Towards a Semantic Specification for GDPR Data Breach Reporting

Harshvardhan J.Pandit[12] , Paul Ryan[125], Georg Philip Krog[4], Martin Crane[2], Rob Brennan[13]

1. ADAPT SFI Research Centre,
2. School of Computing, Dublin City University
3. University College Dublin, Ireland,
4. Signatu AS, Oslo, Norway
5. Uniphar PLC, Dublin, Ireland

Contact : Paul.Ryan76@mail.dcu.ie

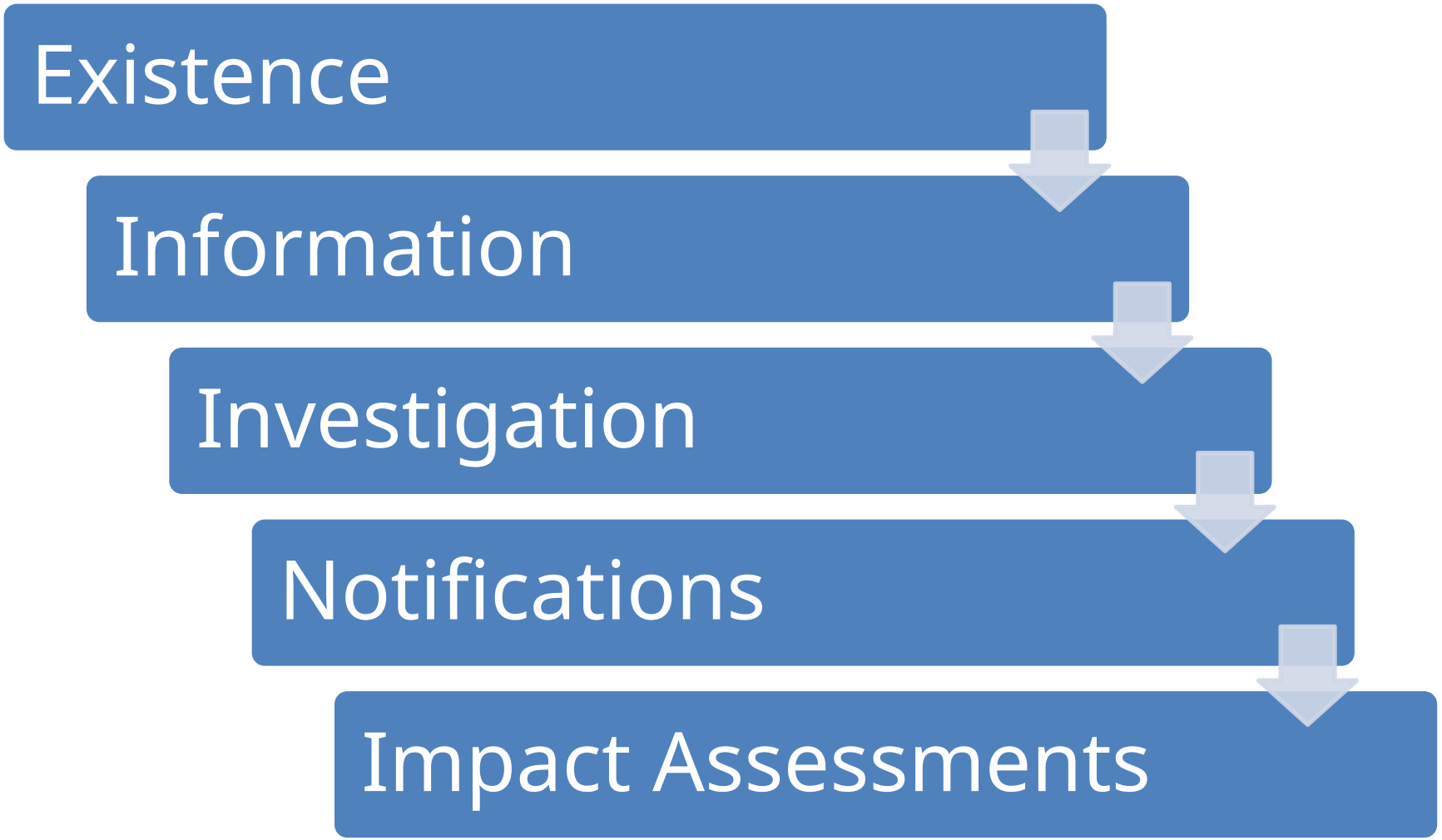
The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

What is a Data Breach ?

A breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed.” GDPR Article 4-12

An emerging challenge in the digital era!

Reporting Obligations - The Data Breach Life Cycle



Motivation

- Impact assessment can be challenging to complete
- Notifications – factual communication to multiple stakeholders
- Regulators reports require supplementary information
- Conventional approaches represent a barrier to effective compliance
- Increased CyberSecurity reporting demands *

* Network and Information Security Directive (NIS2, 2023) and Digital Operational Resilience Act (DORA, 2023)

Our Objective

To define information regarding data breaches in a machine-readable form (by using semantic web standards) to address these challenges and enable the development of interoperable tools for assisting organisations and authorities in their data breach requirements.

