ISO/IEC JTC 1/WG 7
Working Group on Sensor Networks

Document Number:	N051
Date:	2010-07-05
Replace:	
Document Type:	Liaison Organization Contribution
Document Title:	Liaison Statement from JTC 1/SC 27/WG 5 to JTC 1/WG 7 on the ISO/IEC 1st CD 29101
Document Source:	JTC 1/SC 27/WG 5
Document Status:	For consideration at the 2nd WG 7 meeting in US.
Action ID:	FYI
Due Date:	
No. of Pages:	41

ISO/IEC JTC 1/WG 7 Convenor:

Dr. Yongjin Kim, Modacom Co., Ltd (Email: cap@modacom.co.kr)

ISO/IEC JTC 1/WG 7 Secretariat:

Ms. Jooran Lee, Korean Standards Association (Email: jooran@kisi.or.kr)

Committee Draft **ISO/IEC 1st CD 29101**	Reference number: ISO/IEC JTC 1/SC 27 **N8808**
Date: **2010-06-10**	Supersedes document SC 27 N8164

THIS DOCUMENT IS STILL UNDER STUDY AND SUBJECT TO CHANGE. IT SHOULD NOT BE USED FOR REFERENCE PURPOSES.

ISO/IEC JTC 1/SC27 Information technology - Security techniques Secretariat: Germany (DIN)	Circulated to P- and O-members, and to technical committees and organizations in liaison for voting (P-members only) by: **2010-09-10** Please submit your votes and comments via the online balloting application by the due date indicated.

ISO/IEC 1st CD 29101

Title: Information technology -- Security techniques – Privacy reference architecture

Project: 1.27.55 (29101)

Explanatory Report

Status	**SC 27 Decision**	**Reference documents**	
		Input	**Output**
Study Period (SP)	Resolution 30 of 17th SC 27 Plenary (N4599), Apr. 2005	JTC 1 Recommendation (4108=JTC1N7552) to assign responsibility to SC 27 in the area of privacy technologies	Call f. Contr. (N4616)
	Recommendation of Ad Hoc on Privacy (N4880), Nov. 2005.	SoContr (N4723)	Report of Ad Hoc Nov. 2005 (N4880)
NWIP	Recommendation of Ad Hoc on Privacy (N5186rev1), May 2006 and Resolution 1 of 18th SC 27 Plenary (N5199), May 2006	Report on Ad Hoc, Nov. 2005 (N4880rev1) & May 2006 (N5186rev1); Report of teleconf. (N4953rev1); DE NWIP (N5071)	Text f. NWIP (N5212)

NP 29101	Resolution 6 of 1st WG 5 meeting (N5513), Nov. 2006	SoV (N5289);
		SoContr. (N5332).
			DoC (N5222);
			Call f. contr. (N5521);
			Meeting Report (N5585).

	2nd WG 5 meeting May 2007, Resolutions 1 & 7 (N5873) May 2007 and Resolutions 1, 26 of 19th SC 27 Plenary (N5939), May 2007.	ES contr. (N5668)	2nd Call f. Contr. (N5883) due by 2007-09-01.

For details regarding project development at the Working Draft stage please see the text on the 2nd page.

1st CD 29101	9th WG 5 meeting, April 2010, resolutions 3, P3 (N8828rev) & SC 27 resolution 1 (N8916).	SC 37 com. (N8562);
		PICOS com (N870);
		SoCom (N8568).
			Liaisons to
			PICOS (N8839);
			SC 37 (N8847);
			DoC (N8807);
			Text f. 1st CD (N8808).

1st CD Registration and Consideration

In accordance with resolution P3 (in SC 27 N8828rev) of the 9th SC 27/WG 5 meeting held in Melaka (Malaysia), 19th – 23rd April 2010, the attached document SC 27 N8808 has been registered with the ISO Central Secretariat (ITTF) as a 1st Committee Draft (CD) and is hereby circulated for a 3-month 1st CD LB closing by

2010-09-10

Explanatory Report (2nd page)

Status	**SC 27 Decision**	**Reference documents**	
		Input	**Output**
1st WD 29101	3rd WG 5 meeting, Oct. 2007, resolutions 1, 8 (N6251)		Text f. 1st WD (N6259).
2nd WD 29101	5th WG 5 meeting, April 2008, resolutions 1, 8 & P 3 (N6726) & 20th Plenary, April 2008, resolut. 2 (N6799).	SoCom. (N6523); Proposed DoC (N6587).	DoC (N6735); Text f. 2nd WD (N6736).

