027 Feb  3 19:31 enti.tsv

-rw-rw-r-- 1 v v    5637 Feb  3 19:31 LICENSE.txt

-rw-rw-r-- 1 v v     425 Feb  3 19:31 README.md

(.venv) v :~/D/monitorapa/out/enti/2023-02-03 main $ wc -l enti.tsv

22902 enti.tsv
```
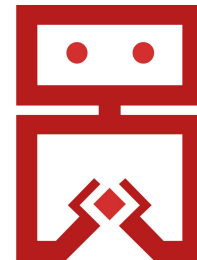
7833 with Google Analytics!

The 2022 experience of **Monitora-PA**: impact

It was also possible because of **two italian perks**

- Exists a national list of all the public agencies, with website and DPO mail address, maintained by AgID.

About us

The Agenzia per l'Italia Digitale - Agency for Digital Italy (AgID) is the technical agency of the Presidency of the Council of Ministers.

The main purpose of the Agency is to guarantee the achievement of the Italian digital agenda objectives and contribute to the diffusion of information and communication technologies, with the aim of fostering innovation and economic growth.

It was also possible because of **two italian perks**

- Exists an email compatible protocol that bound a public official to consider the content of the email (PEC). This has been used to send notifications.

Certified email

Article   Talk

文A Language   ☆ Watch   ✏ Edit

For the whitelisting technique employed by certain internet service providers, see Whitelist § Commercial whitelists.

**Certified email** (known as ***Posta elettronica certificata*** in Italy, or **PEC** in short) is a special type of email in use in Italy,[1] Switzerland,[2] Hong Kong[3] and Germany.[4] Certified email is meant to provide a legal equivalent of the traditional registered mail, where by paying a small fee, users are able to legally prove that a given email has been sent and received.

Registered mail is mainly used in Italy,[5] but there are present efforts to extend its legal validity according to the framework of the European Union.[6]
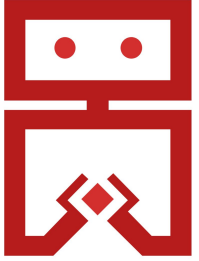
Lessons learnt from **Monitora-PA**

- The Data Protection Authority don't have the capacity to follow up.
- It has also been influential in suggesting technologies to replace Google's services and promote free software alternatives.
- Outreach its achieved also by the sending of thousands of compliance-breach notifications ("good purpose spam").
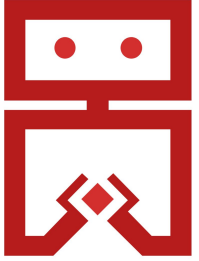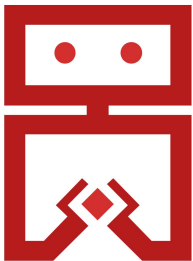
Lessons learnt from **Monitora-PA**

- The DPA issued a statement
  - This was the key
- Technical criticism and quirks
  - See next slide
- Fostered the market for alternatives
  - Still unclear how much we should play a role

