
Designing GDPR compliant software




Agenda

- GDPR Summary
- What does compliance with GDPR mean ?
- Example of GDPR Accountability
- Consent & Purpose Management
- What does security mean in the GDPR?
- Privacy Design Strategies
- Security within the Data Lifecycle Management
- OWASP SAMM & GDPR
- Anonymisation & Pseudonimisation
- Data Breach Management
- Privacy by design

General Data Protection Regulation

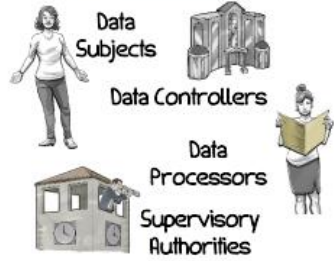
GDPR will be enforce as from 25 may 2018 directly in the 28 EU Member state

TERRITORIAL SCOPE




EU Establishments
Non-EU Established Organizations
Offer goods or services or engaging in monitoring within the EU.

THE PLAYERS

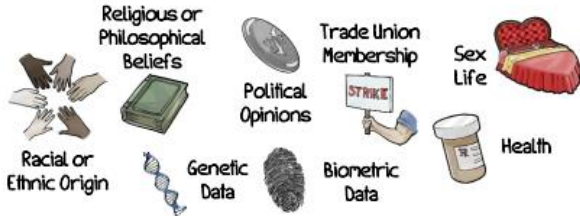


PERSONAL DATA



Identified Identifiable

SENSITIVE DATA




RESPONSIBILITIES OF DATA CONTROLLERS AND PROCESSORS

LAWFUL PROCESSING

Collection and processing of personal data must be for "specified, explicit and legitimate purposes"
— with consent of data subject or necessary for


- performance of a contract
- compliance with a legal obligation
- to protect a person's vital interests
- task in the public interest
- legitimate interests

CONSENT




Consent must be freely given, specific, informed, and unambiguous.

SECURITY




Designate DPO if core activity involves regular monitoring or processing large quantities of personal data.

RECORD OF DATA PROCESSING ACTIVITIES




Maintain a documented register of all activities involving processing of EU personal data.

DATA PROTECTION BY DESIGN



built in starting at the beginning of the design process

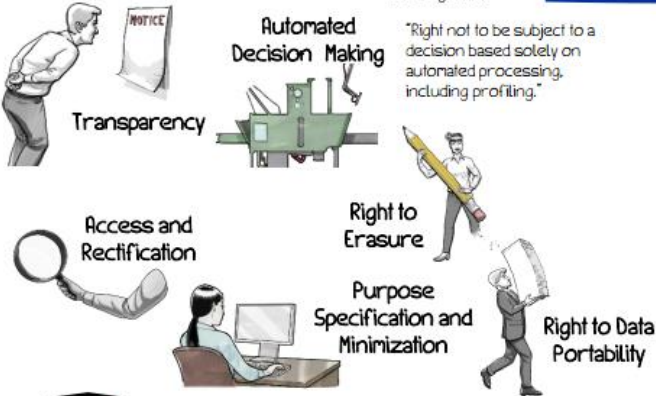
DATA IMPACT ASSESSMENT



For high risk situations


GDPR

RIGHTS OF DATA SUBJECTS



"Right not to be subject to a decision based solely on automated processing, including profiling."


ENFORCEMENT



Fines
Up to 20 million euros or 4% of total annual worldwide turnover. Less serious violations: Up to 10 million euros or 2% of total annual worldwide turnover.

Effective Judicial Remedies:
compensation for material and non-material harm.

DATA BREACH NOTIFICATION




A *personal data breach* is "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed."


If likely to result in a high privacy risk → notify data subjects


Notify supervisory authorities no later than 72 hours after discovery.

INTERNATIONAL DATA TRANSFER



Adequate Level of Data Protection





www.teachprivacy.com

Workforce awareness training by Prof. Daniel J. Solove

Please ask permission to reuse or distribute

What does compliance with GDPR mean ?

The compliance approach implemented is based on the respect for privacy principles:

- respect for legal principles for privacy protection
- management of risks related to the security of personal data and having an impact on data subjects' privacy

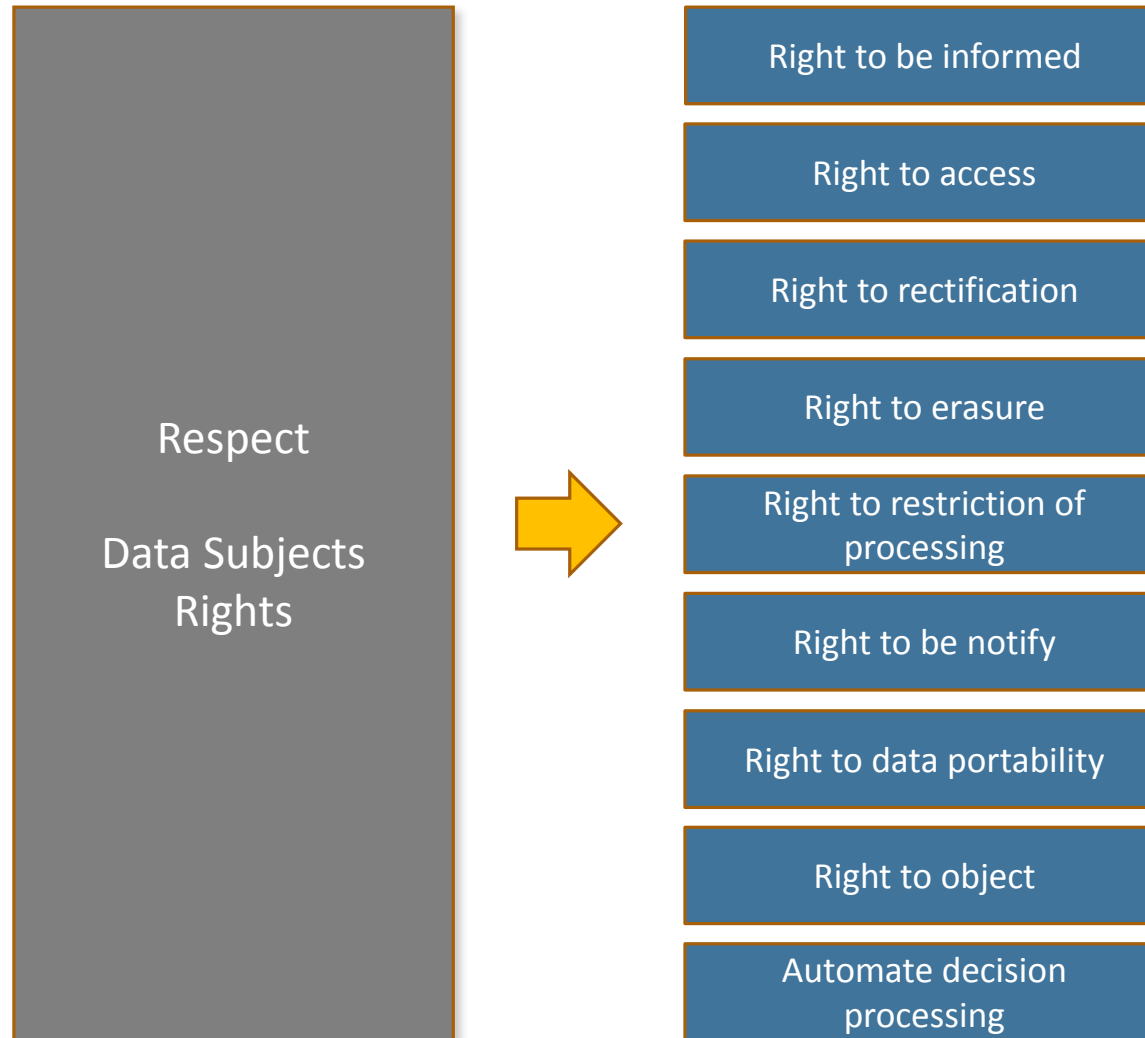


What does compliance with GDPR mean ?

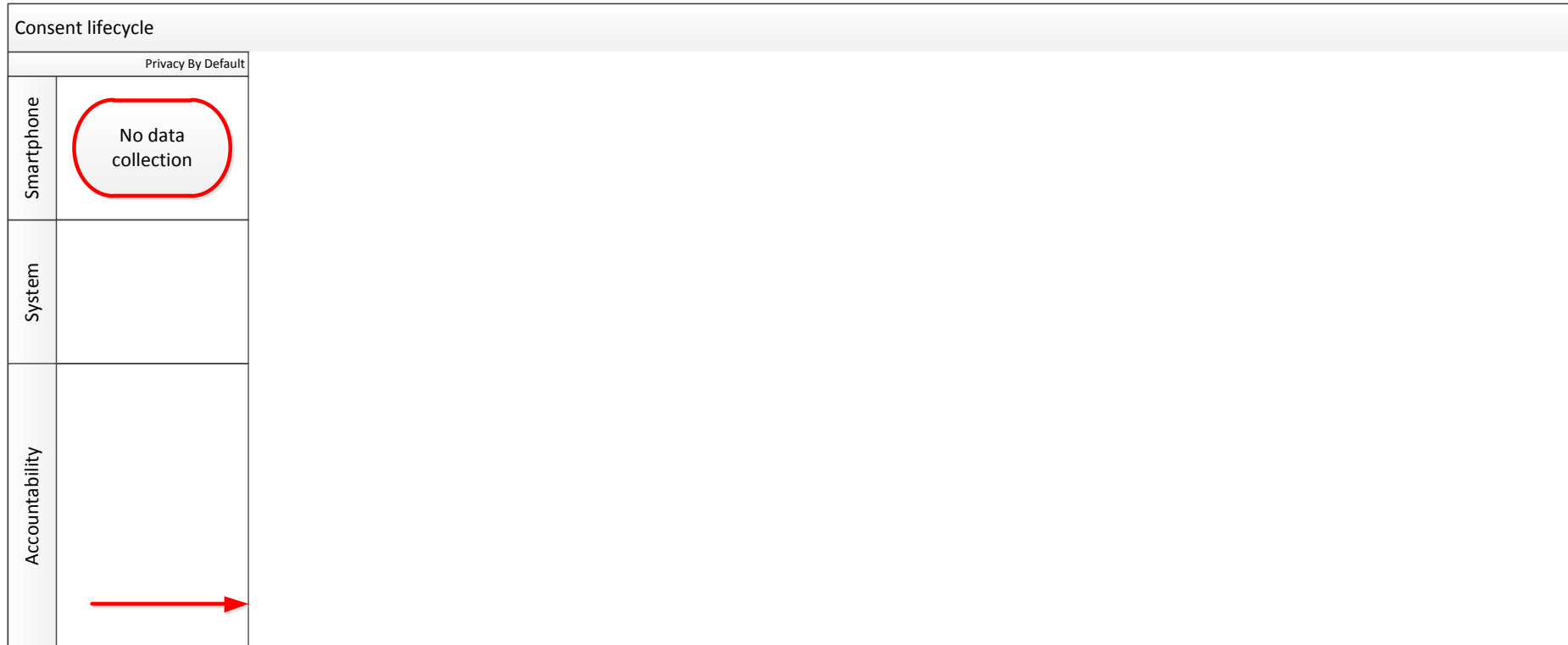
Comply with legal requirements



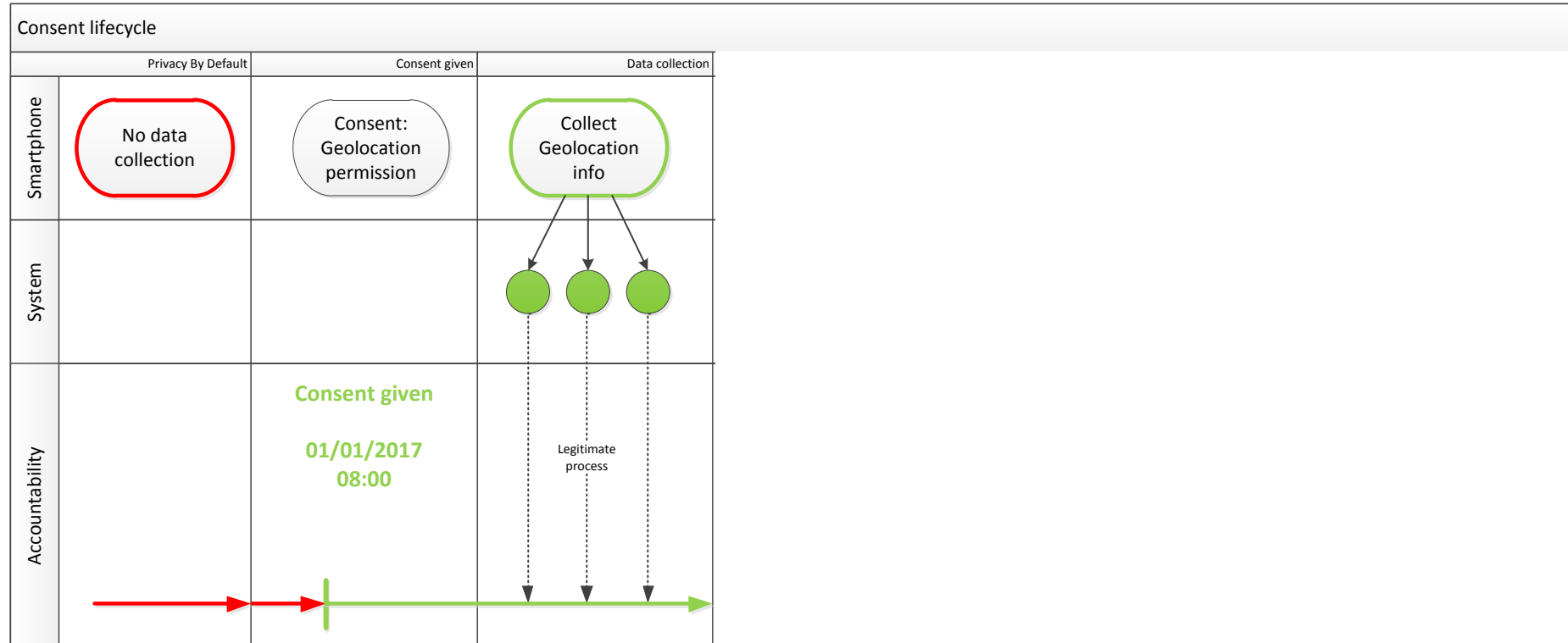
What does compliance with GDPR mean ?



