Information Security and Data Controllability for Collaborative Systems

Manuel MUNIER
Vincent LALANNE

Universite de Pau et des Pays de l'Adour, E2S UPPA, LIUPPA, EA 3000

WorldCIS 2019
Manuel Munier
Vincent Lalanne

December 9-11, 2019 / London, UK

Motivation

- **Information is the most important asset for companies**
  - Business data warehouse
  - Informational capital
  - Business intelligence
  - Decision making systems

Motivation

- **Information is the most important asset for companies**
  - Business data warehouse
  - Informational capital
  - Business intelligence
  - Decision making systems

- **Security « toolbox »**
  - Hardware : computer, storage, energy,…
  - Software : system management (OS & services), software engineering, model driven engineering,…
  - Communications : network architecture, protocols, cryptography,…
  - Security criteria : confidentiality, integrity, availability
  - Risk management (information security) : EBIOS / CRAMM / OCTAVE methods, ISO 27xxx standard

- **Note:** we should often talk about **data** rather than **information**
  - The temperature is 29°C, the weather is sunny → **data**
  - It's a nice day (IMHO) → **information**

Motivation

Where is the problem ?

- Ok for « closed » IS → the CISO[1] controls all elements of the company's information system
- But today's systems are connected to each other and exchange a lot of information
  - We import information from many sources
  - We export our information without really knowing who will use it or how
- Question : what control do we have on this information ?
  - « Quality » of information that we inject into our IS ? (incoming flows)
  - « Use » of information that our IS provides ? (outgoing flows)

Application areas

- IoT, smart buildings, smart cities, smart grids,...
- Interconnection of medical data systems
- Distributed e-learning systems
- Big data processing
- ...

[1] CISO : Chief Information Security Officer

Existing works

- **Data Quality (DQ)**
  - Works date a little (1990s)

  - A classification that can be useful to us to group DQ dimensions into categories :
    - Intrinsic DQ → accuracy, objectivity, believability, reputation
    - Accessibility DQ → accessibility, access security
    - Contextual DQ → relevancy, value-added, timeliness, completeness, amount of data
    - Representational DQ → interpretability, ease of understanding, concise repr., consistent repr.

Existing works

- **Data Quality (DQ)**
  - Works date a little (1990s)

  - A classification that can be useful to us to group DQ dimensions into categories :
    - Intrinsic DQ → accuracy, objectivity, believability, reputation
    - Accessibility DQ → accessibility, access security
    - Contextual DQ → relevancy, value-added, timeliness, completeness, amount of data
    - Representational DQ → interpretability, ease of understanding, concise repr., consistent repr.

- **Provenance, lineage, traceability**
  - Provenance is information that describes the origin or experience of a data
  - Lineage is a flow of how data will move and transform between systems, tables, data domains
  - Traceability provides metadata to track information

Existing works

Trust & reputation

- Trust is the extend to wich a party is prepared to depend on something or someone in a given situation with a sense of relative security, even if negative consequences are possible
- Reputation is what’s usually believed a few person’s or thing’s character or standing
- Trust ≠ reputation → can be illustrated by the following perfectly normal and plausible statements :
  - « *I trust you because of your good reputation* »
  - « *I trust you despite your bad reputation* »

Existing works

Trust & reputation

- Trust is the extend to wich a party is prepared to depend on something or someone in a given situation with a sense of relative security, even if negative consequences are possible
- Reputation is what's usually believed a few person's or thing's character or standing

- Trust ≠ reputation → can be illustrated by the following perfectly normal and plausible statements :
  - « *I trust you because of your good reputation* »
  - « *I trust you despite your bad reputation* »

Contracts, Service Level Agreements

- SLA = agreed document to establish requirements about the quality of a service → QoS
- SLA often refers to the performance and security properties of the service

- Contracts are a more legal vision : they set the rights and duties of each party, the penalties, the evidence requested,…
- Contracts could be a way to set up a usage control policy (for example…)

Our work

- These works are different facets of **« data controllability »**

- Our opinion is that data controllability should now appear in information security risk analysis in the same way as confidentiality, integrity, availability

- **Our approach**
  - Follow a risk management process (eg EBIOS or ISO 27005)
  - Define a new security criterion → (data) controllability
  - Provide ideas for tools to support this new criterion

Risk management process

Scale of needs

- To be able to assess the risks (associated with a given criteria) → definition of a scale of needs
- Example :

1) **Limited** → the quality or reliability of the information is not necessary
fuzzy reasoning, artificial intelligence

2) **Weak** → it is possible to accomodate some erroneous and/or unreliable information without impacting the processes
big data, statistical processes

3) **Strong** → the quality of the information is essential for the proper functioning of the processes
customer orders, software development, smart grids, smart & intelligent systems

4) **Absolute** → uncontrolled information can cause damage to the sustainability of the business
medical data, nuclear power plant

Risk management process

- **Risk identification**
  - Risk = **Threat** × Vulnerability × Impact
  - Main steps :
    - Identification & valuation of assets (context)
    - Identification of **threats**

Risk management process

- **Risk identification**
  - Risk = Threat × **Vulnerability** × Impact
  - Main steps :
    - Identification & valuation of assets (context)
    - Identification of threats
    - Identification of **vulnerabilities** (taking into account existing controls)

Risk management process

- **Risk identification**
  - Risk = Threat × Vulnerability × **Impact**
  - Main steps :
    - Identification & valuation of assets (context)
    - Identification of threats
    - Identification of vulnerabilities (taking into account existing controls)
    - Identification of **consequences**

Risk management process

- **Risk identification**
  - Risk = **Threat × Vulnerability × Impact**
  - Main steps :
    - Identification & valuation of assets (context)
    - Identification of threats
    - Identification of vulnerabilities (taking into account existing controls)
    - Identification of consequences

A threat has the potential to exploit one or more vulnerabilities to harm assets and thereby impact the organization (with regards to security criteria).