The project was first designed by the Hermes Center and then developed by a diverse group of privacy activists.
https://github.com/MonitoraPA/monitorapa

The 2022 experience of **Monitora-PA**: impact

MonitoraPA

⚠ bias ⚠

Improvements to detection software

PA compromesse

variable — Google Analytics — Google Fonts — FontAwesome — Google Maps — Google Hosted Libraries — Google reCAPTCHA

Abstraction of **Monitora-PA** in an **European Perspective**

- We want to **replicate the Italian experience in all EU Member States**.
- Given the lesson learnt, **we can separate the technological effort** and offer a pipeline for use by individuals, NGOs, or even DPAs themselves.
- We enable national stakeholders without the technical-operational complexity.

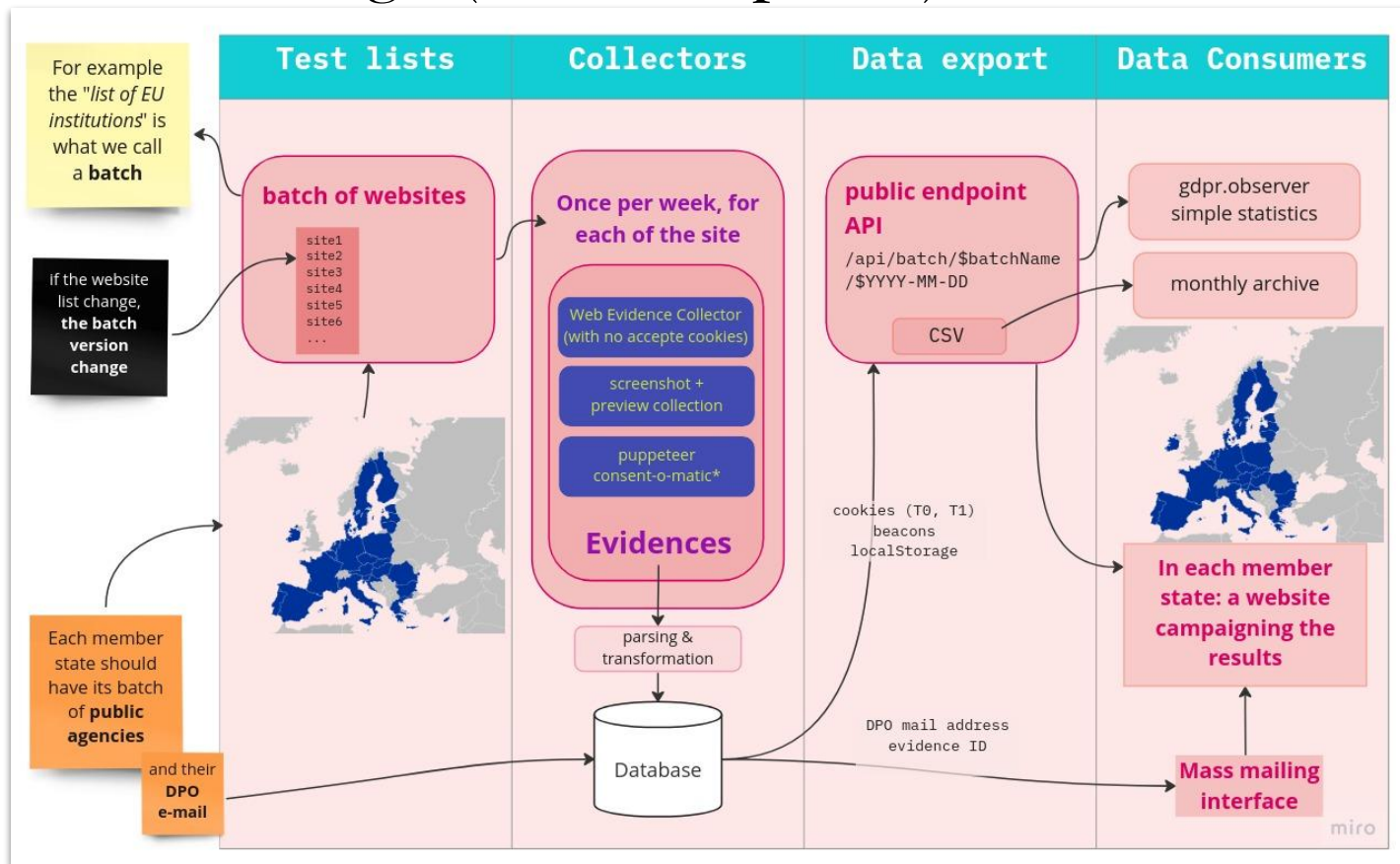| We	Execute weekly GDPR-compliant tests	Publish open data and software interfaces	Manage mass mailing complaint on behalf of national partners
National partner(s)	Provide public agencies websites list with email contacts (DPO)	Promote and advocate nationally the results with awareness campaigns	Write content fitting their advocacy strategy. Issue formal complaints to national DPA.

GDPR.observer

It is a free software-based **infrastructure**.
It **supports many observatories** to ensure that EU Member State institutions are **GDPR**-compliant.
Focus on: web-trackers, informed consent, transport security, cloud location, cookies, and more.

Architecture design (research phase)

For example the "*list of EU institutions*" is what we call a **batch**

if the website list change, **the batch version change**

Each member state should have its batch of **public agencies**

and their **DPO e-mail**

Test lists	Collectors	Data export	Data Consumers

batch of websites

site1
site2
site3
site4
site5
site6
...

Once per week, for each of the site

Web Evidence Collector (with no accepte cookies)

screenshot + preview collection

puppeteer consent-o-matic*

Evidences

parsing & transformation

Database

public endpoint API

/api/batch/$batchName/$YYYY-MM-DD

CSV

cookies (T0, T1)
beacons
localStorage

DPO mail address
evidence ID

gdpr.observer simple statistics

monthly archive

In each member state: a website campaigning the results

Mass mailing interface

miro

Architecture design (first experiments)

Test lists	Collectors	Data export	Data Consumers

```
"European Parliament":
  site: "https://www.europarl.europa.eu/portal/en"

"European Commission":
  site: "https://commission.europa.eu/index_pl"

"Court of Auditors":
  site: "https://www.eca.europa.eu/es/Pages/ecadefault.aspx"