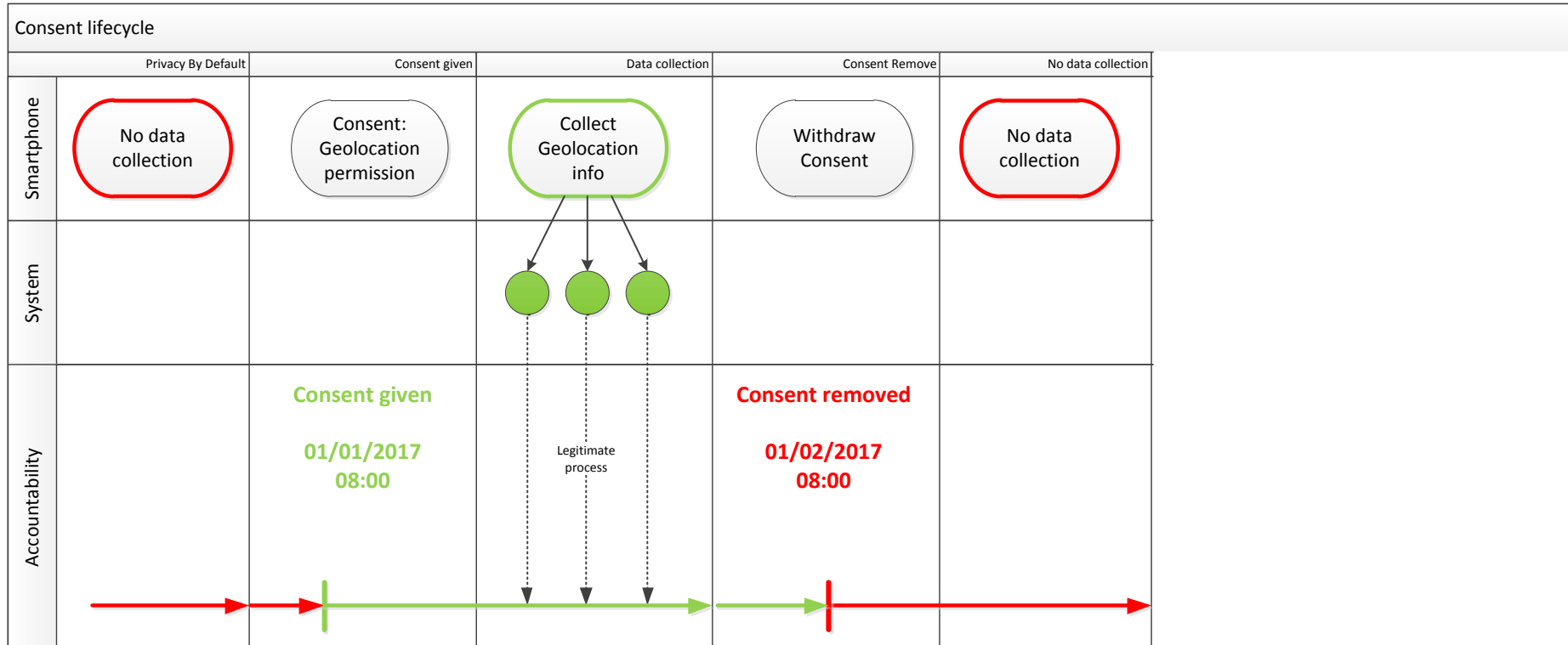
GDPR Accountability



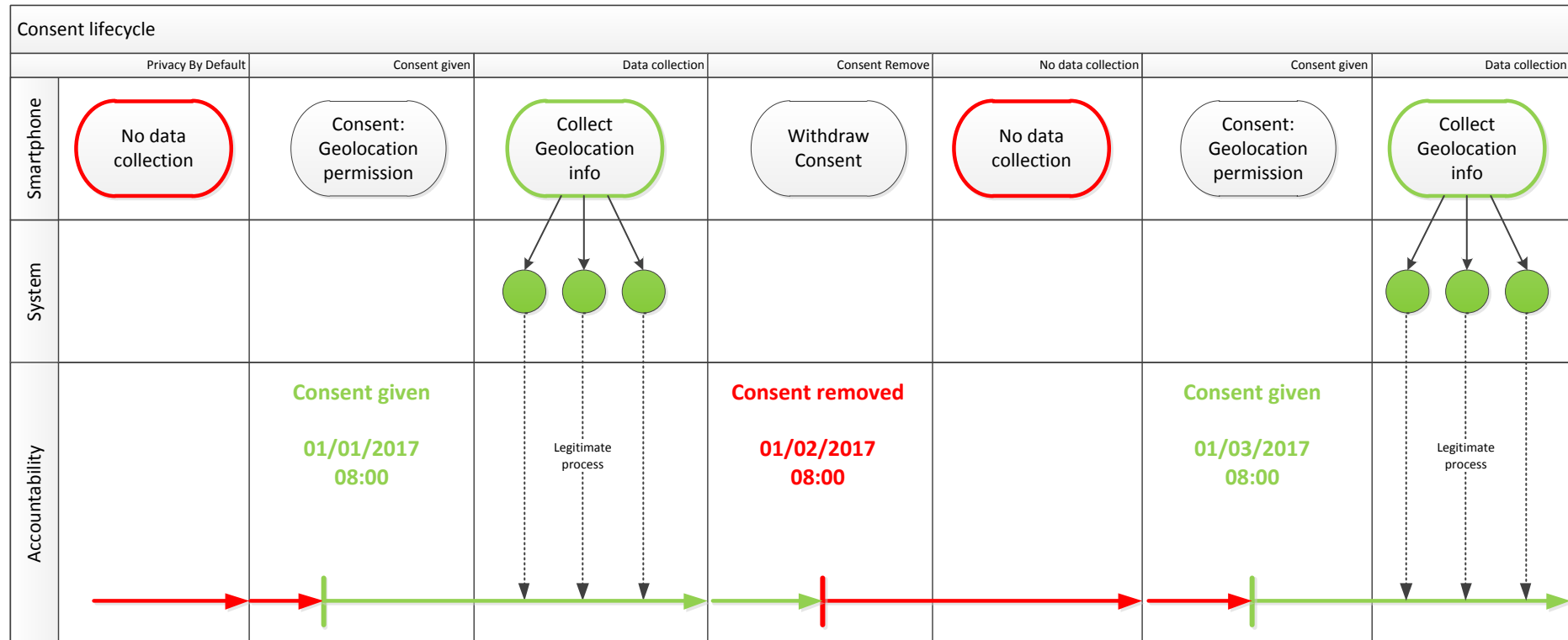
GDPR Accountability



GDPR Accountability



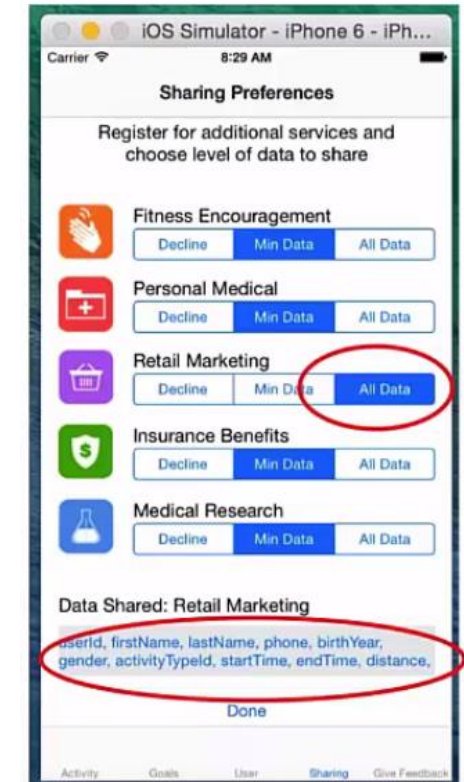
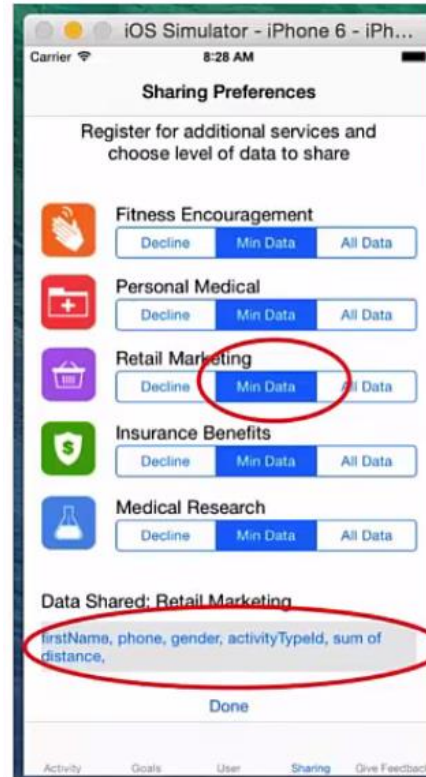
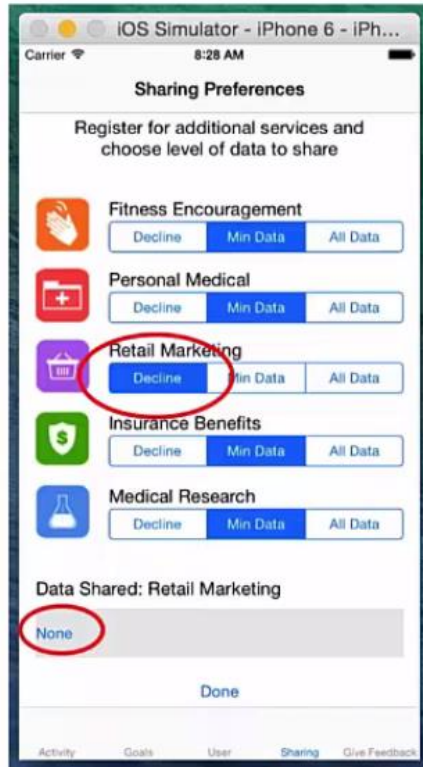
GDPR Accountability



Create evidences to :

- Know when the users give or withdraw consents
- Know which purpose is used when processing PII

