Cooperate in Open Console

OSSYM#5, 4 Oct 2023

Open Console

Cooperate on what we now
about
websites / domains / networks

Open Console Project

- Sub-grantee of Open WebSearch .eu

https://open-console.eu
team@open-console.eu

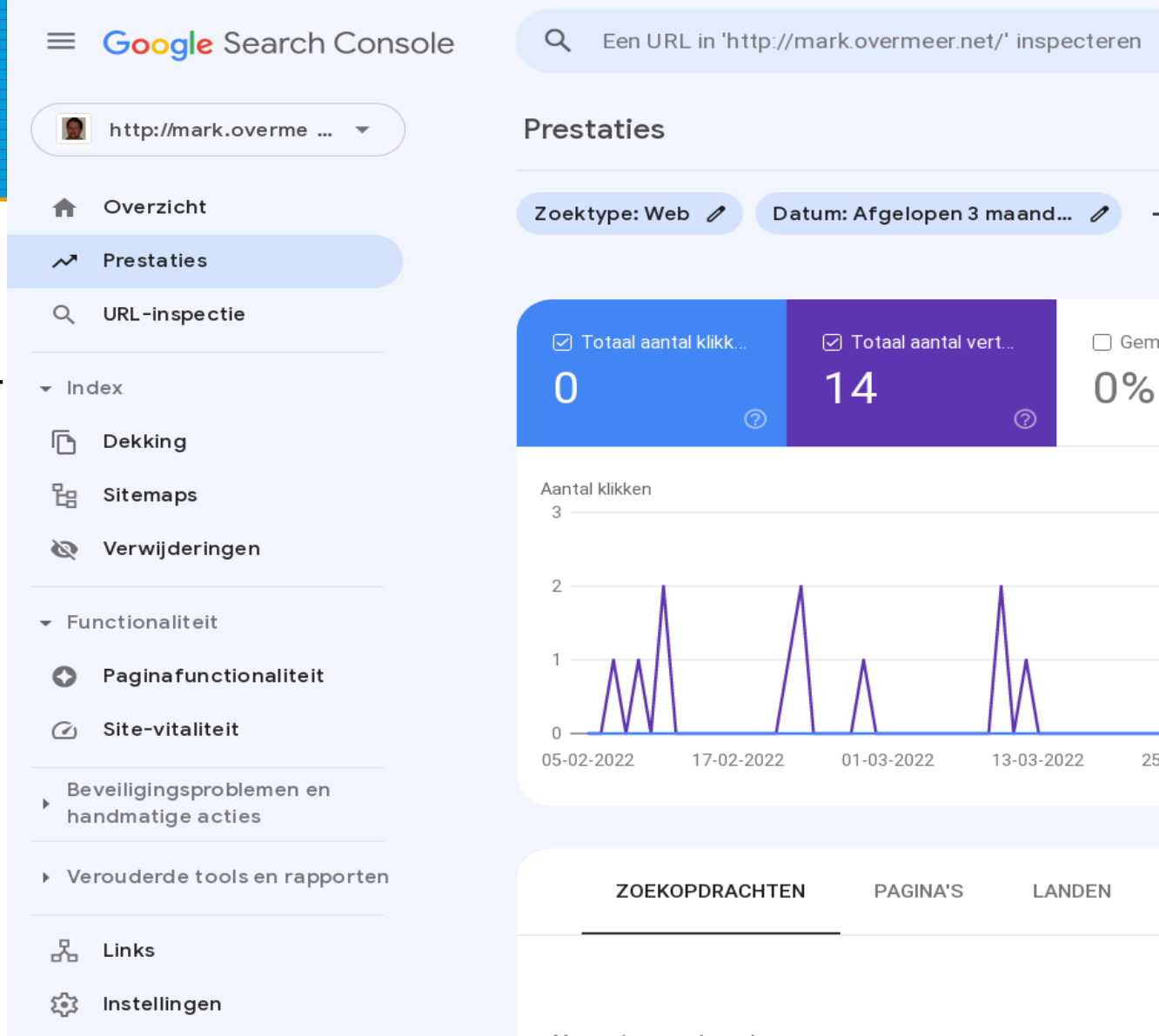
- Run by Markov Solutions

Thao

t Community
er Finances, etc

