

# Open Console

**Mark Overmeer MSc, MarkOv Solutions**  
**OSSYM 11 Oct 2022**

# Open Console

**Share what we now  
about  
websites / domains / networks**



# Open Console

- Entropy of text is far *too small* to pick the best search answers on Google-scale: additional **meta-data** is crucial.
  - Website-owners need to help.
  - Third parties need to help.

# Open Console

- Entropy of text is far too small to pick the best search answers on Google-scale: additional **meta-data** is crucial.
- Some knowledge on domain-names and websites is available, but hard to **share**:
  - SEO spam networks
  - List of phishing sites
  - Child-unfriendly filters
  - Email blacklists
  - Take-down orders

# Open Console

- Entropy of text is far too small to pick the best search answers on Google-scale: additional **meta-data** is crucial.
- Some knowledge on domain-names and websites is available, but hard to **share**.
- Website-owners have the right to correct users of their website content. Copyright owners have the **right to correct** websites.
  - Informed about take-down/copyright claims
  - Removal from black-list
  - Informed about (bad) review

# Open Console

- Entropy of text is far too small to pick the best search answers on Google-scale: additional **meta-data** is crucial.
- Some knowledge on domain-names and websites is available, but hard to **share**.
- Website-owners have the right to correct users of their website content. Copyright owners have the **right to correct** websites.
- Third parties can help improve the **quality** of websites.

# @Google

- Site ownership
- Crawl optimization
- Feedback

**Major competitive advantage**

The screenshot shows the Google Search Console interface for the URL `http://mark.overmeer.net/`. The left sidebar contains navigation options: Overzicht, Prestaties (selected), URL-inspectie, Index (with sub-items: Dekking, Sitemaps, Verwijderingen), Functionaliteit (with sub-items: Paginafunctionaliteit, Site-vitaliteit), Beveiligingsproblemen en handmatige acties, Verouderde tools en rapporten, Links, and Instellingen.

The main content area displays performance metrics for the selected URL. At the top, it shows the search type as 'Web' and the date range as 'Afgelopen 3 maand...'. Below this, there are three summary cards: 'Totaal aantal klikk...' with a value of 0, 'Totaal aantal vert...' with a value of 14, and 'Gem...' with a value of 0%. A line chart titled 'Aantal klikken' shows the number of clicks over time, with peaks around 05-02-2022, 17-02-2022, 01-03-2022, and 13-03-2022. At the bottom, there are tabs for 'ZOEKOPDRACHTEN', 'PAGINA'S', and 'LANDEN'.

# Open Console, Purpose

- One single platform for communication between
  - website-, domain-, and network owners
  - applications using website information, f.i. Crawlers
  - applications knowing facts about websites, f.i. phishing site detectors, geo-location, languages used, review, classification tags, ...
  - legal entities, f.i. copyright holders

“Open“ alternative to Google’s Search Console



# Open Console, mechanism

- *Providers* may **publish** facts about websites
- *Consumers* will **collect** facts about websites, to achieve a Task.
- Providers must show to the website (owner) what facts it shares, and offer a correction interface.
- Consumers interpret the facts: OC does not know. Hopefully, standardized facts.
- Website owners have a graphical interface (named Open Console), which is both Consumer as Provider.

# Open Console, complications

- Huge scale of fact distribution
- Ownership, maintenance of the facts
- Behavior of Consumers and Providers
- EU Privacy and data-processing law
- Seamless upgrades of interfaces and definitions
- Handling out-dated data structures
- ...

# Tech choice Forms

- Fight for the left column
- Permission hierarchy
- Input and output fields
- Uniform and pleasant styling
- 3rd party storage means outdated data stays alive

The screenshot displays the Google Search Console interface for the URL `http://mark.overme...`. The left sidebar contains navigation options: Overzicht, Prestaties (selected), URL-inspectie, Index (with sub-items: Dekking, Sitemaps, Verwijderingen), Functionaliteit (with sub-items: Paginafunctionaliteit, Site-vitaliteit), Beveiligingsproblemen en handmatige acties, Verouderde tools en rapporten, Links, and Instellingen.

The main content area shows performance metrics for the selected URL. The 'Prestaties' section includes filters for 'Zoektype: Web' and 'Datum: Afgelopen 3 maand...'. The metrics are:

- Totaal aantal klikt...: 0
- Totaal aantal vert...: 14
- Gem...: 0%

Below the metrics is a line chart titled 'Aantal klikken' showing the number of clicks over time. The x-axis represents dates from 05-02-2022 to 25-03-2022. The y-axis represents the number of clicks, ranging from 0 to 3. The chart shows several peaks, with the highest peak reaching 3 clicks.

At the bottom, there are tabs for 'ZOEKOPDRACHTEN', 'PAGINA'S', and 'LANDEN', with 'ZOEKOPDRACHTEN' currently selected.