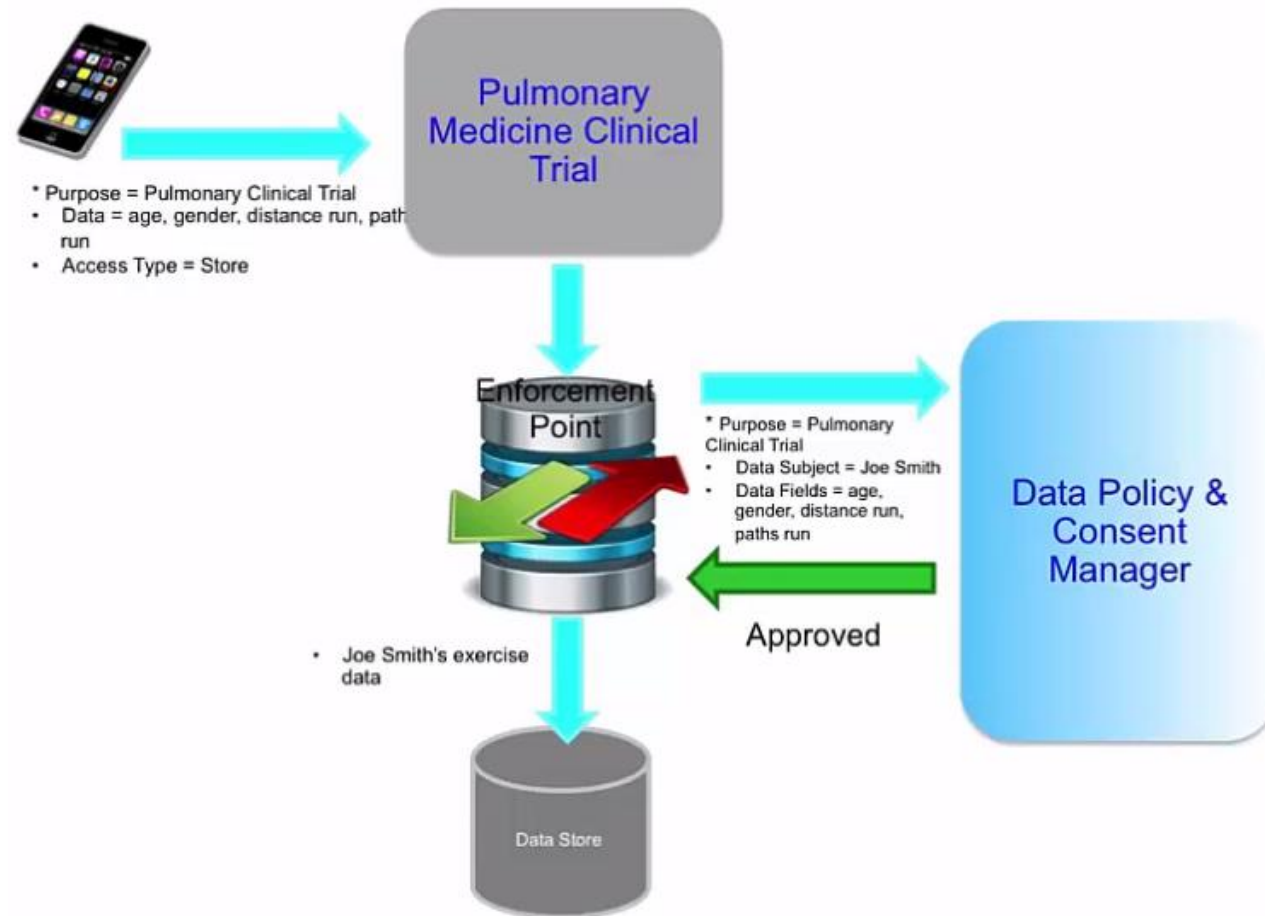
Consent & Purpose Management



The main objective: Give the control to the end user to manage how his personal data can be exchanged

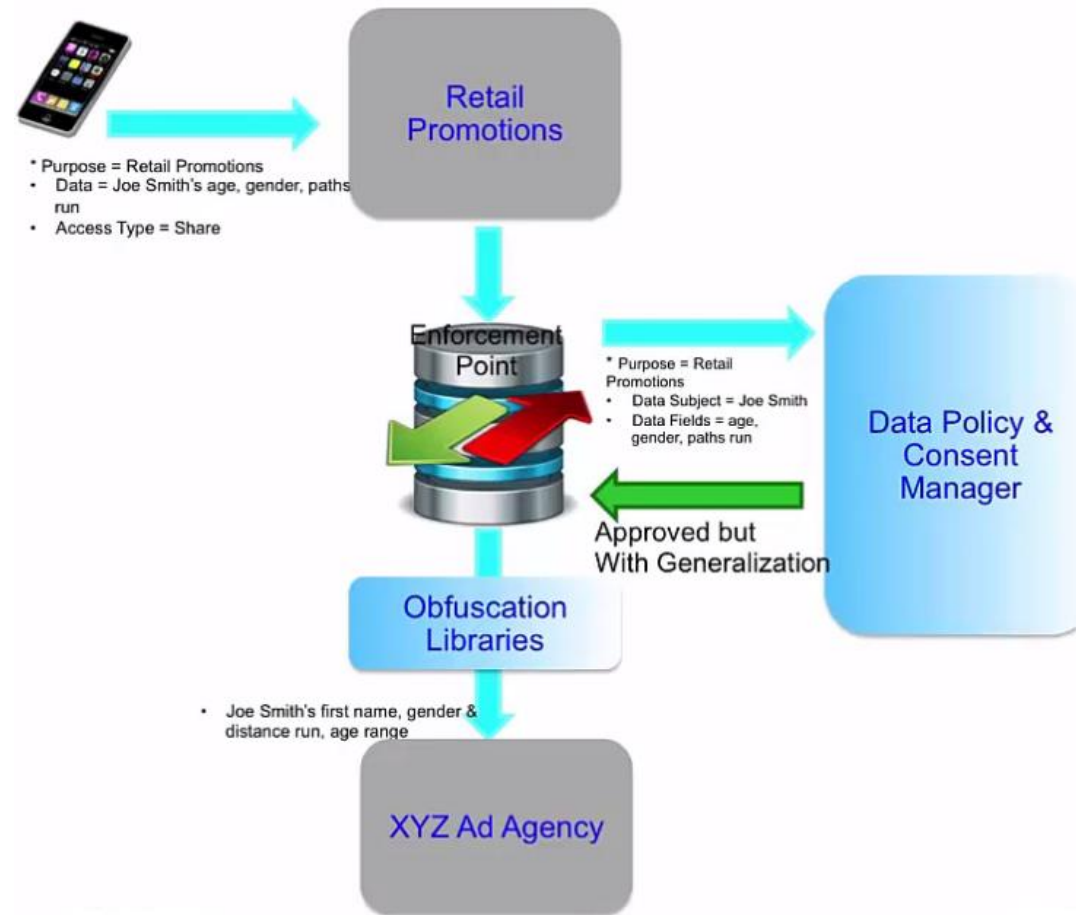
- 5 purposes exist in this sample App

Consent & Purpose Management



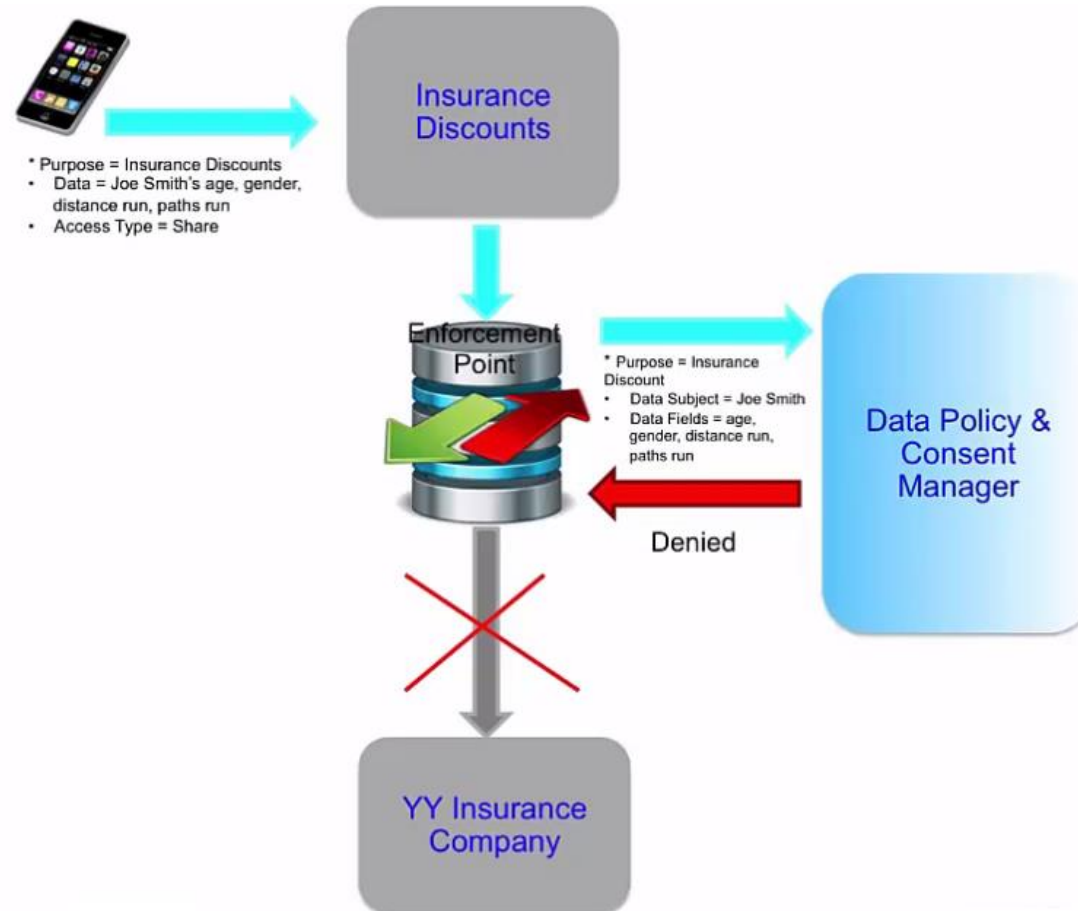
Source: IBM - <https://youtu.be/GDzYry6GCFg?t=2348>

Consent & Purpose Management



Source: IBM - <https://youtu.be/GDzYry6GCFg?t=2348>

Consent & Purpose Management

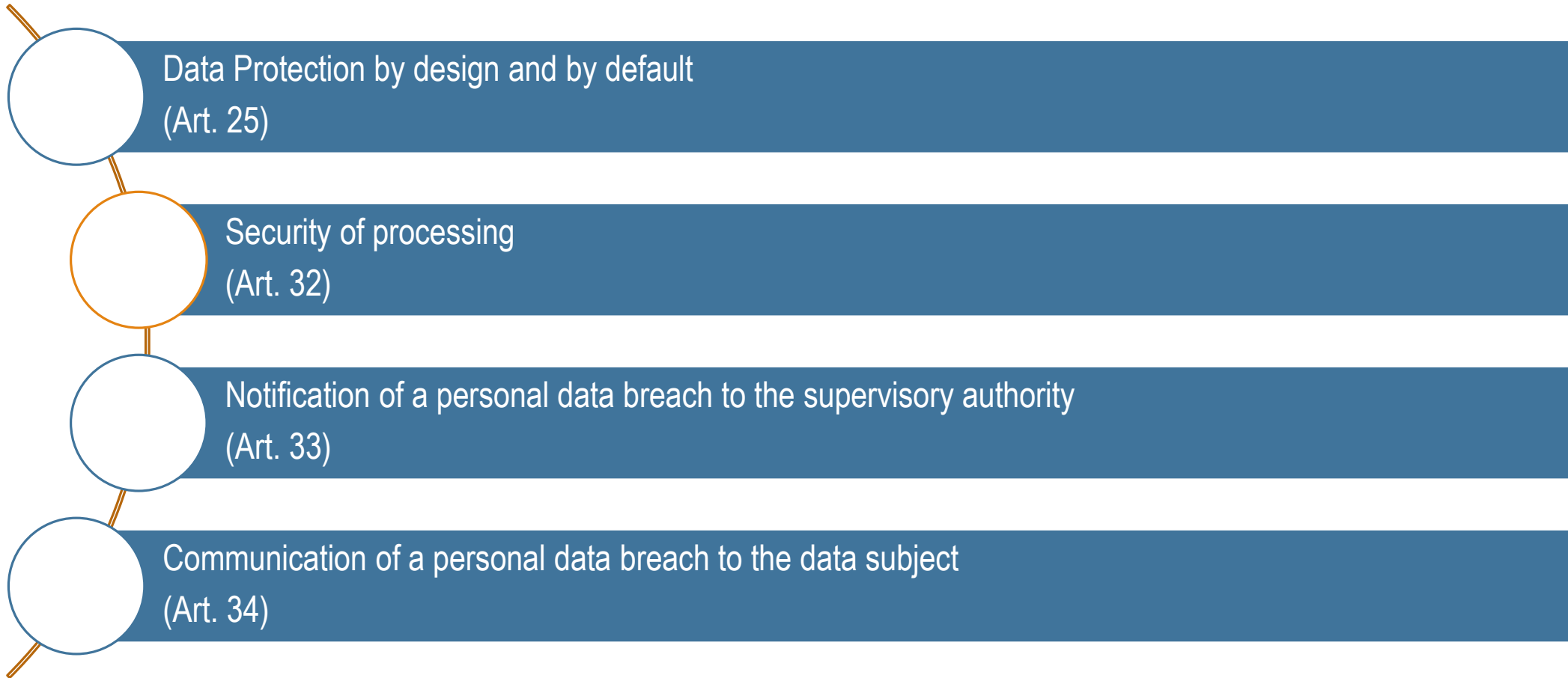


Source: IBM - <https://youtu.be/GDzYry6GCFg?t=2348>

GDPR Challenges

- The term "purpose" is a notion that does not exist in computer systems
- Data Controller must understand for what purpose(s) personal data is collected and processed
- Managing consent can become complex
 - Consent should be collected by purpose
 - Indicate for each purpose what type of data is collected
 - Data controller must maintain an audit trail of how data is used and for what purpose
 - Consent becomes a dynamic and flexible feature in an application
- Authorization mechanisms are not yet design to manage consent

What does security mean in the GDPR?



What does security mean in the GDPR?