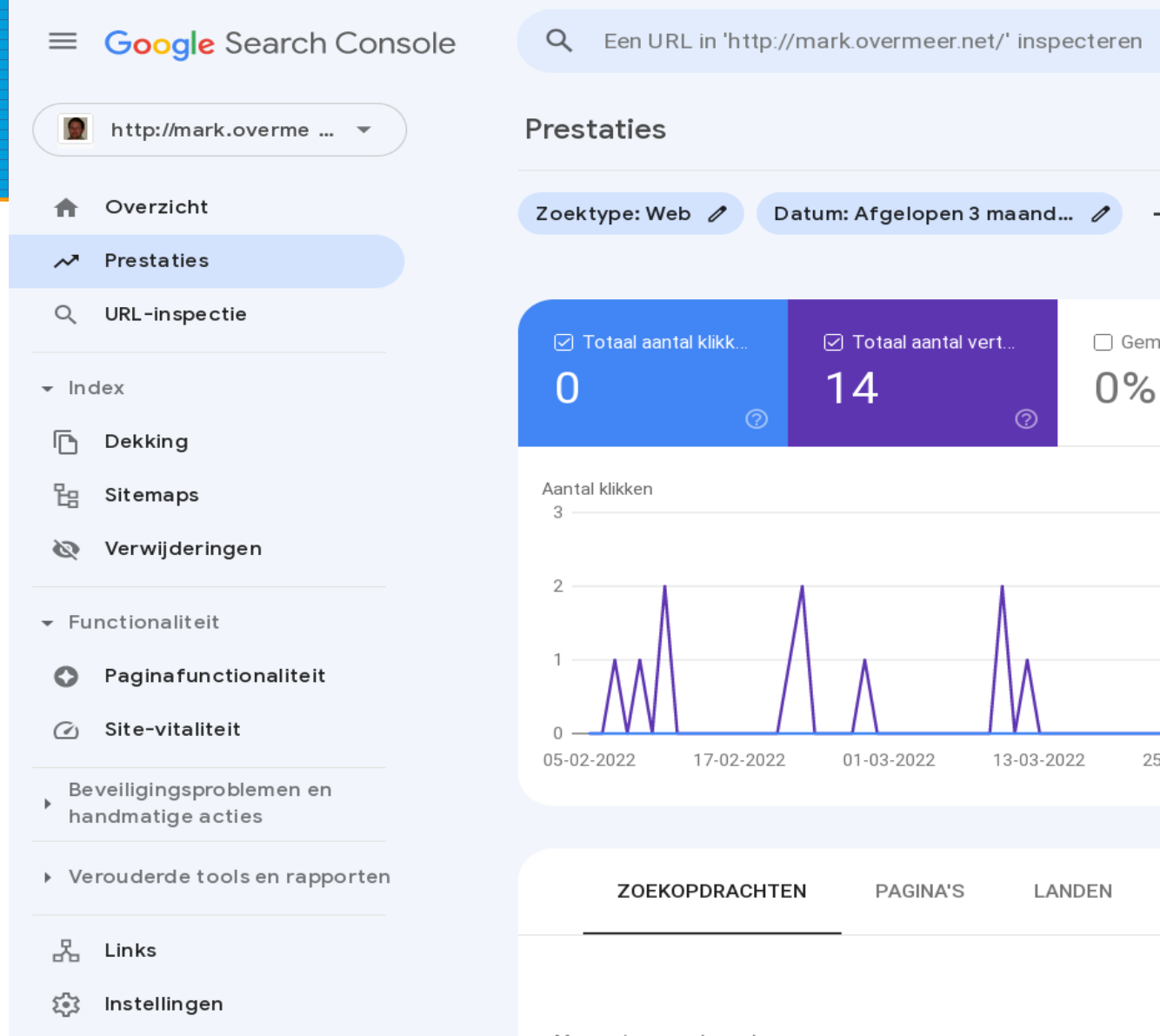
GSC

{Bing, Yandex, Baidu} webmaster



GSC

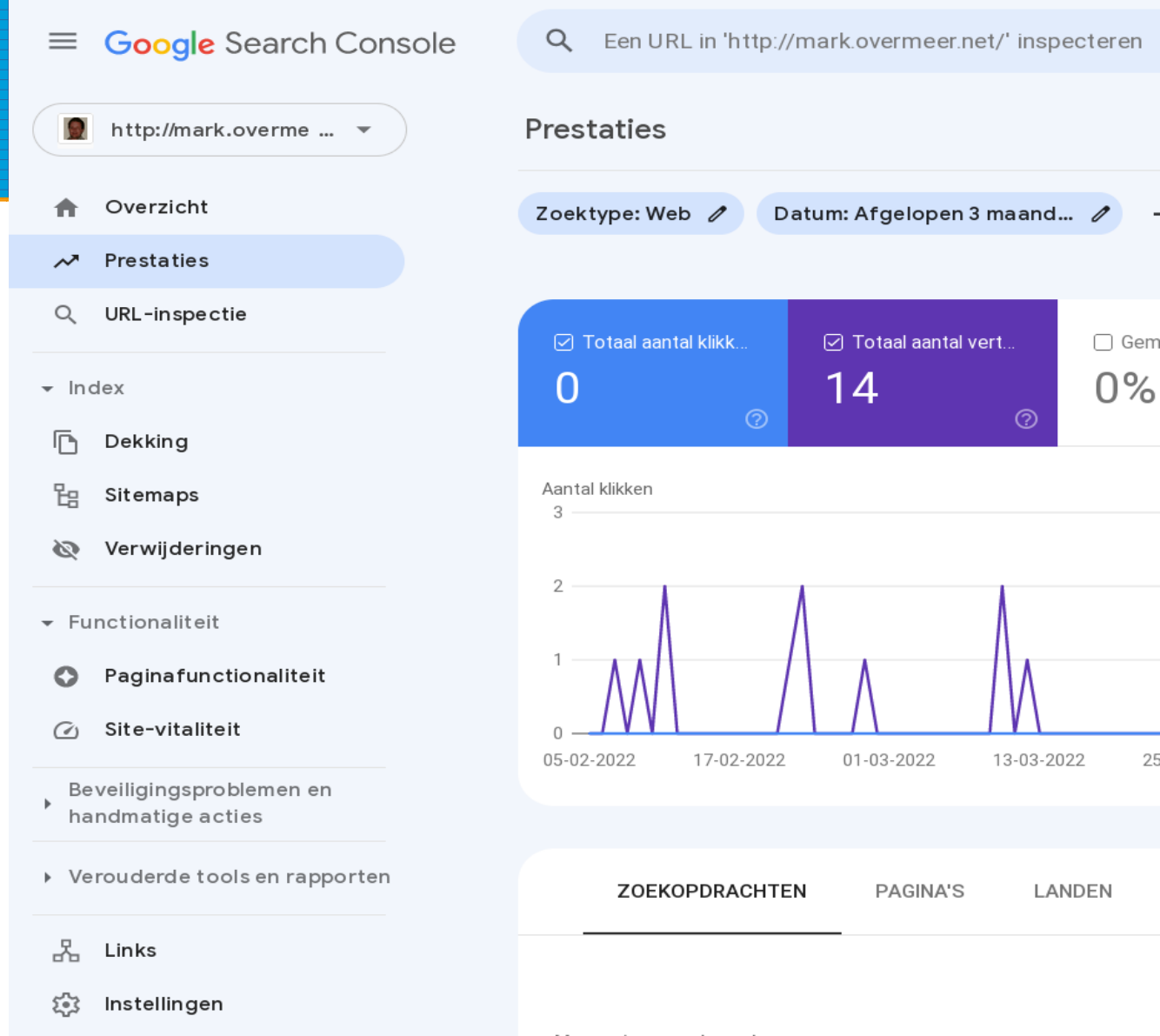
Major competative advantage



Open Console

~~Googles competative advantage~~

(EU) Legal challenges



"Console" purpose

- Improve the **quality** of the Web by sharing **knowledge** about **websites** and domain-names.

"Console" purpose

- Improve the **quality** of the Web by sharing **knowledge** about **websites** and domain-names.

- additional to/overlap with
  - sitemaps
  - robots.txt and security.txt
  - <meta> and <link rel>
  - schema.org, OpenGraph

Full Implementation

discovered
- new websites
- spam senders
- SEO networks
- meta-data
- phishing sites