"Council of the European Union":
  site: "https://www.coe.int"

"European Economic and Social Committee (EESC)":
  site: "https://www.eesc.europa.eu"
  selector: '.agree-button'
  dpo: 'todo@email.mail'

...
```

```
v :~/D/gdpr.observer$ time bin/collect-with-wec.mjs --country EUI --source
input/eui-30.json 2>&1 >&/tmp/collect-with-wec9.log

real    8m27.698s
user    1m41.130s
sys     0m14.863s
```

Beacons before consent

Site	Url
osha.europa.eu	https://osha.europa.eu/modules/contrib/eu_cookie_compliance/js/eu_cookie_compliance.m...
	https://osha.europa.eu/sites/default/files/styles/large/public/2021-03/flickr-icon.png?itok=...
	https://osha.europa.eu/sites/default/files/styles/large/public/2021-03/footer-facebook-new...
	https://osha.europa.eu/sites/default/files/styles/large/public/2021-03/footer-youtube-new...
	https://osha.europa.eu/sites/default/files/styles/large/public/2021-03/ico-twitter.png?itok...
	https://osha.europa.eu/sites/default/files/styles/large/public/2021-03/youtube-icon.png?ito...
	https://piwik.osha.europa.eu/piwik/matomo.js
	https://piwik.osha.europa.eu/piwik/matomo.php?action_name=European%20Agency%20fo...
	https://piwik.osha.europa.eu/piwik/plugins/HeatmapSessionRecording/configs.php?idsite=...
	https://platform.twitter.com/js/timeline.16b53cc33aaa562f8f41a495bf720289.js
	https://platform.twitter.com/widgets/widget_iframe.2b2d73daf636805223fb11d48f3e94f...
	https://syndication.twitter.com/i/jot?l=%7B%22_category_%22%3A%22tfw_client_event...
www.cedefop.europa.eu	https://embed.twitcker.com/ticker/CedWebPortal.js
	https://webstats.cedefop.europa.eu/matomo.js
	https://webstats.cedefop.europa.eu/matomo.php?action_name=CEDEFOP%20%7C%20Euro...
	https://www.cedefop.europa.eu/modules/contrib/eu_cookie_compliance/js/eu_cookie_com...
	https://www.google-analytics.com/analytics.js
	https://www.google-analytics.com/j/collect?v=1&_v=j99&aip=1&a=1204667478&t=pagevie...
	https://www.googletagmanager.com/gtag/js?id=UA-4456942-4
www.coe.int	https://static.coe.int/pics/icons/2014/global/youtube-follow.svg
	https://static.coe.int/pics/theme/coe-2014/footer/icon-facebook.png
	https://static.coe.int/pics/theme/coe-2014/footer/icon-flickr.png
	https://static.coe.int/pics/theme/coe-2014/footer/icon-twitter.png
	https://static.coe.int/pics/theme/coe-2014/footer/icon-youtube.png
	https://static.coe.int/pics/theme/coe-2014/share/icon-facebook-rounded.png
	https://static.coe.int/pics/theme/coe-2014/share/icon-linkedin-rounded.png
	https://static.coe.int/pics/theme/coe-2014/share/icon-share.png