Article 32

Security of processing

1. Taking into account:

- The state of the art, the costs of implementation
- The nature, scope, context and purposes of processing
- The risk of varying likelihood and severity for the rights and freedoms of natural persons

The controller and the processor shall implement appropriate technical & administrative measures to ensure a level of security appropriate to the risk

What does security mean in the GDPR?

Article 32

Security of processing

The controller and the processor shall implement appropriate technical & administrative measures to ensure a level of security appropriate to the risk

- the pseudonymisation and encryption of personal data;
- the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
- the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;
- a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.

2. In assessing the appropriate level of security account shall be taken in particular of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed.

What does security mean in the GDPR?

Article 25

Data Protection by design and by default

(1) Taking into account:

- The state of the art, the costs of implementation
- The nature, scope, context and purposes of processing
- The risk of varying likelihood and severity for the rights and freedoms of natural persons

The controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, **implement appropriate technical and organisational measures**, such as pseudonymisation, which are **designed to implement data-protection principles**, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.

What does security mean in the GDPR?

Article 25

Data Protection by design and by default

(2) The controller shall **implement** appropriate technical and organisational **measures** for **ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed.**

That obligation applies to

- the amount of personal data collected,
- the extent of their processing,
- the period of their storage and their accessibility.

In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons.

What does compliance with GDPR mean ?

Article 35: Data protection impact Assessment

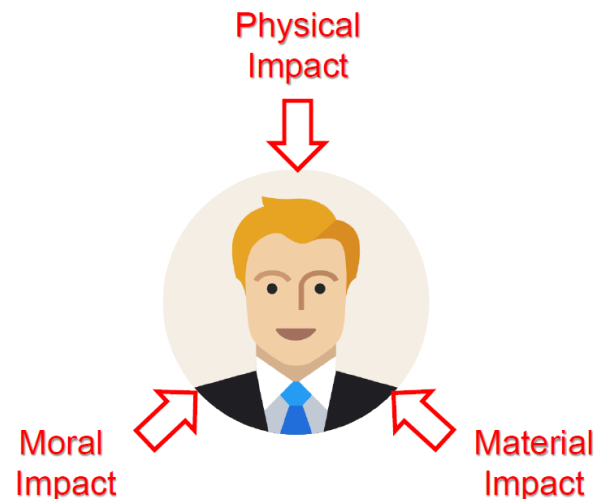
1. Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a **high risk** to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks.

7. The assessment shall contain at least:

- (a) a systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;
- (b) an assessment of the necessity and proportionality of the processing operations in relation to the purposes;
- (c) an assessment of the risks to the rights and freedoms of data subjects referred to in paragraph 1; and
- (d) the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned.

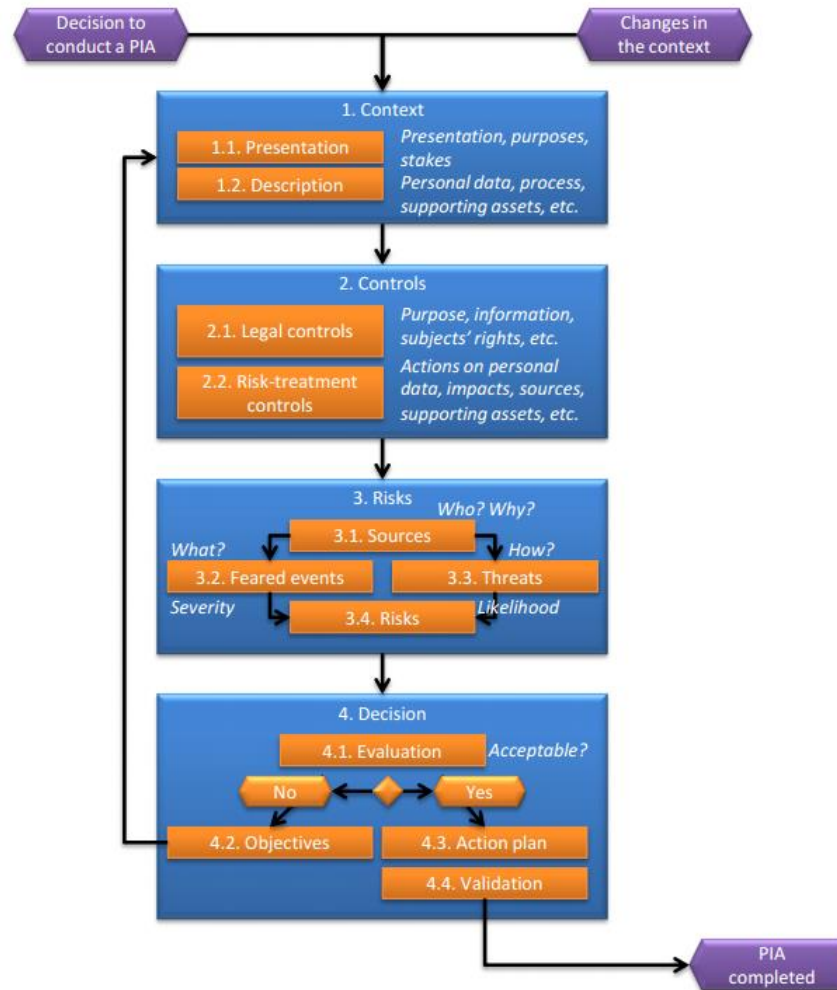
What does compliance with GDPR mean ?

Recital 75: The **risk to the rights and freedoms of natural persons**, of varying **likelihood** and **severity**, may result from personal data processing which could lead to **physical**, **material** or **non-material damage**, in particular: where the processing may give rise to discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy, unauthorized reversal of pseudonymisation, or any other significant economic or social disadvantage;



Existing PIA frameworks

Privacy Impact Assessment (PIA), Commission nationale de l'informatique et des libertés (CNIL), 2015.



Privacy Impact Assessment Process

CNIL Methodology

In summary, to comply with GDPR, it is necessary to:

1. Context: Define and describe the context of the processing of personal data under consideration and its stakes;

2. Controls: Identify existing or planned controls

3. Risks: Assess privacy risks to ensure they are properly treated;

4. Decisions: make the decision to validate the manner in which it is planned to comply with privacy principles and treat the risks,



Figure 4 – General approach for carrying out a PIA

*** The current CNIL Approach is still based on the French regulation and not 100% GDPR oriented at legal point of view.**

Privacy Design Strategies

Privacy Design Strategies

The eight privacy design strategies as derived by Hoepman [*] from the legal principles underlying data protection legislation. This work distinguishes data oriented strategies and process oriented strategies.

Data Oriented strategies

4 data oriented strategies can support the unlinkability protection goal and primarily address the principles of necessity and data minimization

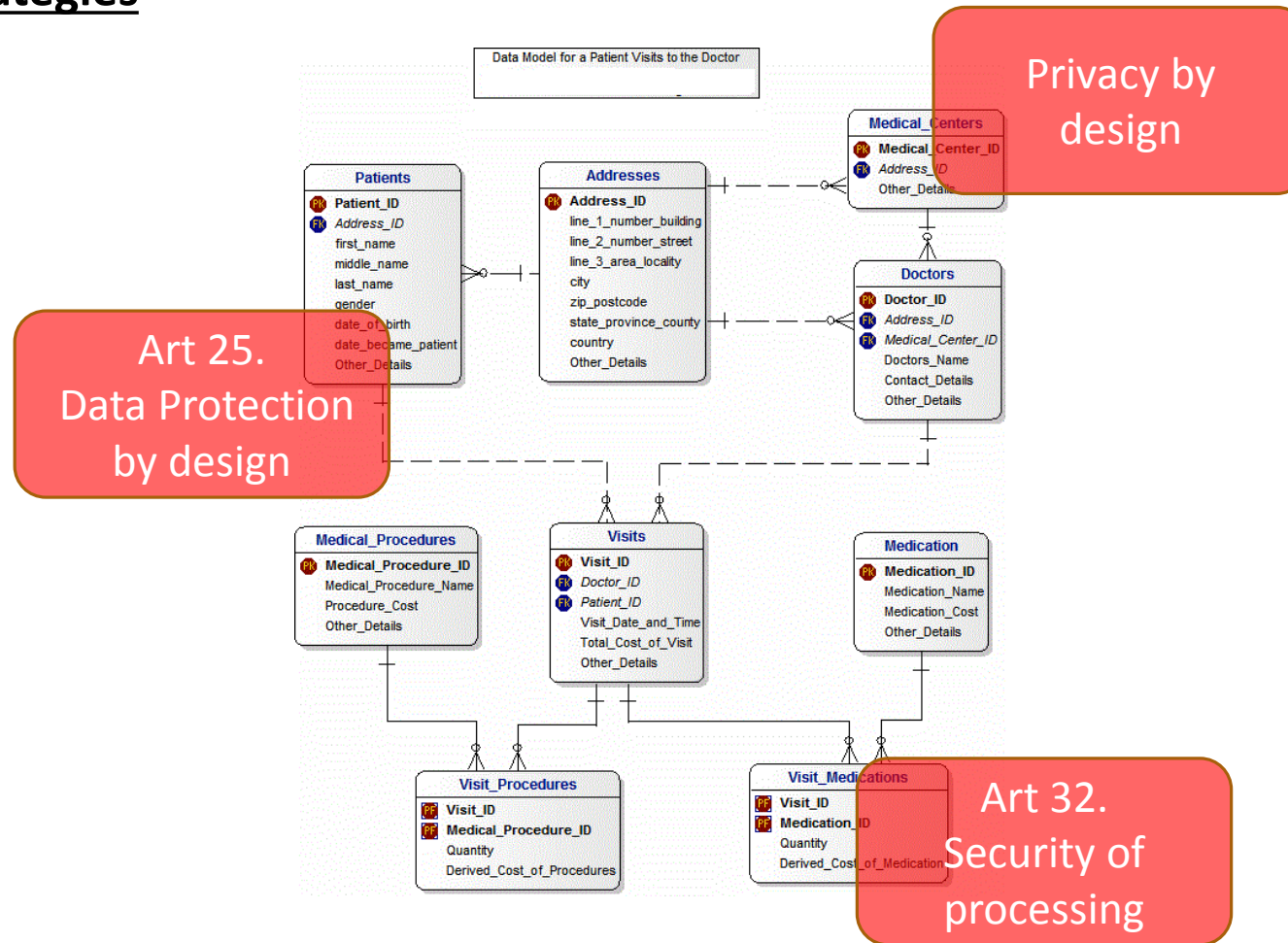
Process Oriented strategies

4 process oriented strategies can support a series of actions or steps taken in order to achieve GDPR compliance.

Jaap-Henk Hoepman. Privacy design strategies – (extended abstract). In ICT Systems Security and Privacy Protection - 29th IFIP TC 11 International Conference, SEC 2014, Marrakech, Morocco, June 2-4, 2014. Proceedings, pages 446–459, 2014.

