INSPIRE in the cloud and GDPR

a spatial proxy solution

T. Chevallier
(AKKA Technologies)
demonstrate the feasibility of employing a cloud-based infrastructure to provide INSPIREd services.

develop a secure framework for storing and processing data outsourced to the cloud so end-users can control their data.

an « INSPIREd » journey to the land of Data Protection.

INSPIRE in the cloud and GDPR: a spatial proxy solution.
GDPR
General Data Protection Regulation

the most important change in data privacy regulation in 20 years

entry into force
25 May 2018
A brief history

Data Protection Directive 95/46/CE

- 1995
- 1996
- 1997
- 1998
- 1999
- 2000
- 2001
- 2002
- 2003
- 2004
- 2005
- 2006
- 2007
- 2008
- 2009
- 2010
- 2011
- 2012
- 2013
- 2014
- 2015
- 2016
- 2017
- 2018

SEARCH ENGINES
WEBMAIL
INSTANT MESSAGING
PEER-TO-PEER FILE SHARING
ONLINE ADVERTISING
ONLINE SOCIAL NETWORKING
CLOUD STORAGE
ON-DEMAND CLOUD COMPUTING PLATFORMS
CROWDSOURCING
WEB 2.0 « SOCIAL WEB »
MOBILE WEB
INTERNET OF THINGS
WEB 3.0 FUTURE WEB

GENERAL PUBLIC GPS NAVIGATION
TILED WEB MAPS
STREET VIEW
LOCATION BASED SERVICES
SEMANTIC WEB
« BIG DATA »

« Safe Harbor »

Commission proposal for a new regulation

GDPR entry into force

« Safe Harbor » invalidated

EU-US Privacy Shield

INSPIRE in the cloud and GDPR: a spatial proxy solution
Key requirements

- Organisations must maintain a « Data Protection Officer » (DPO)
- Organisations must be able to demonstrate compliance at all time
- Data subjects have extended rights (right to be forgotten, data portability, information about data breach)
- Organisations must keep accurate records on data processing activities
- There must be a legal basis for processing personal data of EU citizens
- There are restrictions on the types of data that can be processed (sensitive personal data)
- Gaining consent from data subject is more stringent than it was before
What does it change?

- A regulation instead of a directive: a unique text *directly applicable to all Member States*
- **Cultural change**: from a logic of control (administrative formalities) to a logic of accountability (organisations must ensure compliance)
- Very significantly *increases the penalties* for violating the rules (fines up to 4% annual global turnover or 20 M€)
- **Privacy by design** and privacy by default (e.g. pseudonymise data as soon as nominative identification is not absolutely necessary to fulfill the need)
- Appointing a **Data Protection Officer (DPO)** is now *mandatory* for public authorities and companies processing personal data on a regular basis
What to do, in practice?

In order to comply with the regulation, an organisation should:

1. appoint a Data Protection Officer (DPO)
2. define the policies and processes regarding personal data protection
3. then use the technology which will be able to support these policies
GDPR and the cloud

known issues

<table>
<thead>
<tr>
<th>« Shadow IT »</th>
<th>Data location issues</th>
<th>« Honest-but-Curious » model</th>
</tr>
</thead>
</table>
| knowing where are all the employees data is hard to achieve in the organisation, so ensuring permanent compliance with the GDPR is difficult. | Cloud service providers (CSP) generally have data centers in different countries. Most of these countries have laws to protect privacy, but these laws can vary between countries (e.g. inside and outside the EU). | According to this threat model, cloud service providers (CSP) are:  
- « Honest »: the CSP will never consciously deviate from the protocols or « attack » user’s data  
- But « Curious »: the CSP will exploit all available data and infer private information (for e.g. profiling, targeted advertisement, etc.) |

GDPR compliance is difficult in the context of cloud computing
The European Commission invests in funding for research and innovation on data protection in the cloud.

**Horizon 2020**

**TOPIC: Advanced Cloud Infrastructures and services**

« Significantly higher quality of user experience and trust in clouds through **stronger security and data protection** »
## Related innovative technologies

### Cloud Encryption Gateway

A cloud security proxy which performs encryption or tokenization on an item-by-item basis as data flows through the proxy. The obfuscated data can then be stored in the cloud.

Cloud encryption gateways typically provide a choice of various encryption and tokenization algorithms, depending on the strength of protection required and how much format preservation is necessary.

### Cloud Access Security Broker

On-premises or cloud-based security policy enforcement points, placed between cloud service consumers and cloud service providers to interject enterprise security policies as the cloud-based resources are accessed.

CASBs consolidate multiple types of security policy enforcement. Example security policies include authentication, single sign-on, authorization, credential mapping, device profiling, encryption, tokenization, logging, alerting, malware detection/prevention and so on.