Risk management process

A threat has the potential to exploit one or more vulnerabilities to harm assets and thereby impact the organization (with regards to security criteria).

Risk identification

- Risk = Threat × Vulnerability × Impact
- Main steps :
  - Identification & valuation of assets (context)
  - Identification of threats
  - Identification of vulnerabilities (taking into account existing controls)
  - Identification of consequences
- Output → list of incidents scenarios
  - Ex : « *Traceability of information from external services is not provided. As a result, the reliability and/or quality of the information can not be guaranteed. The company therefore incurs the risk of generating and disseminating information calculated on the basis of unreliable data, and thus engage its own responsability.* »

Risk management process

A threat has the potential to exploit one or more vulnerabilities to harm assets and thereby impact the organization (with regards to security criteria).

- **Risk identification**
  - Risk = Threat × Vulnerability × Impact
  - Main steps :
    - Identification & valuation of assets (context)
    - Identification of threats
    - Identification of vulnerabilities (taking into account existing controls)
    - Identification of consequences
  - Output → list of incidents scenarios
    - Ex : « *Traceability of information from external services is not provided. As a result, the reliability and/or quality of the information can not be guaranteed. The company therefore incurs the risk of generating and disseminating information calculated on the basis of unreliable data, and thus engage its own responsability.* »

- **Remaining steps of the methodology**
  - Assessment of incident scenario likelihood
  - Level of risk determination
  - Risk treatment (risk reduction, risk retention, risk avoidance, risk sharing)

Risk management process

- Do we still always the full development of a risk analysis and its various stages ?
    - It quickly become a tedious job to complete for a non-expert person
        - simultaneity of the analyzed security concepts : confidentiality, integrity, availability criteria, different gravity scales,…
    - It's mostly a way to demonstrate how our proposal can fit into a well-known methodology

[2] GDPR : General Data Protection Regulation (May 25, 2018)

Risk management process

- Do we still always the full development of a risk analysis and its various stages ?
  - It quickly become a tedious job to complete for a non-expert person
    - simultaneity of the analyzed security concepts : confidentiality, integrity, availability criteria, different gravity scales,…
  - It's mostly a way to demonstrate how our proposal can fit into a well-known methodology

  - We prefer a position such as that of the GDPR[2] which requires a Privacy Impact Assessment (PIA), which is a security risk analysis focused solely on the risks affecting personal data and their impact on the rights and freedoms of the persons concerned by these data
  - A PIA can be summarized in 4 steps :
    - **Context** → describe processing(s) of data under consideration, its (their) purposes and stakes ; supporting assets (hardware, software, networks, people, papers)
    - **Controls** → identify existing or planned controls
    - **Risks** (potential breaches) → 2 parts
      - Evaluate what is feared about the processes and the level of severity of these dreaded events
      - Identify the threats targeting the assets and which can lead to the dreaded events
    - **Decision** → risk treatment
  - Such an analysis will take up the points **specific to data controllability** (vulnerabilities, threats, consequences)

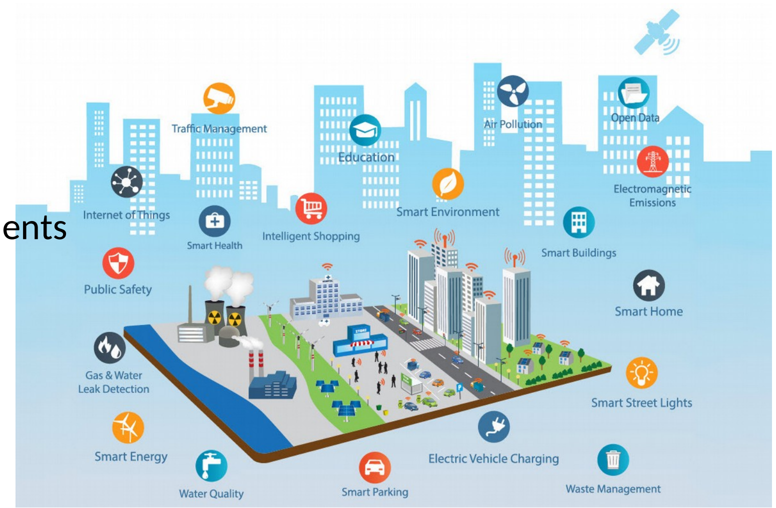
[2] GDPR : General Data Protection Regulation (May 25, 2018)

Future work

- Define mechanisms to provide data controllability indicators
  - Metadata for proactive traceability
  - Usage control policies
  - Peer-to-peer contractual agreements to set the security rules and to identify the responsabilities of each other

- Develop tools to support (data controllability) supervision
  - Decision support tool to monitor data exchanges (according to contracts)
  - Preprocess received data to prevent information insufficiently controlled → business processes do not have to deal with that

- Application areas
  - IoT, smart buildings, smart cities
  - Right to informational self-determination in smart environments
  - Smart grids, energy mix



Thank you for your attention

CONTACT

Manuel MUNIER
Universite de Pau et des Pays de l’Adour, E2S UPPA, LIUPPA, EA 3000
Mont-de-Marsan, France

manuel.munier@univ-pau.fr
http://munier.perso.univ-pau.fr/

UNIVERSITÉ DE PAU ET DES PAYS DE L’ADOUR | COLLÈGE STEE SCIENCES ET TECHNOLOGIES POUR L’ÉNERGIE ET L’ENVIRONNEMENT

e2s UPPA energy environment solutions