	https://static.coe.int/pics/theme/coe-2014/share/icon-twitter-rounded.png
	https://www.facebook.com/tr/?id=1013102268891157&ev=fb_page_view&dl=https%3A%2...
	https://www.facebook.com/v10.0/plugins/page.php?adapt_container_width=true&app_id=...
www.eacea.ec.europa.eu	https://webanalytics.europa.eu/ppms.js
	https://webanalytics.europa.eu/ppms.php?action_name=EACEA&idsite=e546631c-9ccd-454...
www.eca.europa.eu	https://cdn.matomo.cloud/eca.matomo.cloud/matomo.js
	https://eca.matomo.cloud/matomo.php?action_name=Home%20%7C%20TRIBUNAL%20DE...
	https://www.eca.europa.eu/en/PublishingImages/icons/rss.png
	https://www.eca.europa.eu/Style%20Library/eucookielaw.css?rev=20170713

Architecture design (first experiments)
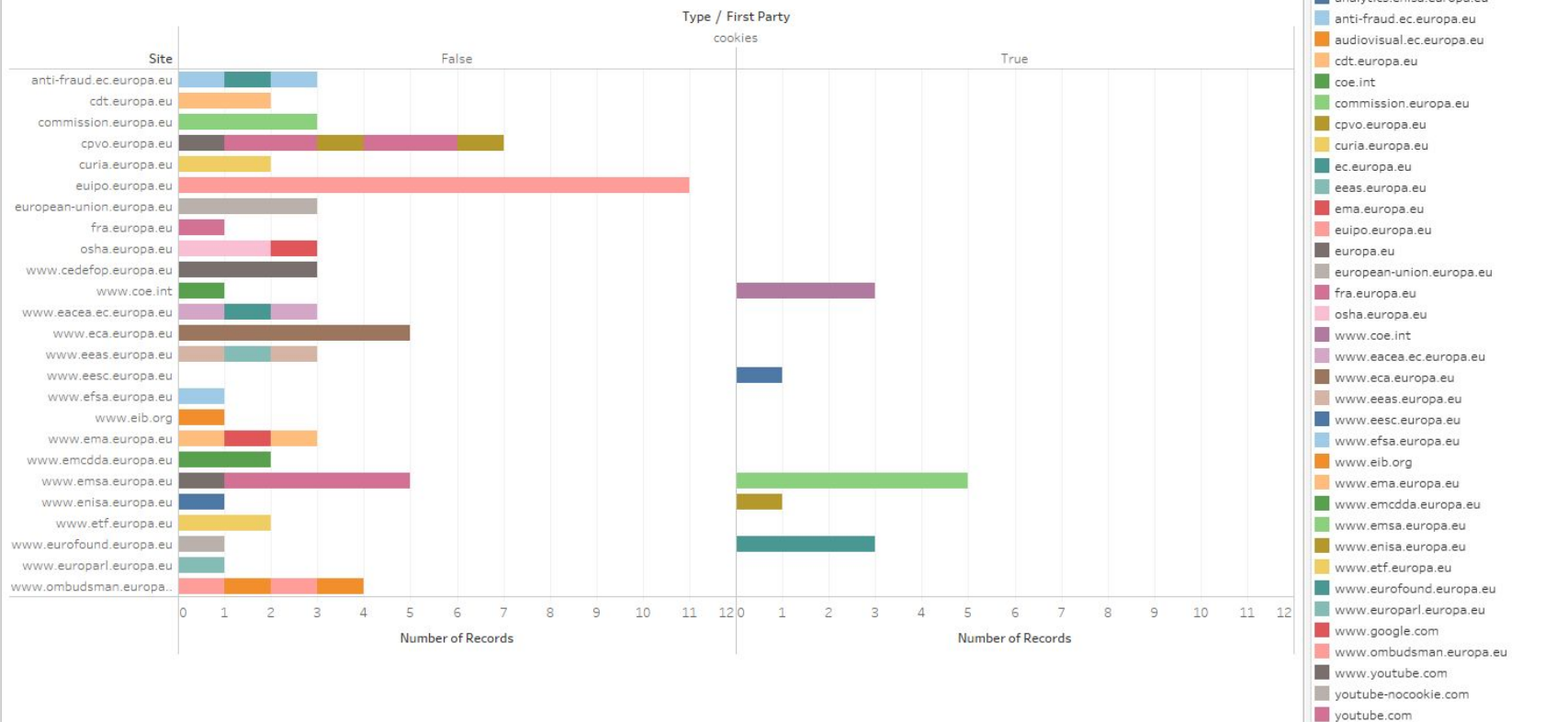
Action plan

- Complete the website with some informative material.
- Keep repeating the test on weekly basis, and complete performance measurement.
  - How much disk we need, how many sites we can quickly test, performances, reliability.
- Assist the *early adopters* and understand **how much is it replicable outside of Italy**.

Next steps

1. Extend and stabilize first full batch of GDPR–compliance tests.
2. Prepare outreach material for partners
3. Extend and support localization and collaborative improvement of Consent automation.
4. Improve third party integration to easily consume results (analyze, visualize, etc).
5. Integrate with airtable to pull and update results.
6. Advice GDPR-compliant and digitally sovereign IT tools commonly used by public agencies (e.g. analytics, fonts, …)

Thanks for listening

Claudio Agosti
claudio.agosti@hermescenter.org

https://gdpr.observer

Hermes Center is an Italian NGO active since 2012 on the topics of **digital privacy**, **secure whistle-blowing** and **human rights in the digital sphere**. It produces software and develops campaigns and outreach.

Technical Challenges

- Automatic reliable consent-clicker is challenging:
  - Language dependent (different language require national partner review/cooperation)
  - `Consent-O-Matic` plugin provide a good design pattern and framework
- Each website analyzed may require deeper scanning of multiple web pages to identify GDPR compliance issues (not always triggered in main homepage)
  - Or engage with supporters that navigates on website (data donation model)

Technical Challenges

- Find programmatically the DPO name and email

Existing limit in current analysis method

Existing technical limit in current analysis method: