Open Internet Search – a paradigm shift in and for Europe

Why and how Europe should build an open Internet search Infrastructure

Web-meeting on Open Internet Search, CERN, 26 May 2020
Dr. Stefan Voigt, open search foundation
Just as we can orientate freely in the physical sphere we also have to be able orientate freely ...

... in the digital sphere!
We are meeting at the cradle of the web

Tim Berners-Lee said 2019:
„If we give up on building a better web now, then the web will not have failed us. We will have failed the web.“

„The web is broken: Three sources of dysfunction affecting today’s web
• Deliberate, malicious intent => laws and code to minimize this
• System design that creates perverse incentives => redesign systems in a way that change incentives
• Unintended negative consequences of benevolent design => research to understand existing systems and model possible new ones or tweak those we already have

To fix it, we need: laws and regulations for the digital age that ensure markets remain competitive, innovative and open.
• laws and regulations for the digital age that ensure markets remain competitive, innovative and open.
• Platforms and products designed with privacy, diversity and security in mind.”
The Open Search Foundation

“" We raise awareness for the need of a neutral and open web search infrastructure for Europe.

“" We inform people: the general public currently knows too little about the search problem.

“" We encourage experts to research and cooperate on solutions across different disciplines.
The Open Search Foundation takes on three roles.

**Catalyst**
“We educate, make people reflect and initiate change.”

**Ideator**
“We create and suggest new approaches for open search. We encourage and inspire others to take action.”

**Connector**
"We bring together major players from different disciplines and build a solid network."
Who are currently the gatekeepers of this information?

- USA: 86.4%
  - Google: 75%
  - Bing: 5.1%
  - Yahoo!: 1%
  - Others: 7.5%

- Europe: 91.4%
  - Google: 37%
  - Bing: 2.4%
  - Yandex: 2.5%
  - Others: 3.7%

- China: 73.4%
  - Baidu: 4.1%
  - Shenma: 7.5%
  - Haosu: 15%
  - Others: 4.1%

Source: Statista (2019)
Four global web-indices. None in Europe.
As Europeans we need to regain autonomy and sovereignty in navigating and operating the digital sphere.

We need to work on concepts for a cooperative, open and distributed Internet search infrastructure in and for Europe.
An Open Web Data Pool and Web Index, as a fundamental and indispensable basis for a large variety of public and private information services.

→ Europe's „Galileo“ Programme for Internet Autonomy
We need to connect our data spaces, computing spaces and web spaces... for setting up a joint Open Web Data Pool and Web Index in Europe.

- User
- User
- User
- User
- User
- User
- User
- User
- User
- User
- User
- User
- User
- User
- User

Service 1
Service 2
Service 3
Service n

Open Search Infrastructure

Standardized Interface / API

Distributed Open Web Data Pool and Web Index

Standardized and cooperative Crawlers gather information and aggregate it

Information available on the Internet

- CERN
- ESA
- ECMWF
- JRC
- Univ.
- .eu
- .de
- .fr
- .it
- .com
- ...
European Open Search Infrastructure:

Open Search Infrastructure

Standardized Interface / API

Distributed Open Web Data Pool and Web Index

Standardized and cooperative Crawlers gather information and aggregate it

Web Data Pool

Index

<table>
<thead>
<tr>
<th>URL</th>
<th>feature</th>
<th>key</th>
<th>key</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ontology

Graph Data Base

Image Credits:
https://medium.com/neo4j/hands-on-graph-data-visualization-bd1f055a492d
We suggest three basic principles to build such an open search ecosystem in and for Europe.

1. Collaborative, distributed computing and hosting
2. Open source
3. Public moderation

Europe inside!
=> ethics, culture, languages, jurisdiction, economy, ….
Where do we stand?

✶ **Ideas and approaches** are developing.
✶ First **tests and trials** have been done.
✶ An **interdisciplinary community** is shaping.
✶ **Politics** are getting more and more interested.
What are the next steps?

- We need to further research and work on ‘open search’.
- We should start building key parts and examples.
- We need to raise funding for the necessary R&D projects.
- We need to further grow the interdisciplinary (!) community.
- We have to gain support and active engagement by politics and administration in Member States and European Commission.
How can I help to get this going?

✶ **Spread the word** among colleagues, your organisation and your peers to further build the network.

✶ **Contribute** with your research, master/PhD projects, studies, colloquia, workshops etc. – **speak to politicians**!

✶ Contribute with papers, presentations, suggestions for sessions to the **OSSYM 2020** hosted by CERN on 12-14 October 2020.

✶ Contribute to the **osf working groups**:
  tech, awareness, economy, ethics and legal...

... to do so, contact us at: yourinfo@opensearchfoundation.org
70 years ago, 5 years after WW II, Robert Schuhmann, the French Minister of foreign affairs, said:

"World peace cannot be safeguarded without the making of creative efforts proportionate to the dangers which threaten it.

Today, with regards to the challenges Europe faces in the digital era we may translate this to:

"Open and fair access to Internet information and markets cannot be safeguarded without the making of creative efforts proportionate to the dangers which threaten it."
It’s not about money, it’s about cooperation!