Towards better search for trustworthy content and object discovery

Through tools and concepts in Slovenian Interoperability Framework

Aleš Černivec, Project Manager, XLAB d.o.o.
XLAB
Software Development & Research

- Founded 2001
- Employees 100+
- 54 Research Projects
- 6 products & 8 start ups

2018 Among top 3 SMEs in Slovenia in terms of H2020 signed grants
What I do?

• National projects
  • NIO – national interoperability portal
  • OPSI – Open data portal

• H2020 Research projects related to security and privacy
  • FORTIKA, CyberWISER

• CISO (taking care of ISO controls for the company)
Building a national interoperability framework

- OpenData portal (podatki.gov.si)
- Portal NIO (nio.gov.si)
Why?

• Legislation on openness of public sector’s data, information and solutions

• Advantages
  • allows citizens a look into the operation increasing trust and revision
  • it makes possible the development of new solutions based on the data
  • The institutions producing the data become more aware of their responsibility --> improvement in data quality, quantity and management
  • It provides a single point of access for easier search, analysis and linking of the data
Tools/OPSI

- CKAN API and plugins allow simple harvesting between instances, making possible even an intra-national hierarchy of portals
- Provides a simple interface for searching and filtering through published datasets
- The resources are regularly checked for broken links and scored based on the openness of their data format. Previews of some common data formats are also possible.
- Within XLAB we have extended the features to support:
  - Standardized publishing workflow
  - CAS and SAML2 authT and authZ schemes
  - CSV harvesting capability from intra-national portals
  - Keywords suggestions from Eurovoc and Gemet tesauri
  - Tracking via Piwik instance (GDPR compliance)
Tools/NIO

• Rewritten from monolithic architecture to microservice‘s architecture

• **Complete** search of the CMS‘ content and data content (full text) using ElasticSearch, Swift (block storage)

• **Standardized** workflow for publishing, ISO compliant workflows for editors/reviewers

• Allowing export of the data in standard formats (ADMS-AP)

• Extendible into **Knowledge Management Service** (KMS), allowing different modules for existing cloud-based storages

• Usage of standard authT and authZ schemes, e.g. SAML2, OAuth2, usage of JWT for authZ
Have you noticed words in bold?

• Human values: openness, trust, security, revision (transparency), responsibility,

• Processes demanding knowledge: search, analysis, linking, (knowledge) management

• Technology allows us to manage the knowledge easier and faster, produce new content faster

• Human values can now be powered by the technology and Next Generation Internet should make knowledge management easier, more dynamic, discoverable
Expected output of this workshop

• Human values and challenges
• Potential solutions to the challenges
• Are there any known initiatives tackling the challenges?
• What are the gaps
• What R&I is needed to resolve the gaps
Expected output of this workshop

• What tools and technology used for providing the values of the similar portals/services?
• Where are the gaps?
• Are you aware of some initiatives related to the search & discovery of trustworthy content and objects discovery?
• What is the next logical step? (e.g. mashups through different run-time environments (cloud or self-hosted); NLP, AI/ML)
Thank you!

• Questions?
Company info

XLAB d.o.o.
Pot za Brdom 100
SI-1000 Ljubljana

P: +386 1 244 77 50
E: ales.cernivec @xlab.si
W: http://www.xlab.si