3rd WD 29101	6th WG 5 meeting, Oct. 2008, resolutions 1, 8 (N7097rev1).	FIDIS com. (N7060); SoCom. (N7007);
		KR com. (N7219).
			Liaison to FIDIS (N7102);
			DoC (N7240);
			Request f. title change (JTC 1 N9549);
			Text f. 3rd WD (N7241).

4th WD 29101	7th WG 5 meeting, May 2009, resolutions 1, (N7724); 21st SC 27 Plenary, May 2009, resolution 2 (N7777).	FIDIS com. (N7543);
		SoCom (N7546);
		FR com. (N7547).
			JTC 1 endorsement of title change (N7457);
			Liaisons to PICOS (N7730);
			SC 37 (N7734);
			DoC (N7752);
			Text f. 4th WD (N7753).

5th WD 29101	8th WG 5 meeting, Nov. 2009, resolutions 1, 5, P2, P5, P11 (N8138).	SC 37 com. (N8045);
		PICOS com (N8072);
		SoCom (N8050).
			Liaisons to
			PICOS (N8150);
			SC 37 (N8141);
			DoC (N8163);
			Text f. 5th WD (N8164).

ISO/IEC JTC 1/SC 27 **N8808**

Date: 2010-06-10

ISO/IEC CD 29101

ISO/IEC JTC 1/SC 27/WG 5

Secretariat: DIN

Information technology — Security techniques — A privacy reference architecture

Élément introductif — Élément central — Élément complémentaire

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Document type: International Standard
Document subtype:
Document stage: (20) Preparatory
Document language: E

X:\TA3\TG3-3\NA043\NA043_Sekretariate\JTC1_SC27\03_Projekte\PROJECT_admin\29101_NP_Privacy_ref_architecture\03_01_1stCD_29101_20100610\N8808_1stCD_29101_20100610\ISO-IECJTC1-SC27_N8808_1stCD_29101_20100610.doc STD Version 2.2

Contents

Page

Figures

Tables

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 29101 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 27, *Security techniques*.

Introduction

This International Standard provides a high-level reference architecture for planning and building information and communication technology (ICT) systems that facilitate the proper handling of personally identifiable information (PII). This privacy reference architecture can be used as a best practice to build necessary privacy controls into an ICT environment in a way that is compatible with information security controls.

The privacy reference architecture provided within this International Standard:

a) provides a consistent, high-level approach to the implementation of privacy safeguarding requirements to safeguard the processing of PII in ICT systems;

b) provides guidance for planning, designing and building ICT system architectures that more effectively facilitate the privacy of individuals by preventing inappropriate use of an individual's PII; and

c) shows how privacy enhancing technologies (PETs) can be used to enhance the implementation of privacy controls.

ISO/IEC 29101 builds on the privacy framework provided by ISO/IEC 29100 which is intended to help an organization to define its privacy safeguarding requirements as they relate to PII processed by any ICT system in a data processing life cycle.

Information technology — Security techniques — A privacy reference architecture

1 Scope

1.1 Purpose

This International Standard provides a reference architecture that should guide individuals and organizations who specify, procure, architect, design, develop, implement, test, maintain, administer, and operate ICT systems on how to

- address privacy safeguarding requirements when processing PII,
- ensure the proper handling of PII within such ICT systems, and
- apply consistent architectural decisions to accomplish compliance with specific privacy safeguarding requirements, rules and regulations.

However, this International Standard is limited to the processing of PII in ICT systems.