Privacy Design Strategies

Data Oriented strategies



Oracle Data Security Architecture

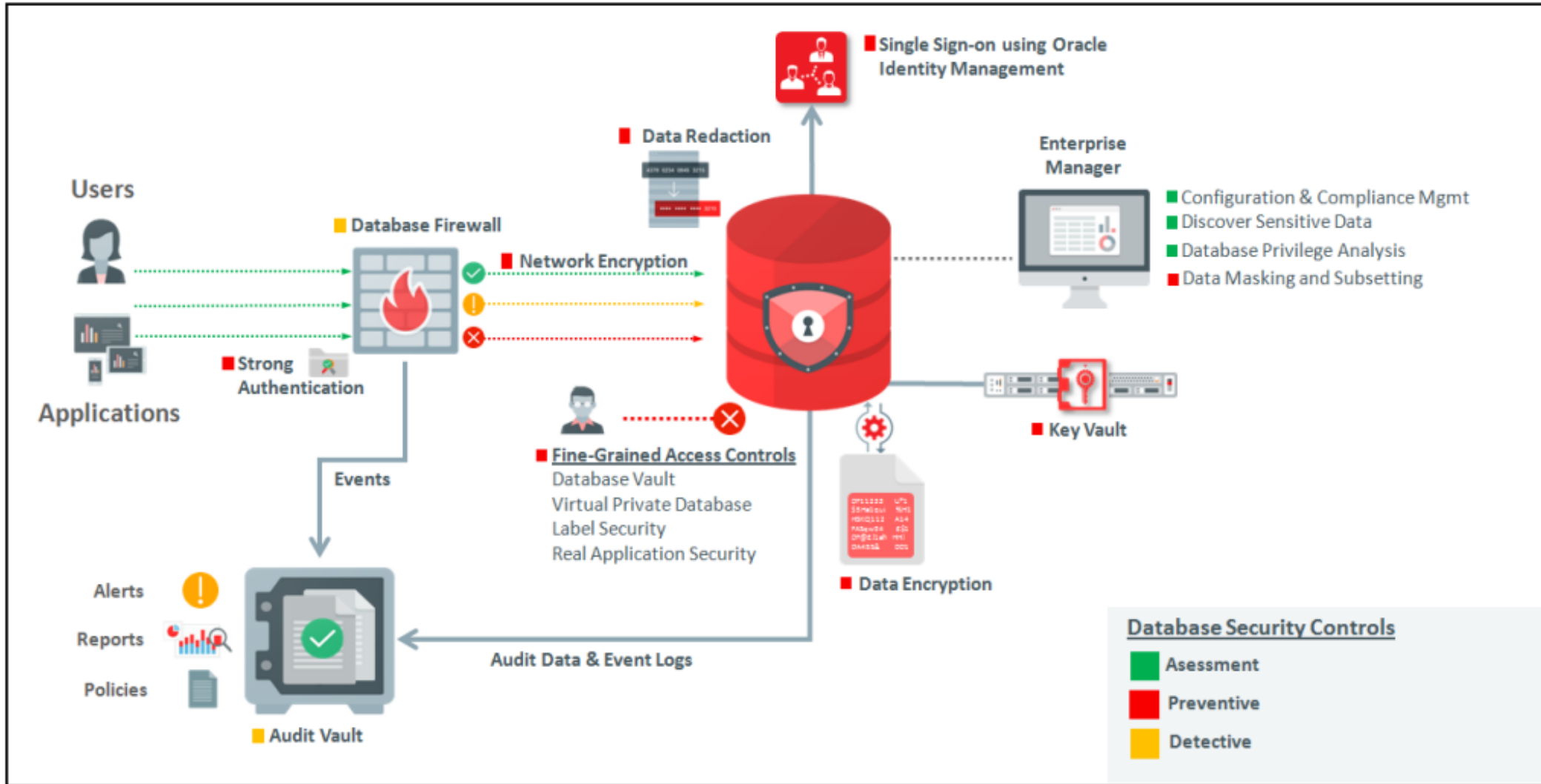


Figure 8: Oracle Maximum Data Security Architecture

Privacy Design Strategies

Data Oriented strategies: MINIMIZE

the amount of personal data that is processed should be restricted to the minimal amount possible. By ensuring that no, or no unnecessary, data is collected, the possible privacy impact of a system is limited

In many cases information about the individual may be collected without him/her even being aware (e.g. relating to his/her web searches and overall online behaviour). Privacy enhancing technologies that can support internet and mobile users' privacy are available today, including anti-tracking, encryption, identity masking and secure file sharing tools.

Implementation: Define what data are needed before collection, select before collect (reduce data fields, define relevant controls, delete unwanted information, etc), Privacy Impact Assessments.

Privacy Design Strategies

Data Oriented strategies: MINIMIZE

STRATEGY

EXCLUDE: refraining from processing a data subject's personal data, partly or entirely, akin to blacklisting or optout.

SELECT: decide on a case by case basis on the full or partial usage of personal data, akin to whitelisting or opt-in.

STRIP: removing unnecessary personal data fields from the system's representation of each user.

DESTROY: completely removing a data subject's personal data.

Privacy Design Strategies

Data Oriented strategies: HIDE

Any personal data, and their interrelationships, should be hidden from plain view.

It can be achieved by the use of encryption of data (when stored, or when in transit), anonymisation or the use of pseudonyms

Implementation: Privacy enhancing end-user tools, e.g. anti-tracking tools, encryption tools, identity masking tools, secure file sharing, etc.

Encryption of data at rest. Authentication and access control mechanisms. Other measures for secure data storage

Privacy Design Strategies

Data Oriented strategies: HIDE

STRATEGY

RESTRICT: preventing unauthorized access to personal data.

MIX: processing personal data randomly within a large enough group to reduce correlation.

OBFUSCATE: preventing understandability of personal data to those without the ability to decipher it.

DISSOCIATE: removing the correlation between different pieces of personal data.

Privacy Design Strategies

Data Oriented strategies: SEPARATE

Personal data should be processed in a distributed fashion, in separate compartments whenever possible. By separating the processing or storage of several sources of personal data that belong to the same person, complete profiles of one person cannot be made

STRATEGY

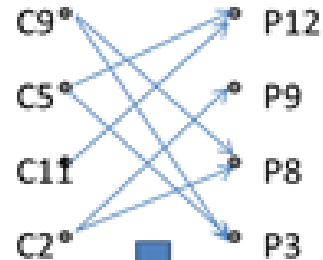
DISTRIBUTE: partitioning personal data so that more access is required to process it.

ISOLATE: processing parts of personal data independently, without access or correlation to related parts.

Privacy Design Strategies

Data Oriented strategies: SEPARATE

Cid	DOB	Sex	ZIP
C1	7/18/79	F	30323
C2	2/17/83	M	30323
C3	5/07/77	M	30327
C4	1/5/76	F	30328
C5	8/4/82	M	30330
C6	3/9/79	M	30331
C7	4/10/64	M	30331
C8	2/6/81	F	30334
C9	7/14/72	F	30337
C10	9/25/74	M	30338
C11	4/28/80	M	30338
C12	3/12/78	M	30339



Cid	Did
C2	P8
C2	P9
C5	P3
C5	P12
C9	P3
C9	P8
C11	P12

Did	Drug name	Category
P1	epinephrine	bronchodilator
P2	ibuprofen	analgesic
P3	Zovirax	antiviral
P4	Tylenol	analgesic
P5	erythromycin	antibiotic
P6	cortisone	anti-inflammatory
P7	gentamicin	antibiotic
P8	insulin	hypoglycemic
P9	sertraline	antidepressant
P10	tramadol	analgesic
P11	cetirizine	antihistamine
P12	zolpidem	hypnotic

Revealing the associations between Cid and Did violates privacy

Privacy Design Strategies

Data Oriented strategies: AGGREGATE

Personal data should be processed at the highest level of aggregation and with the least possible detail in which it is (still) useful.

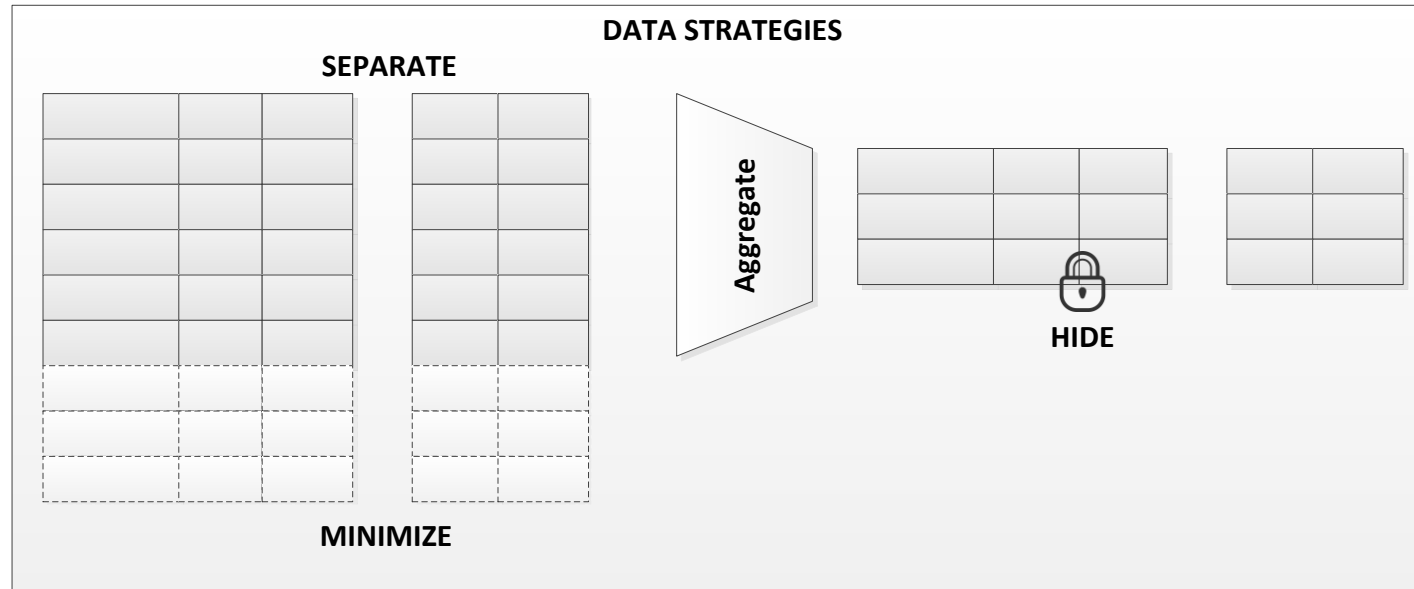
STRATEGY

SUMMARIZE: extracting commonalities in personal data by finding and processing correlations instead of the data itself.

GROUP: inducing less detail from personal data prior to processing, by allocating into common categories.

Privacy Design Strategies

Data Oriented strategies



MINIMIZE	HIDE	SEPARATE	AGGREGATE
EXCLUDE	RESTRICT	DISTRIBUTE	SUMMARIZE
SELECT	MIX	ISOLATE	GROUP
STRIP	OBFUSSATE		
DESTROY	DISSOCIATE		

Privacy Design Strategies

Process Oriented strategies : INFORM

The INFORM strategy corresponds to the important notion of transparency

Whenever data subjects use a system, they should be informed about which information is processed, for what purpose, and by which means

STRATEGY

SUPPLY: making available extensive resources on the processing of personal data, including policies, processes, and potential risks.

NOTIFY: alerting data subjects to any new information about processing of their personal data in a timely manner.

EXPLAIN: detailing information on personal data processing in a concise and understandable form.

Privacy Design Strategies

Process Oriented strategies : INFORM

Microsoft Privacy Statement

History
Privacy notice
Summary

Table of contents

Topic summary

[Last Updated: November 2016](#) [What's new?](#)

[Expand All](#)
[Print](#)

Your privacy is important to us. This privacy statement explains what personal data we collect from you and how we use it. We encourage you to read the summaries below and to click on "Learn More" if you'd like more information on a particular topic.

The product-specific details sections provide additional information relevant to particular Microsoft products. This statement applies to the Microsoft products listed below, as well as other Microsoft products that display this statement. References to Microsoft products in this statement include Microsoft services, websites, apps, software and devices.

Personal Data We Collect
 How We Use Personal Data
 Reasons We Share Personal Data
 How to Access & Control Your Personal Data
 Cookies & Similar Technologies
 Microsoft account
 Other Important Privacy Information [v](#)

Product-specific details:
[Bing](#)
[Cortana](#)
[Groove Music/Movies & TV](#)
[Microsoft Cognitive Services](#)
[Microsoft Health Services](#) [v](#)
[Microsoft Translator](#)

Personal Data We Collect

Microsoft collects data to operate effectively and provide you the best experiences with our products. You provide some of this data directly, such as when you create a Microsoft account, submit a search query to Bing, speak a voice command to Cortana, upload a document to OneDrive, purchase an MSDN subscription, sign up for Office 365, or contact us for support. We get some of it by recording how you interact with our products by, for example, using technologies like cookies, and receiving error reports or usage data from software running on your device. We also obtain data from third parties.

[Learn More](#)
[Top of page](#) [^](#)

How We Use Personal Data

Microsoft uses the data we collect to provide you the products we offer, which

Privacy Design Strategies

Process Oriented strategies : INFORM

Table of contents

- Personal Data We Collect
- How We Use Personal Data
- Reasons We Share Personal Data
- [How to Access & Control Your Personal Data](#)
- Cookies & Similar Technologies
- Microsoft account
- Other Important Privacy Information ▾
- Product-specific details:**
- Bing
- Cortana
- Groove Music/Movies & TV
- Microsoft Cognitive Services
- Microsoft Health Services ▾
- Microsoft Translator
- MSN
- Office
- OneDrive
- Outlook
- Silverlight
- Skype
- Store
- SwiftKey
- Windows ▾
- Xbox
- Enterprise Products

Personal Data We Collect

Microsoft collects data to operate effectively and provide you the best experiences with our products. You provide some of this data directly, such as when you create a Microsoft account, submit a search query to Bing, speak a voice command to Cortana, upload a document to OneDrive, purchase an MSDN subscription, sign up for Office 365, or contact us for support. We get some of it by recording how you interact with our products by, for example, using technologies like cookies, and receiving error reports or usage data from software running on your device.

We also obtain data from third parties. For example, we supplement the data we collect by purchasing demographic data from other companies. We also use services from other companies to help us determine a location based on your IP address in order to customize certain products to your location.

You have choices about the data we collect. When you are asked to provide personal data, you may decline. But if you choose not to provide data that is necessary to provide a product or feature, you may not be able to use that product or feature.

The data we collect depends on the products and features you use, and can include the following:

Name and contact data. We collect your first and last name, email address, postal address, phone number, and other similar contact data.

Credentials. We collect passwords, password hints, and similar security information used for authentication and account access.

Demographic data. We collect data about you such as your age, gender, country, and preferred language.

Payment data. We collect data necessary to process your payment if you make purchases, such as your payment instrument number (such as a credit card number), and the security code associated with your payment instrument.