Producers

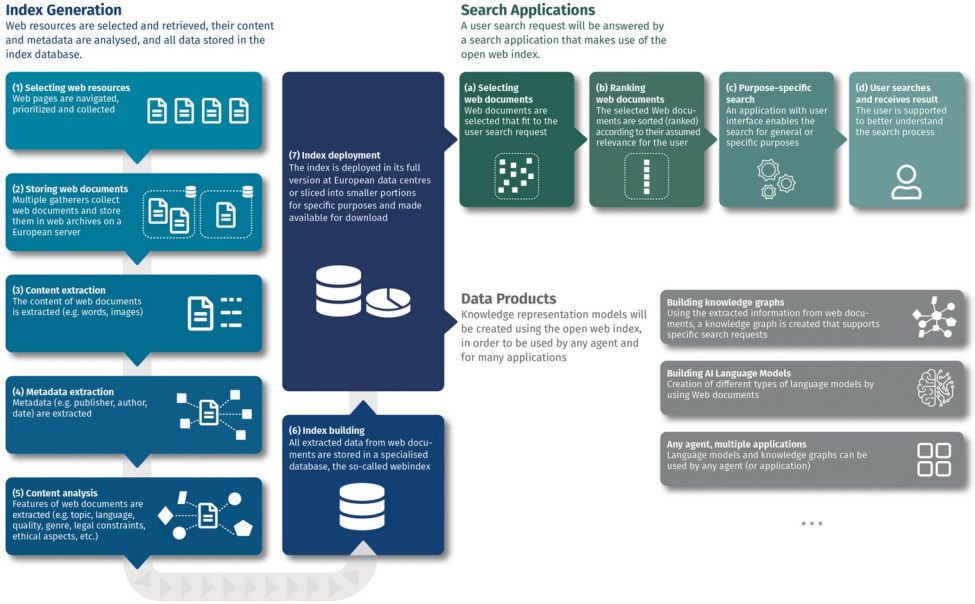
website facts

applications
- crawlers
- spam filters
- indexing
- research
- protection

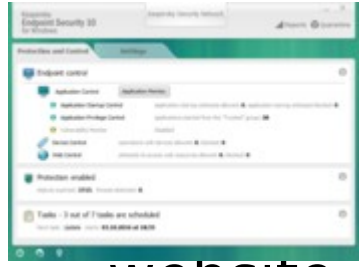
Consumers

website owner
domain owner
network owner

OpenWebSearch – Phase 1

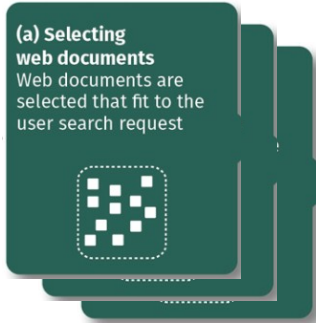


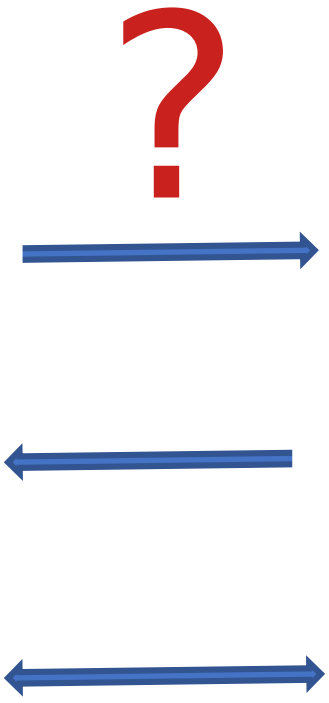
?

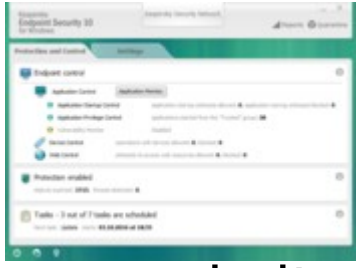


website owner

OpenWebSearch – Phase 1



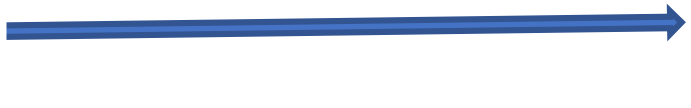


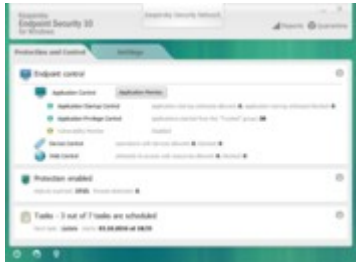


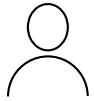
website owner

Phase 1

(1) Selecting web resources
Web pages are navigated, prioritized and collected

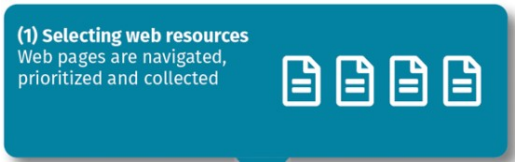




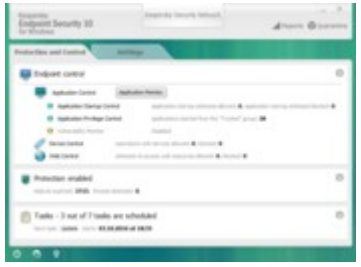


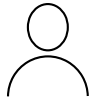
found broken link

Phase 1



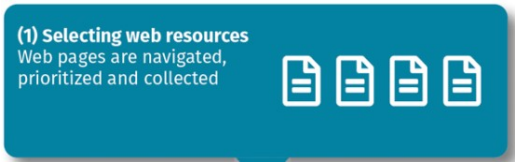
found broken link





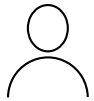
- push notification?
- display as web form
- verify correction
- silence issues