This International Standard establishes:

- organizational provisions that should be established;
- PII protection mechanisms that should be integrated, and
- available PETs that should be used in privacy-enhanced ICT systems.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 29100 – *Information technology – Security techniques - Privacy framework*

3 Terms and definitions

For the purposes of this document, the majority of terms in ISO/IEC 29101 are used either according to their accepted dictionary definitions, specific privacy terminology defined in ISO/IEC 29100 or according to commonly accepted definitions that may be found in existing standards, ISO security glossaries or other well-known collections of security and privacy terms. Some combinations of common terms used in ISO/IEC 29101, while not meriting inclusion in this clause 3, are explained for clarity in the context where they are used.

4 Symbols and abbreviated Terms

The following abbreviations are common to ISO/IEC 29101:

ICT	Information and Communication Technology
IT	Information Technology
PET	Privacy Enhancing Technology
PII	Personally Identifiable Information

5 Overview of the privacy reference architecture

The privacy reference architecture provides guidelines on how to develop, implement and operate ICT systems with built-in privacy controls; is a resource containing a consistent set of architectural best practices for managing PII in ICT systems; and extends on the privacy framework derived from ISO/IEC 29100.

The privacy reference architecture depicted by Figure 1 shows the main elements (1) organizational provisions, (2) PII protection mechanisms and (3) privacy enhancing technologies that are necessary in order to implement privacy controls that are effective in a specific business process that includes the processing of PII.

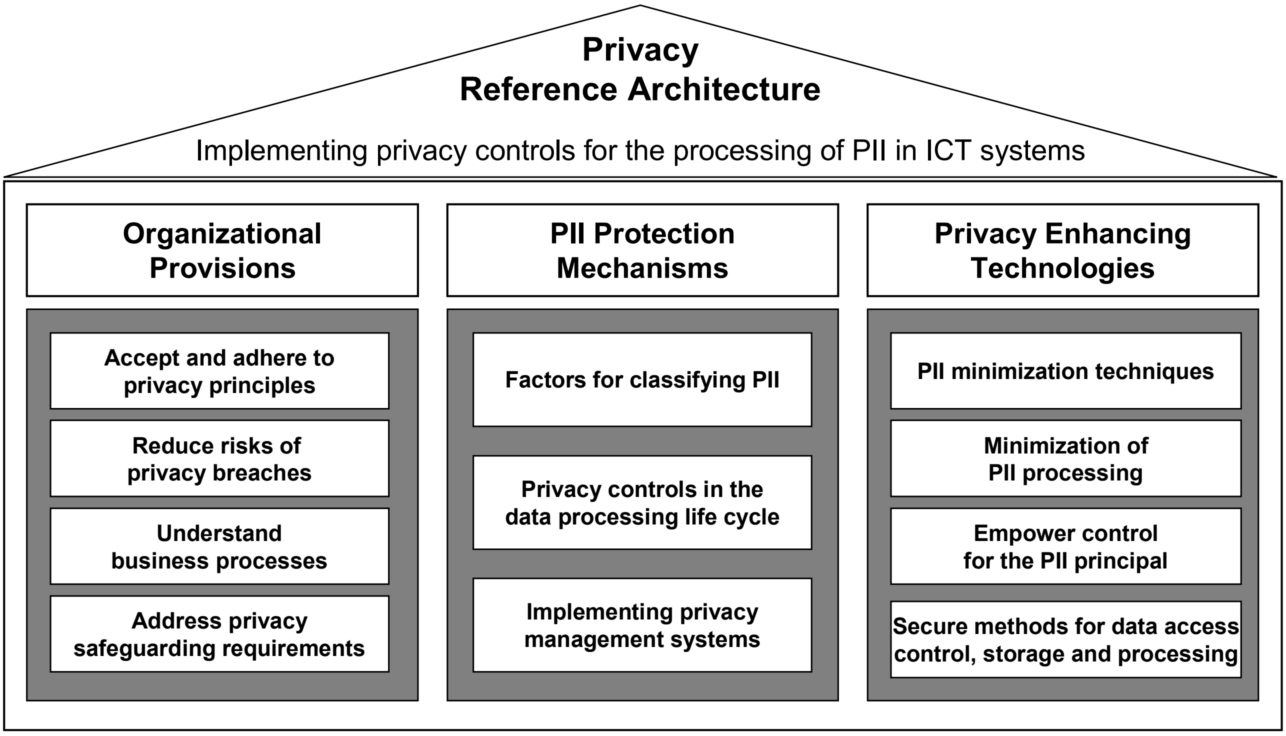
Privacy
Reference Architecture
Implementing privacy controls for the processing of PII in ICT systems
Organizational Provisions
PII Protection Mechanisms
Privacy Enhancing Technologies
Accept and adhere to privacy principles
Reduce risks of privacy breaches
Understand business processes
Address privacy safeguarding requirements
Factors for classifying PII
Privacy controls in the data processing life cycle
Implementing privacy management systems
PII minimization techniques
Minimization of PII processing
Empower control for the PII principal
Secure methods for data access control, storage and processing