Usage data. We collect data about how you and your device interact with Microsoft and our products. For example, we collect:

Topic detail

Privacy Design Strategies

Process Oriented strategies : INFORM



Change History for Microsoft Privacy Statement

[Back to the privacy statement](#)

November 2016

- In **How We Use Personal Data**, we updated **Advertising** to better clarify the use of your data by third parties to customize the ads you see.
- In **How to Access & Control Your Personal Data**, we updated **Your Communications Preferences**, clarifying how to modify your preferences.
- In **Other Important Information**, we updated the **Where We Store and Process Personal Data** section to reflect Microsoft's participation in the EU-U.S. Privacy Shield program.
- In **Bing**, we removed the Bing Rewards Program section, as Bing Rewards has been replaced by Microsoft Rewards.
- We added a new **Microsoft Cognitive Services** section to explain how we collect and use data when developers use the services. We also clarified that Microsoft Cognitive Services are not Enterprise Products under this privacy statement.
- We added a new **Microsoft Translator** section, to explain how Microsoft Translator, Collaborative Translations Framework, and Microsoft Translator Hub collect and use data.
- In **Windows**, we revised the **Telemetry & Error Reporting** section to reflect that wireless network identifiers are collected at the optional "Enhanced" level of telemetry rather than at the "Basic" level.
- We added a new **captioning** section in **Xbox** to explain how Microsoft incorporates a voice-to-text feature to provide captioning of in-game chat for users who need it.

Changes introduced by a specific version of a privacy notice

September 2016

In **Enterprise Products**, we added links to privacy notices that still apply to certain enterprise offerings.

Privacy Design Strategies

Process Oriented strategies: CONTROL

Allow the users to control what kind of information is processed about them

STRATEGY

CONSENT: only processing the personal data for which explicit, freely-given, and informed consent is received.

CHOOSE: allowing for the selection or exclusion of personal data, partly or wholly, from any processing.

UPDATE: providing data subjects with the means to keep their personal data accurate and up to date.

RETRACT: honoring the data subject's right to the complete removal of any personal data in a timely fashion.

Privacy Design Strategies

Process Oriented strategies: ENFORCE

A privacy policy compatible with legal requirements should be in place and should be enforced.

This strategy supports the accountability principles

STRATEGY

CREATE: acknowledging the value of privacy and deciding upon policies which enable it, and processes which respect personal data.

MAINTAIN: considering privacy when designing or modifying features, and updating policies and processes to better protect personal data.

UPHOLD: ensuring that policies are adhered to by treating personal data as an asset, and privacy as a goal to incentivize as a critical feature.

Privacy Design Strategies

Process Oriented strategies: DEMONSTRATE

The final strategy, DEMONSTRATE, requires a data controller to be able to demonstrate compliance with the privacy policy and any applicable legal requirements. This strategy supports the accountability principles

STRATEGY

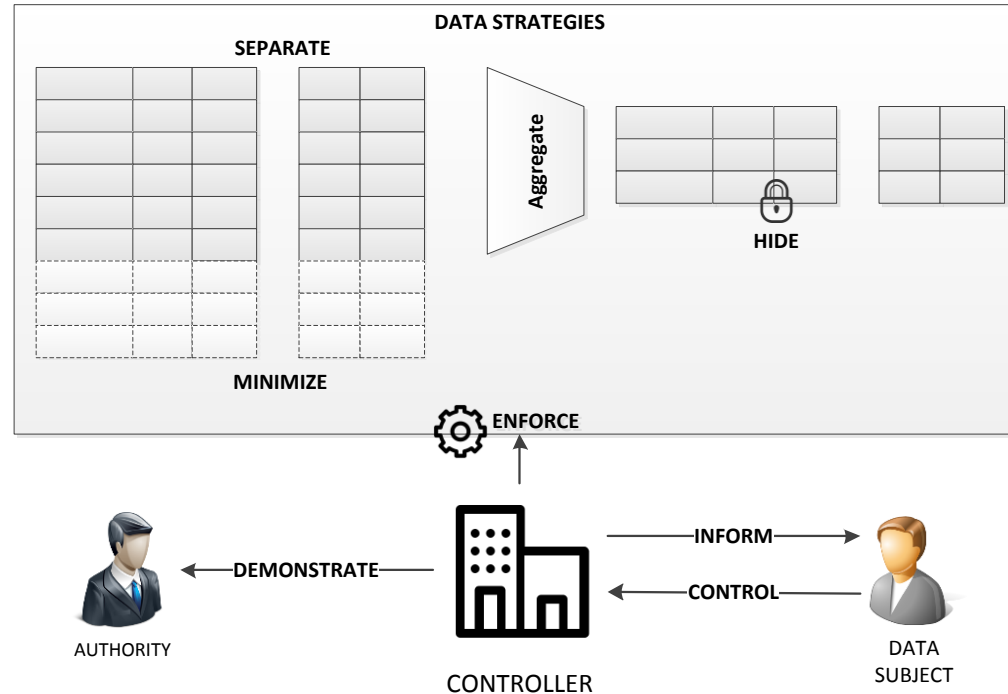
LOG: tracking all processing of data, without revealing personal data, securing and reviewing the information gathered for any risks.

AUDIT: examining all day to day activities for any risks to personal data, and responding to any discrepancies seriously.

REPORT: analyzing collected information on tests, audits, and logs periodically to review improvements to the protection of personal data.

Privacy Design Strategies

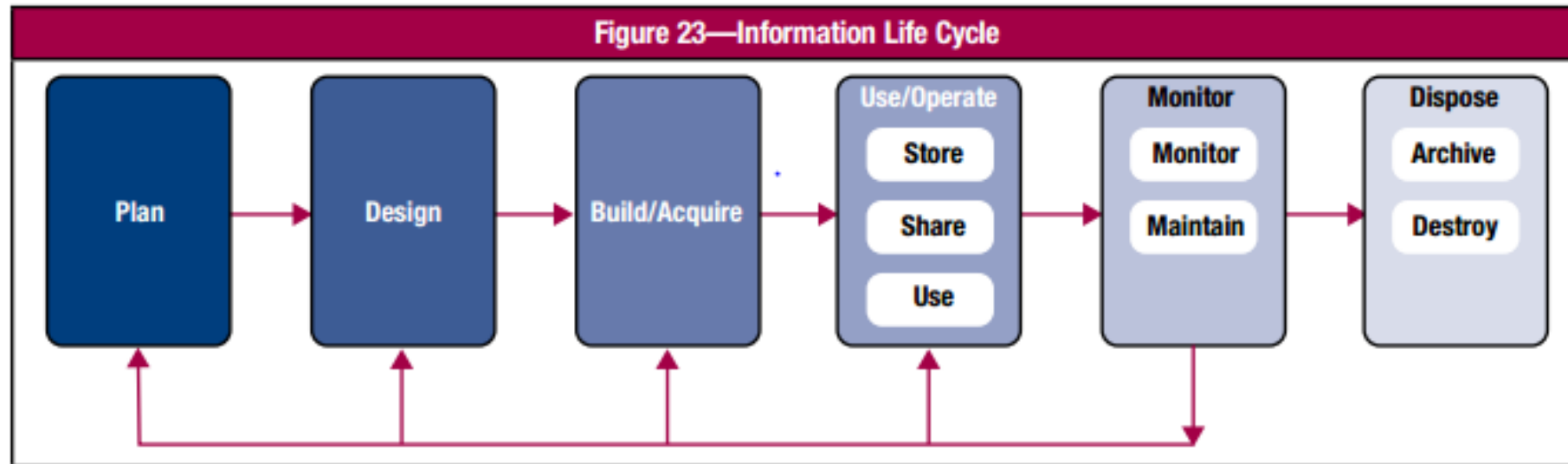
Data Oriented strategies



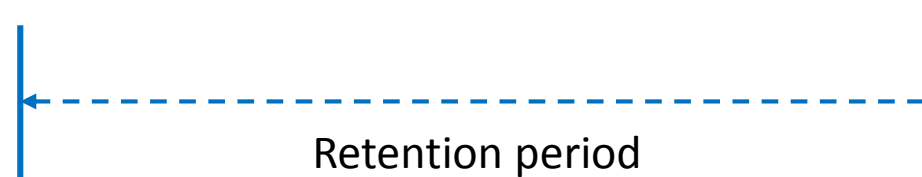
INFORM	CONTROL	ENFORCE	DEMONSTRATE
EXCLUDE SELECT STRIP DESTROY	RESTRICT MIX OBFUSCATE DISSOCIATE	DISTRIBUTE ISOLATE	SUMMARIZE GROUP

Security within the Data Lifecycle Management

Data protection during the full life cycle of information needs to be considered.



Cobit 5 – Enabling Information – p 33



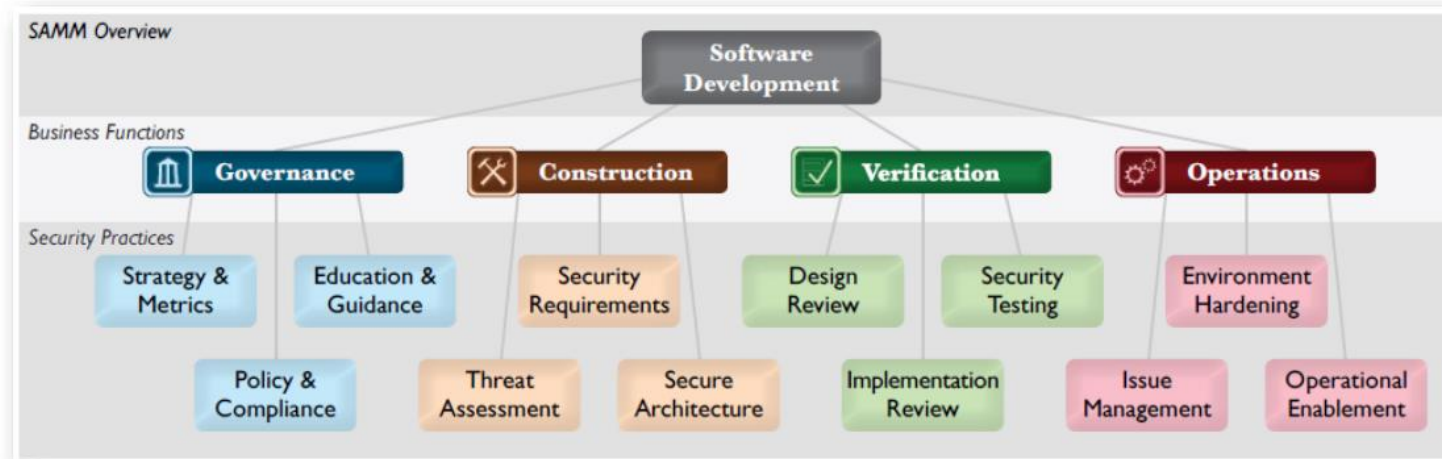
Security within the Data Lifecycle Management

	BIG DATA VALUE CHAIN	KEY PRIVACY BY DESIGN STRATEGY	IMPLEMENTATION
1	Data acquisition/collection	MINIMIZE	Define what data are needed before collection, select before collect (reduce data fields, define relevant controls, delete unwanted information, etc), Privacy Impact Assessments.
		AGGREGATE	Local anonymization (at source).
		HIDE	Privacy enhancing end-user tools, e.g. anti-tracking tools, encryption tools, identity masking tools, secure file sharing, etc.
		INFORM	Provide appropriate notice to individuals – Transparency mechanisms.
		CONTROL	Appropriate mechanisms for expressing consent. Opt-out mechanisms. Mechanisms for expressing privacy preferences, sticky policies, personal data stores.
2	Data analysis & data curation	AGGREGATE	Anonymization techniques (k-anonymity family, differential privacy).
		HIDE	Searchable encryption, privacy preserving computations.
3	Data storage	HIDE	Encryption of data at rest. Authentication and access control mechanisms. Other measures for secure data storage.
		SEPARATE	Distributed/ de-centralised storage and analytics facilities.
4	Data use	AGGREGATE	Anonymisation techniques. Data quality, data provenance.
5	All phases	ENFORCE/ DEMONSTRATE	Automated policy definition, enforcement, accountability and compliance tools.