Source: Gartner IT Glossary
INSPIRE in the cloud and GDPR: a spatial proxy solution

Cloud Service Provider

Honest-but-Curious

UNTRUSTED ZONE

TRUSTED ZONE

Proxy

User $u$

1. Dataset
2. Search query
3. Clear results
4. Transformed search
5. Obscured results
6. Dataset
Security policy (example)

What to protect

<data>
  <attribute
    name="patient_gender"
    attribute_type="quasi_identifier"
    data_type="categoric">
  </attribute>
  <attribute
    name="patient_name"
    attribute_type="identifier"
    data_type="categoric">
  </attribute>
  <attribute
    name="episode_id"
    attribute_type="non_confidential"
    data_type="categoric_ordinal">
  </attribute>
  <attribute
    name="episode_age"
    attribute_type="quasi_identifier"
    data_type="numeric_discrete">
  </attribute>
  <attribute
    name="diagnose_id"
    attribute_type="confidential"
    data_type="categoric">
  </attribute>
</data>

How to protect

<protection
  module="anonymization">
  <attribute_types>
    <attribute_type
      type="identifier"
      protection="supression">
    </attribute_type>
    <attribute_type
      type="quasi_identifier"
      protection="k-anonymity"
      k="5">
    </attribute_type>
    <attribute_type
      type="non_confidential"
      protection="null">
    </attribute_type>
  </attribute_types>
</protection>

What to protect

<protection
  module="anonymization">
  <attribute_types>
    <attribute_type
      type="identifier"
      protection="supression">
    </attribute_type>
    <attribute_type
      type="quasi_identifier"
      protection="k-anonymity"
      k="5">
    </attribute_type>
    <attribute_type
      type="non_confidential"
      protection="null">
    </attribute_type>
  </attribute_types>
</protection>
Data Protection and INSPIRE

- personal or non-personal
- sensitive or non-sensitive
- confidential or non-confidential
- geographical or non-geographical

which Data?
The INSPIRE directive (reminder)

Mandatory to publish data
Within the scope of 34 themes as described in 3 annexes

With exceptions
When the access to spatial datasets and services would adversely affect:

- public security, national defence or the course of justice
- the confidentiality of commercial or industrial information
- the confidentiality of personal data relating to a natural person

but ... what is personal data?
But... what is « personal data »?

« Personal Data »
any information relating to an identified or identifiable natural person ("data subject »)

an identifiable person is one who can be identified, directly or indirectly, in particular by reference to :
- an identifier (such as a name, an identification number, location data, online identifier),
- one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person

« Sensitive Personal Data »
personal data, revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership; data concerning health or sex life and sexual orientation; genetic data or biometric data.

Data relating to criminal offences and convictions are addressed separately (as criminal law lies outside the EU's legislative competence).
ACCESS DATA

information on **people** who have **accessed, read or downloaded** environmental data or who have **used a particular service** in relation to environmental information

Organisations managing these personal informations, and particularly public authorities, are requested to protect them (data breach notification, right to be forgotten, etc.)
« Personal data » in the context of INSPIRE (geographical)

Several presentations about Data Privacy during past INSPIRE conferences

aerial and satellite imagery

Personal !

depending on the product resolution

Aalborg 2014
Lisbon 2015
Barcelona 2016
Do you have any legal/security barriers relating to your use of geodata? If yes, which ones?

Result of a questionnaire to a panel of geological surveys, conducted by the EGDI scope project.
The CLARUS « Geo-Proxy »

- Protocol parsing
- Sensitive data identification
- Data protection

Clear data → Proxy → Protected data

---

INSPIRE in the cloud and GDPR: a spatial proxy solution
One technology, different applications

- Developing a PostGIS/WFS proxy is a complex task.
- However, the possible applications are wide and numerous.
  - e.g. Metadata creation proxy
  - e.g. Semantic enhancement proxy
  - e.g. Authentication proxy
  - etc.
Conclusions

- The GDPR is a « quantum leap » in the field of Data Protection
- The requirements for personal data protection are stricter and the fines heavier
- Privacy by design will be the rule
- Data Protection Officer accountable for a GDPR-compliant Data Protection policy

New rules

- In order to comply with the GDPR, new data protection technologies are needed (particularly in the context of cloud solutions)
- These technologies should be able to protect geo-referenced data (as the notion of personal data encompasses location data)

New technologies

- Not so many INSPIRE themes deal with « personal data »
- But INSPIRE actors often deal with « confidential data » (for e.g. national security or commercial reasons)
- Technological advances in the field of personal data protection can be applied to wider purposes (confidential data protection, semantic enhancement, etc.)

New applications
THANK YOU

www.clarussecure.eu | contact@clarussecure.eu | @Clarussecure