# Fact Distribution

discovered

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

*Producers*



applications

- crawlers
- spam filters
- indexing
- research
- protection

*Consumers*



website owner  
domain owner  
network owner

# Fact Distribution

discovered

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

*Producers*



applications

- crawlers
- spam filters
- indexing
- research
- protection

*Consumers*

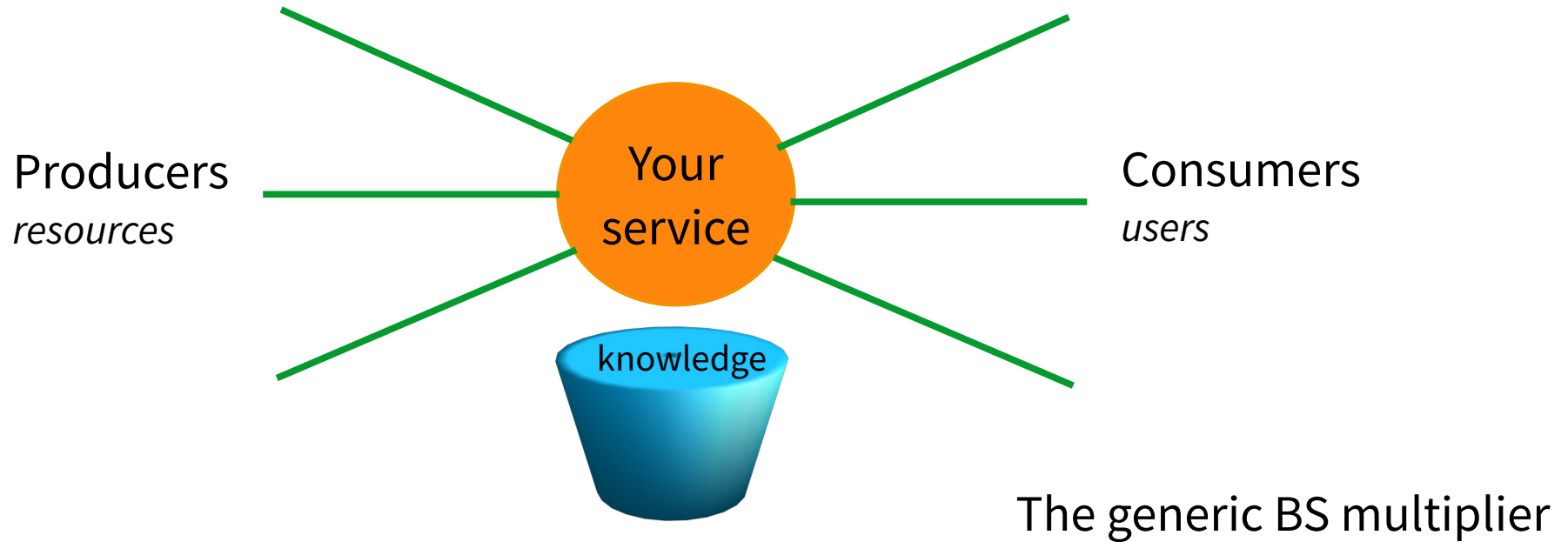


website owner  
domain owner  
network owner

- **Not in a central database:**
  - central storage too slow
  - legal status of facts
  - monopoly

# Research

# Fact Distribution



producers



consumers

Open Console requirement	FTP servers	(Shared network) filesystem, DB	Cloud	MicroServices (SOAP, REST)	
Huge amount of facts	++	++	~	--	
Small facts	-- (use zip, tar)	+	-- (zip, tar)	~	
Fine grained permissions	--	++	-	+	
Bulk behavior	++	++	-	--	
Life cycle management	-	+ (notifications)	?	--	
Redundancy	~ (mirrors)	~	++	~	
Service discovery	--	--	--	+ (WSDL, UDDI)	
Resource efficiency	++	++	-	--	

producers



consumers

Open Console requirement	FTP servers	(Shared network) directory, DB	Cloud	MicroServices (SOAP, REST)	Meshy Space
Huge amount of facts	++	++	~	--	++
Small facts	-- (use zip, tar)	+	-- (zip, tar)	~	++
Fine grained permissions	--	++	-	+	++
Bulk behavior	++	++	-	--	++
Life cycle management	-	+ (notifications)	?	--	++
Redundancy	~ (mirrors)	~	++	~	++
Service discovery	--	--	--	+ (WSDL, UDDI)	++
Resource efficiency	++	++	-	--	++



# Meshy Space

- Hide how we transport (FTP, HTTP, S3)
- Hide how it is stored (Zip, WARC, DB, files)
- 



- **Major disadvantage:**

# Meshy Space

- Hide how we transport (FTP, HTTP, S3)
- Hide how it is stored (Zip, WARC,

Generic solution for  
Data Sharing

**Think like a Librarian**

- **Major disadvantage:**



# Meshy Space

- Hide how we transport (FTP, HTTP, S3)
- Hide how it is stored (Zip, WARC, DB, files)
- Facts are in *Collections*
  - Every Collection maintains *Units*
  - Every Unit has *payload* and life-cycle management
  - Sub-collections, sub-setting, merging, caching etc on Collections
- Sliding migrations: no coordinated upgrades required
- **Major disadvantage: ...**



# Meshy Space

- Hide how we transport (FTP, HTTP, S3)
- Hide how it is stored (Zip, WARC, DB, files)
- Facts are in *Collections*
  - Every Collection maintains *Units*
  - Every Unit has *payload* and life-cycle management
  - Sub-collections, sub-setting, merging, caching etc on Collections
- Sliding migrations: no coordinated upgrades required
- **Major disadvantage: ...**



# Meshy Space

- Hide how we transport (FTP, HTTP, S3)
- Hide how it is stored (Zip, WARC, DB, files)
- Facts are in *Collections*
  - Every Collection maintains *Units*
  - Every Unit has *payload* and life-cycle management
  - Sub-collections, sub-setting, merging, caching etc on Collections
- Sliding migrations: no coordinated upgrades required
- **Major disadvantage: !! ME !!**



# Open Console, complications

- Huge scale of fact distribution
- Ownership of the facts
- Behavior of Consumers and Providers
- EU Privacy and data-processing law
- Seamless upgrades of interfaces and definitions
- Handling out-dated data structures
- ...

*Meshy Space*

*keep distributed*



*Meshy Space*

*Forms*