Table 2: Privacy by design strategies in the big data value chain

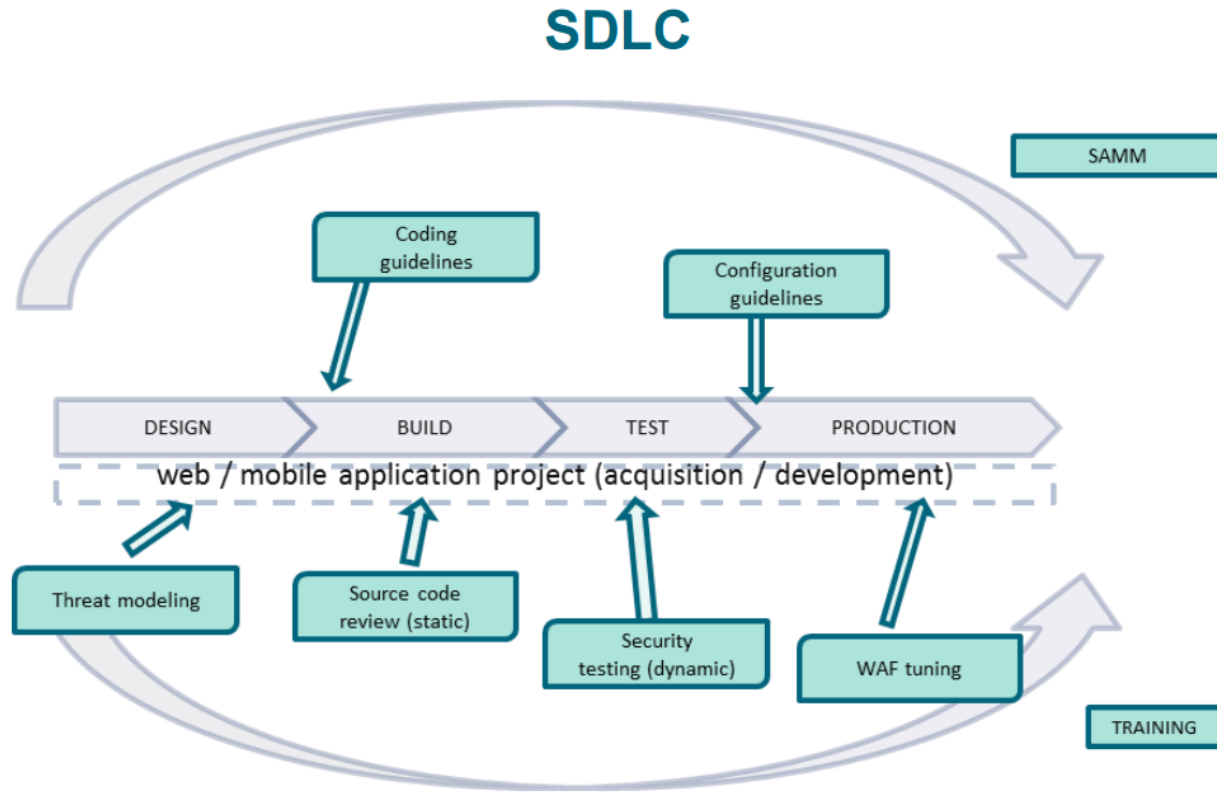
OWASP SAMM

- For each of the Business Functions, 3 Security Practices are defined
- The Security Practices cover areas relevant to software security assurance



<https://www.dp-institute.eu/wp/wp-content/uploads/2017/10/TOREON-Embedding-GDPR-into-the-SDL-C-Privacy-Caf%C3%A9-23-okt-2017.pdf>

OWASP SAMM & GDPR



Embedding GDPR into the SDLC (Toreon - Privacy Café 23 Okt 2017)

<https://www.dp-institute.eu/wp/wp-content/uploads/2017/10/TOREON-Embedding-GDPR-into-the-SDLC-Privacy-Caf%C3%A9-23-okt-2017.pdf>



Mapping GDPR / SAMM

SAMM Domains		GDPR Articles
SM	Strategy & Metrics	5, 24, 32, 33
PC	Policy & Compliance	7, 24, 32, (12-21)
EG	Education & Guidance	37, 39
TA	Threat Assessment	25, 35
SR	Security Requirements	24, 28, 32
SA	Secure Architecture	25
DR	Design Review	24, 25, 30, 32
IR	Implementation Review	24, 25, 32
ST	Security Testing	24, 25, 32
IM	Issue Management	33, 34, 39
EH	Environment Hardening	25, 33
OE	Operational Enablement	32, 33

Embedding GDPR into the SDLC (Toreon - Privacy Cafe 23 Okt 2017)

<https://www.dp-institute.eu/wp/wp-content/uploads/2017/10/TOREON-Embedding-GDPR-into-the-SDL-C-Privacy-Caf%C3%A9-23-okt-2017.pdf>

Anonymisation & Pseudonimisation

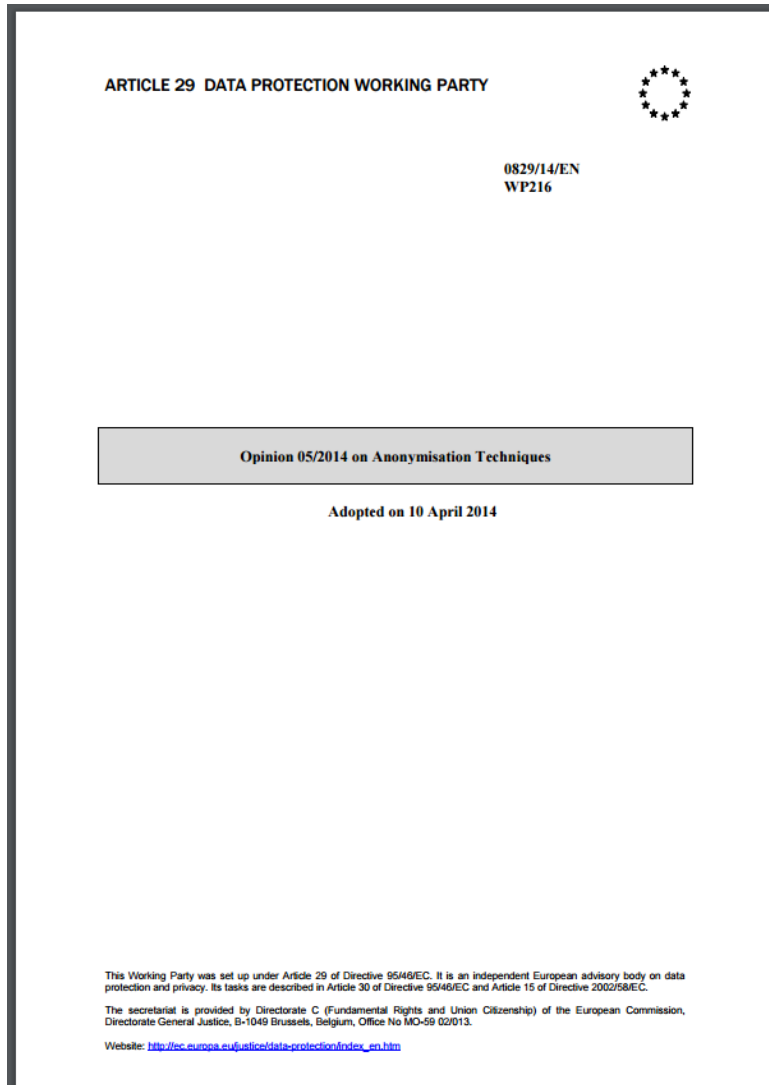


Table of contents

Legal Analysis

Lawfulness of the Anonymization Process

Risks of the Use of Anonymized Data

Robustness of Technologies

Different approaches to anonymization

Pseudonymisation

Conclusions & Recommendations

https://cnpd.public.lu/content/dam/cnpd/fr/publications/groupe-art29/wp216_en.pdf

Anonymization - Legal analysis

Analysis of the wording related to anonymization in the leading EU data protection instruments allows highlighting four key features:

- Anonymization can be a result of **processing personal data with the aim of irreversibly preventing identification** of the data subject.
- **Several anonymization techniques** may be envisaged, there is no prescriptive standard in EU legislation.
- Importance should be attached to contextual elements: account must be taken of “all” the means “likely reasonably” to be used for identification by the controller and third parties, paying special attention to what has lately become, in the current state of technology, “likely reasonably” (given the **increase in computational power and tools available**).
- A risk factor is inherent to anonymization: **this risk factor is to be considered in assessing the validity of any anonymization technique** - including the possible uses of any data that is “anonymized” by way of such technique - and severity and likelihood of this risk should be assessed.

Lawfulness of the Anonymization Process

Anonymisation is **a technique applied to personal data** in order to achieve irreversible deidentification.

Therefore, the starting assumption is that the **personal data must have been collected and processed in compliance with** the applicable **legislation** on the retention of data in an identifiable format.

In this context, the anonymisation process, meaning the processing of such personal data to achieve their anonymization, is an instance of “further processing”.

Anonymisation - Risks of the Use of Anonymized Data

When considering using anonymization techniques, data controllers have to take into account the following risks

- A specific pitfall is to consider pseudonymised data to be equivalent to anonymized data.

- Another negligence would also result from not considering the impact on individuals, under certain circumstances, by properly anonymized data, especially in the case of profiling.

The use made of datasets anonymized and released **for use by third parties** may give rise to a loss of privacy

Anonymisation - Robustness of Technologies

- **Singling out:** which corresponds to the possibility to **isolate some or all records which identify an individual** in the dataset;
- **Linkability:** which is the **ability to link, at least, two records concerning the same data subject** or a group of data subjects (either in the same database or in two different databases).
- **Inference:** which is the possibility to **deduce, with significant probability, the value of an attribute from the values of a set of other attributes**

Anonymisation - Different approaches to anonymization

Anonymization irreversibly destroys any way of identifying the data subject.

Randomization

Randomization is a family of techniques that **alters the veracity** of the data in order to remove the strong link between the data and the individual.

- Noise addition
- Permutation
- Differential privacy

Generalization

Generalization is the second family of anonymization techniques. **This approach consists of generalizing, or diluting**, the attributes of data subjects **by modifying the respective scale** or order of magnitude

- Aggregation and K-anonymity
- L-diversity/T-closeness

Pseudonymisation

Pseudonymization substitutes the identity of the data subject in such a way that additional information is required to re-identify the data subject

The most used pseudonymization techniques are as follows:

Encryption with secret key: the holder of the key can trivially re-identify each data subject through decryption of the dataset. Decryption can only be possible with the knowledge of the key.

Hash function: this corresponds to a function which returns a fixed size output from an input of any size (the input may be a single attribute or a set of attributes) and cannot be reversed;

Keyed-hash function with stored key: this corresponds to a particular hash function which uses a secret key as an additional input

Tokenization: this technique is typically applied in (even if it is not limited to) the financial sector to replace card ID numbers by values that have reduced usefulness for an attacker

Anonymization - Conclusions & Recommendations

Techniques of de-identification and anonymization are the subject of intense research, and this paper has shown consistently that **each technique has its advantages and disadvantages**. In most cases it is not possible to give minimum recommendations for parameters to use as each dataset needs to be considered on a case-by-case basis

In many cases, **an anonymized dataset can still present residual risk to data subjects**.

- Some anonymization techniques show inherent limitations
- **Each technique** described in this paper **fails to meet** with certainty **the criteria of effective anonymization**

The table below provides an overview of the strengths and weakness of the techniques considered in terms of the three basic requirements:

	Is Singling out still a risk?	Is Linkability still a risk?	Is Inference still a risk?
Pseudonymisation	Yes	Yes	Yes
Noise addition	Yes	May not	May not
Substitution	Yes	Yes	May not
Aggregation or K-anonymity	No	Yes	Yes
L-diversity	No	Yes	May not
Differential privacy	May not	May not	May not
Hashing/Tokenization	Yes	Yes	May not

Table 6. Strengths and Weaknesses of the Techniques Considered

Data Breach Management

Data Breach Management

Article 33

Notification of a personal data breach to the supervisory authority

1. In the case of a personal data breach:

- the controller shall without undue delay and, where feasible, not later than 72 hours after having become aware of it
- notify the personal data breach to the supervisory authority competent
- unless the personal data breach is unlikely **to result in a risk** to the rights and freedoms of natural persons