Our Approach

- Identify Breach concepts found in the GDPR
- Utilise existing concepts in Data Privacy Vocabulary (DPV)
- Utilise concepts found in other vocabularies
- Add new concepts to DPV
- Utilise Legal and Data Protection expertise of Data Privacy Vocabulary Community Group
- Model the Breach Reporting requirements

What is the Data Privacy Vocabulary ?

- The DPV is a vocabulary (terms) and an ontology (relationships) serialised using semantic-web standards to represent concepts associated with privacy and data protection, primarily derived from GDPR

- A community specification through the W3C Data Privacy Vocabulary and Controls Community Group (DPVCG).

- A machine-readable representation of personal data processing and can be adopted in relevant use-cases such as legal compliance documentation and evaluation, policy specification, consent representation and requests, taxonomy of legal terms, and annotation of text and data.

- Links to DPV and community group
- https://w3.org/ns/dpv
- https://www.w3.org/community/dpvcg/

Describing a Data Breach

Record keeping requirement

Describing a Breach

```
ex:Incident1A a dpv-breach:DataBreach ;
   rdf:type dpv-breach:ConfidentialityBreach ; # type of breach
   dct:temporal dpv-breach:Unknown ; # start and end are unknown
   dpv:hasRiskSource dpv-breach:Unknown ; # what caused the
breach
   dpv:hasThreatActor dpv-breach:Unknown ; # who caused the
breach
   dpv:hasStatus dpv-breach:DataBreachOngoing . # status of breach
```

Data Breach Detection report

```
ex:IncidentReport2A a dpv-breach:DataBreachDetectionReport ;
   dct:subject ex:Incident1A ; # which data breach this report refers to
   dct:created "2023-05-26T14:38:00" ; # when this report was created
   dct:creator ex:CompanyAlpha ; # who created the report
   dpv:hasDataSource ex:Employee ; # breach was reported by an
employee
   dpv:hasDataSource <https://nytimes.com> ; # breach was reported in a
news
   dpv:hasDataSource ex:Processor ; # breach was reported by a Processor
   dpv:hasActivityStatus dpv:ActivityCompleted . # status of the detection
reporting
```

Data Breach Impact Assessment (DBIA)

DBIA requirement	Content
Information about the Breach	Type of breach; Nature, sensitivity, and volume of personal data; Special characteristics of the individual; Special characteristics of the data controller; Number of affected individuals
Risk Assessment	risks and impacts to rights and freedoms, risk levels, likelihoods, severity of consequences, and specific risks (e.g. ease of identification)
Outcomes:	activities to be undertaken based on the assessment

Describing a Data Breach Impact Assessment

```
ex:DBIA20230628 a data-breach:DataBreachImpactAssessment ;
   dct:title "DBIA for Incident2023-2"@en ;
   # annotations
   dct:subject ex:Incident2023-2 ; # reference to data breach
   dct:creator "Anon. Anon." ; # authorship or contributors
   # temporal
   dct:created "2023-06-28"^^xsd:date ; # creation date
   dct:created "2023-06-28"^^xsd:date ; # last modification date
   dct:dateSubmitted "2023-06-28"^^xsd:date ; # submission for
approval
   dct:dateAccepted "2023-06-28"^^xsd:date ; # approval date
   # versions
   dct:isReplacedBy ex:DBIA20230629 ; # next version
   dct:replaces ex:DBIA20230627 . # previous version
```

Notifications

Multiple stakeholders may require notifications

Processor describing a Data Breach event to a Controller

- ex:ProcessorReportsBreach a dpv-breach:ProcessorDataBreachNotice, schema:Message ;
- rdfs:comment "Processor to Controller" ;
- dct:subject ex:Incident1A ;
- schema:dateReceived "2023-05-24" ; # when the message was sent
- schema:sender ex:Processor ; # who sent it
- schema:recipient ex:Controller ; # who received it
- schema:messageAttachment <report.pdf> . # what were the contents

Controller describing a Data Breach event to an Authority

```
ex:ControllerReportsBreach a dpv-breach:AuthorityDataBreachNotice, schema:Message ;
   rdfs:comment "Controller to Authority" ;
   dct:subject ex:Incident1A ;
   schema:dateSent "2023-05-24" ; # when the message was sent
   schema:sender ex:Controller ; # who sent it
   schema:recipient ex:DPA-IE ; # who received it
   schema:messageAttachment <report.pdf> . # what were the contents
```

Conclusions

- A semantic specification for representing information about data breaches based on the requirements of the GDPR
- A machine-readable vocabulary to represent critical information regarding the data breach event, how it was detected, the consequent analysis of its impact on systems, data, and data subjects, and its communication to other entities
- to enable efficient tools and processes to handle obligations regarding data breaches, and to be interoperable with other security vocabularies.

Future Work and Directions

- Creating communal 'knowledge graphs' to identify relevant risks and impacts similar to existing security initiatives such as MITRE and VERIS
- Data breaches are also security incidents; ontology is extensible to other regulatory requirements such as NIS2 and DORA.

Contact: Paul.Ryan76@mail.dcu.ie

The VERIS Framework
CVE - Common Vulnerabilities and Exposures;. Available from: https://cve.mitre.org/
Network and Information Security Directive (NIS2, 2023)
Digital Operational Resilience Act (DORA, 2023)

Engaging Content
Engaging People

Towards a Semantic Specification for GDPR Data Breach Reporting

Thank you

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.