Phase 1 → 2



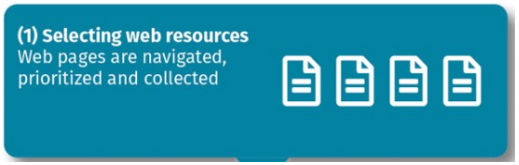
found broken link

- de-duplication
- re-report after some time
- implementation differences/bugs



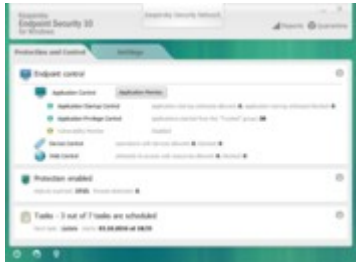
- push notification?
- display as web form
- verify correction
- silence issues

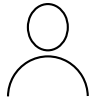
Phase 2+



*

- de-duplication
- re-report after some time
- implementation differences/bugs



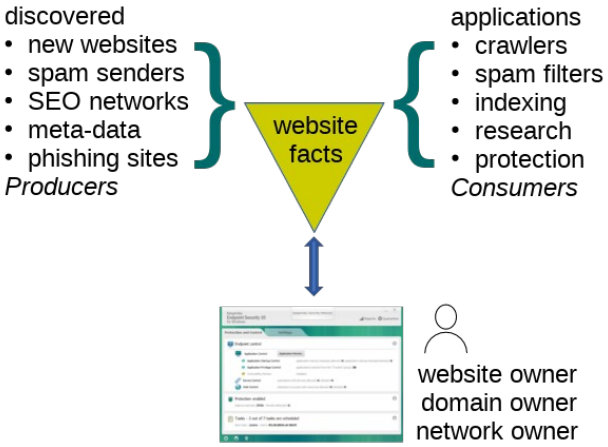


- push notification?
- display as web form
- verify correction
- silence issues

- **>1k sources**
- **>1M per day**
- **>300M websites**
- **>100M owners**

multi-purpose

- report black-listing
- 3rd party site descriptions
- tagging of explicit content
- ...



European Commission

DSA Transparency Database

Home | Search for statements of reasons | Analytics | Documentation | FAQ

Home

Welcome to the DSA Transparency Database!

Regulation (EU) 2022/2065 on a Single Market for Digital Services and amending Directive 2000/31/EC (**Digital Services Act**), or the 'DSA', is a **comprehensive set of new rules** regulating the responsibilities of digital services that act as intermediaries within the EU to connect consumers with goods, services and content. These intermediary services include online platforms, such as **online marketplaces** or social media networks, that store and disseminate information to the public at the request of the recipients of the service. You can find more information about the DSA on the dedicated Commission website.

Article 17 of the DSA obliges providers of hosting services to send clear and specific **statements of reasons** to any affected recipient when they remove or otherwise restrict availability of and access to information provided by the recipient. In other words, providers of hosting services need to inform their users of the content moderation decisions they take and explain the reasons behind those decisions.

To ensure transparency and to enable scrutiny over the content moderation decisions of the providers of online platforms and to monitor the spread of illegal and harmful content online, the **DSA Transparency Database** collects the **statements of reasons submitted by providers of online platforms** to the Commission, in accordance with Article 24(5) of the DSA.

The database, managed by the Directorate-General for Communications Networks, Content and Technology of the Commission is publicly accessible and machine-readable. Providers of online

OpenWebSearch – Phase 1

- For each component in

?

- Which information can **owners** contribute to OWS?
  - license, copyrights
  - crawling configuration
  - AI content declaration
  - contact information
  - (overrule) geographical details

OpenWebSearch – Phase 1

- For each component in

?

- Which information does **OWS** have for the website owners?
  - crawling reports
  - performance reports
  - intelligent extractions
    - location
    - spell checking
    - mobility score
    - schema.org / openGraph

Complications

- **Merge** (compatible?) data from 100's of sources.
- **Unified** presentation.

- **Large** infrastructure.
  - >300M active websites.
  - >100M website owners.

Complications

- Merge (compatible?) data from 100's of sources.
- Unified presentation.

- Large infrastructure.
  - >300M active websites.
  - >100M website owners.

- Fair competition to the user interface.
  Association?

No standard solutions!
(on the required scale)

Complications

- **Long term** development and maintenance (**costs**)

Complications

- **Building cooperation.**

  what useful information do you have to share?
  what data would improve your service?

  Who can we contact to discuss this?
  team@open-console.eu