Data Breach Management

Article 34

Notification of a personal data breach to the supervisory authority

1. In the case of a personal data breach:

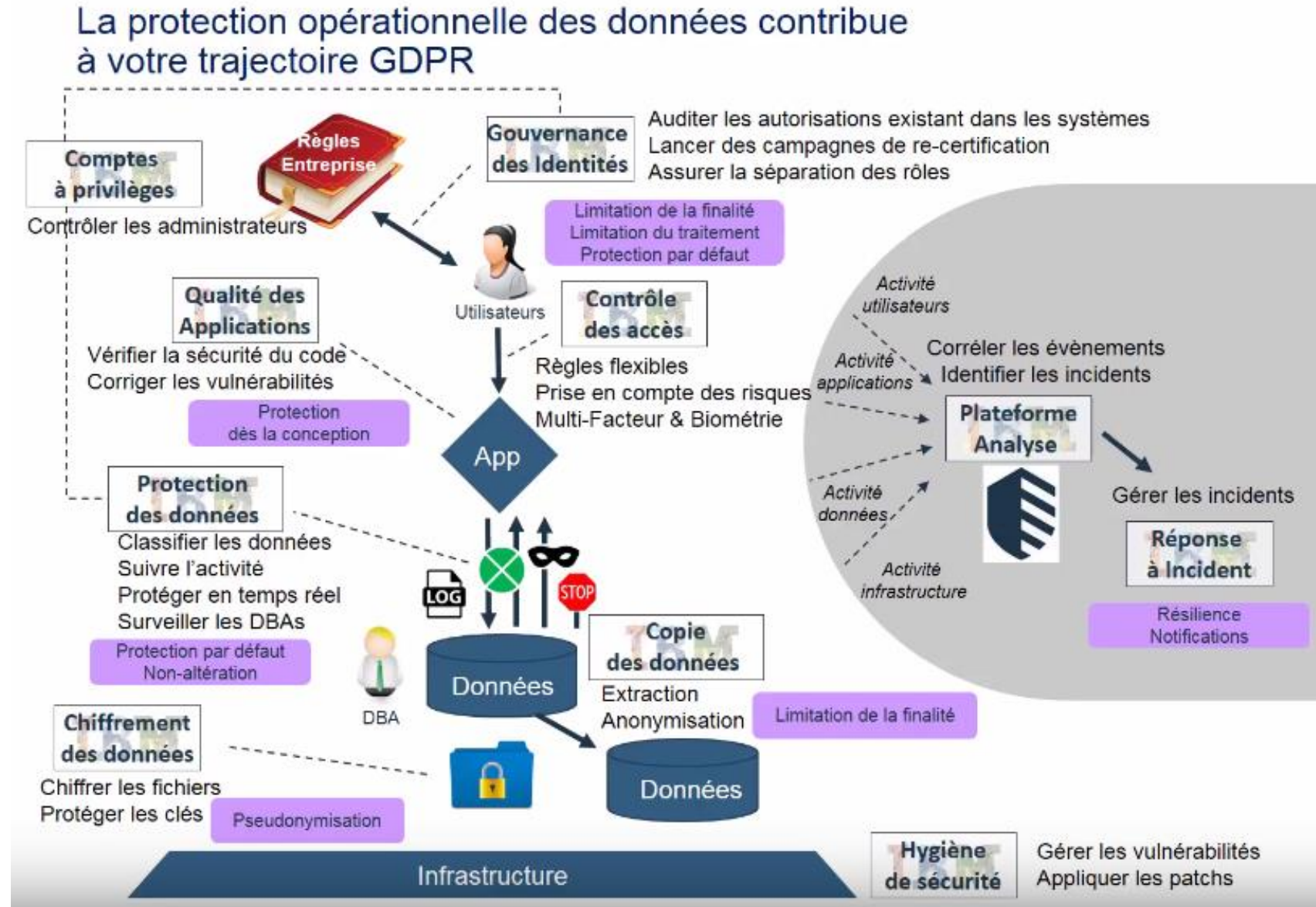
- Likely to result in **a high risk to the rights and freedoms** of natural persons
- the controller shall communicate the personal data breach to the data subject without undue delay.

3 The communication to the data subject referred to in paragraph 1 shall not be required if

- the controller has implemented appropriate protection measures, and those measures were applied to the personal data affected by the personal data breach, in particular those that render the personal data unintelligible to any person who is not authorised to access it, such as encryption;
- it would involve disproportionate effort. In such a case, there shall instead be a public communication or similar measure

Source: <https://www.cocc.edu/its/infosec/concepts/cia-triad/>

High level architecture



Privacy by design

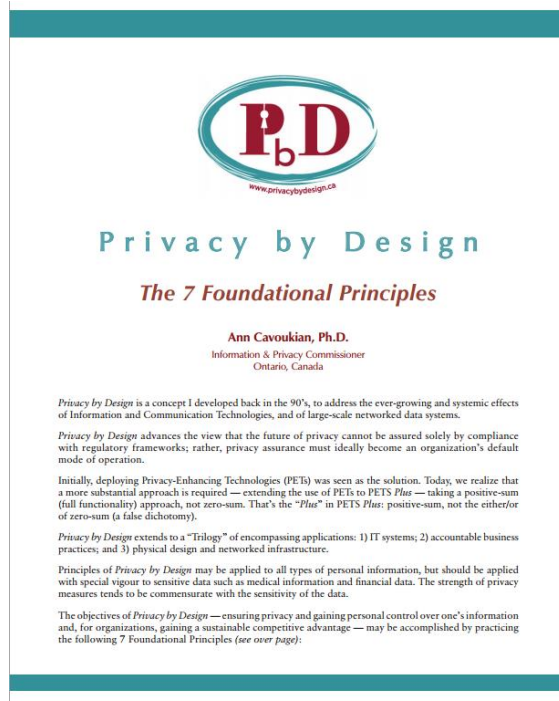
Privacy By Design

The principle “Privacy/data protection by design” is based on the insight that building in privacy features from the beginning of the design process is preferable over the attempt to adapt a product or service at a later stage.

The involvement in the design process supports the consideration of the full lifecycle of the data and its usage.

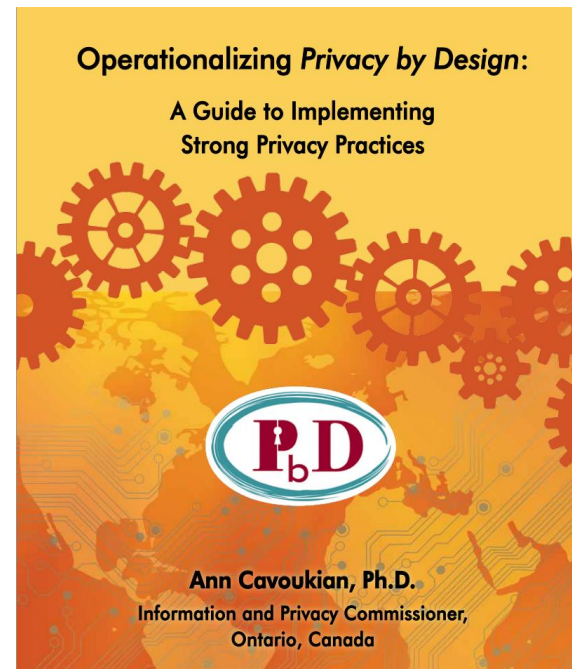
The principle “Privacy/data protection by default” means that in the default setting the user is already protected against privacy risks. This affects the choice of the designer which parts are wired-in and which are configurable. In many cases, a privacy-respecting default would not allow an extended functionality of the product, unless the user explicitly chooses it.

Privacy By Design



Privacy by design

<https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>



Operationalizing Privacy by Design

<http://www.ontla.on.ca/library/repository/mon/26012/320221.pdf>

Privacy By Design: Principle 1

Principle 1

Proactive not ***Reactive***; ***Preventative*** not ***Remedial***

Operational Guidance: *These actions anticipate and prevent privacy invasive events before they happen. Do not wait for privacy risks to materialize – the aim is to prevent the breaches from occurring.*

Actions	Responsibility
1. Affirm senior leadership commitment to a strong, proactive privacy program.	Leadership/Senior Management
2. Ensure that concrete actions, not just policies, reflect a commitment to privacy. Monitor through a system of regularly reviewed metrics.	
3. Develop systematic methods to assess privacy & security risks and to correct any negative impacts, well before they occur.	
4. Encourage privacy practices demonstrably shared by diverse user communities and stakeholders, in a culture of continuous improvement.	
e.g. Board of Directors, CEO, CPO, CIO, COO, CSO, Company Owner(s)	

Privacy By Design : Principle 2

Principle 2

Privacy as the *Default Setting*

Operational Guidance: These methods seek to provide privacy assurance – delivering the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. No action should be required on the part of the individual user to protect their privacy – it should be built into the system, automatically – by default.

Actions	Responsibility
1. Adopt as narrow and specific a purpose(s) for data collection as possible – begin with no collection of personally identifiable information – data minimization.	Software Engineers & Developers
2. Minimize the collection of data at the outset to only what is strictly necessary.	Application & Program Owners
3. Limit the use of personal information to the specific purposes for which it was collected.	Line of Business & Process Owners
4. Create technological, policy and procedural barriers to data linkages with personally identifiable information.	Line of Business & Process Owners

Privacy By Design : Principle 3

Principle 3

Privacy *Embedded* into Design

Operational Guidance: These actions embed privacy requirements into the design and architecture of IT systems and business practices. They are not bolted on as add-ons, after the fact. Privacy should be an essential component of the core functionality being delivered.

Actions	Responsibility
1. Make a Privacy Risk Assessment an integral part of the design stage of any initiative, e.g. when designing the technical architecture of a system, pay particular attention to potential unintended uses of the personal information.	Application & Program Owners
2. Base identity metasystems on the "Laws of Identity," intended to codify a set of fundamental principles to which universally adopted, sustainable identity architecture must conform.	Line of Business & Process Owners
3. Consider privacy in system development lifecycles and organizational engineering processes. System designers should be encouraged to practice responsible innovation in the field of advanced analytics.	Software Engineers & Developers
4. Embed privacy into regulatory approaches that may take the form of self-regulation, sectoral privacy laws, omnibus privacy legislation and more general legislative frameworks, calling for an approach guided by "flexibility, common sense and pragmatism."	Regulators

Privacy By Design : Principle 4

Principle 4

Full Functionality – *Positive-Sum*, not Zero-Sum

Operational Guidance: These actions seek to accommodate legitimate interests and objectives in a positive-sum, 'win-win' manner, not through a zero-sum (win/lose) approach, where unnecessary trade-offs to privacy are made. Avoid the pretense of false dichotomies, such as privacy vs. security – demonstrate that it is possible to have both.

Actions	Responsibility
1. Acknowledge that multiple, legitimate business interests must coexist.	Leaders/Senior Management
2. Understand, engage and partner – Practice the 3Cs – communication, consultation and collaboration, to better understand multiple and, at times, divergent interests.	Application & Program Owners Line of Business & Process Owners
3. Pursue innovative solutions and options to achieve multiple functionalities.	Software Engineers & Developers

Privacy By Design : Principle 5

Principle 5

End-to-End Security – *Full Lifecycle Protection*

Operational Guidance: Security is the key to privacy. These actions ensure cradle-to-grave, lifecycle management of information, end-to-end, so that at the conclusion of the process, all data are securely destroyed, in a timely fashion.

Actions	Responsibility
1. Employ encryption by default to mitigate the security concerns associated with the loss, theft or disposal of electronic devices such as laptops, tablets, smartphones, USB memory keys and other external media. The default state of data, if breached, must be “unreadable.”	Software Engineers & Developers Application & Program Owners
2. Deploy encryption correctly and carefully integrate it into devices and workflows in an automatic and seamless manner.	Line of Business & Process Owners
3. Ensure the secure destruction and disposal of personal information at the end of its lifecycle.	

Privacy By Design : Principle 6

Principle 6

Visibility and Transparency – Keep it Open

Operational Guidance: Stakeholders must be assured that whatever the business practice or technology involved, it is, in fact, transparent to the user, and operating according to the stated promises and objectives, subject to independent verification. Remember, trust but verify.

Actions	Responsibility
1. Make the identity and contact information of the individual(s) responsible for privacy and security available to the public and well known within the organization.	
2. Implement a policy that requires all “public-facing” documents to be written in “plain language” that is easily understood by the individuals whose information is the subject of the policies and procedures.	Leadership/Senior Management Software Engineers
3. Make information about the policies, procedures and controls relating to the management of Personal Information readily available to all individuals.	Application Developers
4. Consider publishing summaries of PIAs, TRAs and independent, third party audit results.	Systems Architect
5. Make available a list of data holdings of Personal Information maintained by your organization.	
6. Make audit tools available so that users can easily determine how their data is stored, protected and used. Users should also be able to determine whether the policies are being properly enforced.	

Privacy By Design : Principle 7

Principle 7

Respect for User Privacy **– Keep it *User-Centric***

Operational Guidance: This method requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice and empowering user-friendly options. Keep it user-centric.

Actions	Responsibility
1. Offer strong privacy defaults.	Leadership/Senior Management
2. Provide appropriate notice.	
3. Consider user-friendly options:	Software Engineers & Developers
a. Make user preferences persistent and effective.	Application & Program Owners
b. Provide users with access to data about themselves.	
c. Provide access to the information management practices of the organization.	Line of Business & Process Owners

Any questions



Alain Cieslik - ac@ictc.eu



Reference

Privacy By Design and Consent Management

- <https://www.youtube.com/watch?v=GDzYry6GCFg&feature=youtu.be&t=7>

Privacy By Design Strategies

- <https://blog.xot.nl/2012/09/10/eight-privacy-design-strategies/>
- https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design/at_download/fullReport
- <http://ec-wu.at/spiekermann/publications/Engineering%20Privacy.pdf>
- <https://www.cs.ru.nl/~jhh/publications/iwpe-privacy-strategies.pdf>
- <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-by-design/>

WP29 - Opinion 05/2014 on Anonymisation Techniques

➤ https://cnpd.public.lu/content/dam/cnpd/fr/publications/groupe-art29/wp216_en.pdf

Oracle security Architecture

➤ <http://www.oracle.com/technetwork/database/security/wp-security-dbsec-gdpr-3073228.pdf>

OWASP: Integrate GDPR into SAMM

➤ https://www.toreon.com/wp-content/uploads/2017/05/TOREON_Embedding_GDPR_into_the_SDLC_OWASP_AppSecEU_Sebastien_Siebe_V20170511.pdf