Figure 1 – Main elements of privacy reference architecture

6 Organizational provisions

6.1 Accept and adhere to privacy principles

The PII controller is responsible for the protection of PII and the fair and lawful handling of it at all times, throughout the organization and for data handling processes outsourced to PII processors. Care in the processing of PII is essential to continued consumer confidence and good will.

Ultimately, the PII controller is responsible for implementing privacy controls in an ICT system. Privacy controls are intended to ensure that the privacy safeguarding requirements set for a specific user, transaction, and scenario are addressed and consistently fulfilled. Safeguarding controls could also be named safeguarding mechanisms or features but the word “controls” implies that the mechanism actually results in more control over PII and the fulfilling of privacy safeguarding requirements.

Examples of privacy safeguarding requirements and their respective controls are given in Annex B.3 in ISO/IEC 29100. Evidence of implementation is provided by properly documenting the privacy controls that are in place and having an internal or external auditor verify that the controls exist, have been properly implemented and are accurately documented.

Ultimately, the PII controller needs to accept and adhere to the privacy principles that are described in ISO/IEC 29100 and that form the basis for this International Standard. Table 1 lists these privacy principles.

Table 1 – The privacy principles of ISO/IEC 29100

1. Consent and choice
2. Purpose legitimacy and specification
3. Collection limitation
4. Data minimization
5. Use, retention and disclosure limitation
6. Accuracy and quality
7. Openness, transparency and notice
8. Individual participation and access
9. Accountability
10. Information security controls
11. Compliance

The following, self-imposed data handling or privacy rules provide a checklist of tasks that the PII controller should follow:

a) Obtain consent and provide choice should be achieved by establishing the following self-imposed data handling or privacy rules:

- inform the PII principal or its authorized agent in a meaningful way of the purposes for the PII processing;

- obtain the PII principal’s (or its authorized agent’s) consent before or at the time of collection whenever possible, and in a manner consistent with applicable law;

- where possible, actively provide choices for the PII principal or its authorized agent to exercise control over the purpose for which PII is used or not used (e.g., whether the PII can be used for secondary purposes such as marketing);

- generate and keep (secure) records of consent obtained from the PII principal or its authorized agent; and

- the PII principal should be informed of any change in context of the retention of their PII including change in ownership of the PII controller.

b) Identify the purpose with the following self-imposed data handling or privacy rules:

- Identify, at or before the time of collection, why the PII is needed and how it will be used at or before the time of collection;

- document why the information is collected and how it will be used;

- inform the PII principal or its authorized agent from whom the information is collected, why it is needed and how it will be used; and

- identify any new purpose for the information and obtain the PII principal's or its authorized agent's consent for this new purpose before using it.

c) Limit the collection of PII by implementing the following self-imposed data handling or privacy rules:

- require collection of only the PII that is necessary for the specified purpose;

- do not collect PII indiscriminately; and

- do not deceive or mislead individuals about the reasons for collecting PII.

d) Minimize data by implementing the following self-imposed data handling or privacy rules:

- minimize the use, creation, transfer, storage, or archiving of PII wherever possible, including, for example, in monitoring or logging processes.

e) Limit the use, retention and disclosure by implementing the following self-imposed data handling or privacy rules:

- use or disclose PII only for the purpose for which it was collected, unless the individual consents, or disclosure is authorized by law;

- keep PII only as long as necessary to satisfy the purpose;

- put guidelines and procedures in place for retaining and destroying or anonymizing PII;

- keep PII used to make a decision about an individual for a reasonable time period; and

- securely destroy or render anonymous in an effective manner information that is no longer required for an identified purpose or a legal requirement.

f) Ensure accuracy of the PII by implementing the following self-imposed data handling or privacy rules:

- take care that information is accurately captured, (e.g., example by validating it with the PII principal or its authorized agent);

- minimize the possibility of using incorrect information when using an individual's PII or when disclosing information to third parties; and

- periodically review PII holdings for accuracy.

g) Be open in your privacy management program by implementing the following self-imposed data handling or privacy rules:

- inform customers, clients and employees that you have policies and practices for the management of PII; and

- make these policies and practices understandable and easily available.

h) Give individuals access to their PII by implementing the following self-imposed data handling or privacy rules:

- when requested, inform authenticated individuals if you have any PII about them;

- give authenticated individuals access to their information;

- grant the PII principal the right to object, on request and free of charge, to the processing of PII relating to him;

- delete PII upon the request of authenticated individuals, if not prohibited by applicable law, contractual requirements or the circumstances of the original transaction;

- correct or amend any PII if its accuracy and completeness is challenged and found to be deficient; and

- communicate the reasons for any refusal to correct, amend or delete PII.

i) Be accountable for the processing of PII by implementing the following self-imposed data handling or privacy rules:

- appoint a person(s) to be responsible for your organization's compliance;

- protect all PII held by your organization or transferred to a third party for processing;

- develop and implement PII policies and practices;

- set up redress procedures; and

- build awareness among employees, third parties and the PII principal so that they are aware of the relevant privacy safeguarding requirements and their specific responsibilities for the proper handling of PII.

j) Use appropriate information security controls by implementing the following self-imposed data handling or privacy rules:

- control PII against unauthorized access, disclosure, copying, use or modification;

- protect PII regardless of the format in which it is held; and

- implement international standards (e.g., ISO/IEC 27001) that contain specific requirements for information security that may similarly be applied to PII handling and are, therefore, recommended for entities handling PII.

k) Assure compliance by implementing the following self-imposed data handling or privacy rules:

- fulfil external requirements relevant to the data handling or privacy rules such as requirements described in applicable legislation, relevant employee or consumer protection laws, or other rules and regulations pertaining to the protection of PII;

- develop a set of internal requirements relevant for adhering to corporate governance and compliance management rules;

- implement internal control systems that can monitor and assure the adherence to self-imposed data handling or privacy rules (e.g., logging access to and alteration of sensitive PII);

- create complaint and redress mechanisms; and

- establish a regular process of frequently assessing the risk landscape around data handling processes and the handling of PII and updating existing rules.

6.2 Reduce risks of privacy breaches

PII and the processing of PII should be protected against the possible consequences of a privacy breach. PII controllers should determine the appropriate privacy safeguarding requirements depending on these factors, so privacy controls can be implemented within the PII controller’s ICT systems, if necessary.

Privacy risk assessments should determine risks levels for a particular instance of PII. Risk levels such as “low, moderate, and high” indicate the potential negative impact or the potential harm that could result to the PII principal and the PII controller from a privacy breach. The PII controller can also determine risk levels that have a broader range, such as a 5-point scale, but they should include at least the three levels “low, moderate, high”.

Potential harm includes any adverse effects that would be experienced by a PII principal who’s PII was the subject of a privacy breach, as well as any adverse effects experienced by the PII controller that maintains the PII. Harm to a PII principal includes any negative or unwanted effects (i.e., that could be embarrassing and also physically or financially damaging). Some examples of privacy-invasive activities that could cause harm to PII principals are listed in Table 2. PII controllers could also experience harm including but not limited to administrative burden, financial losses, loss of public reputation and public confidence, and civil liability.

Table 2 – Examples of privacy-invasive activities

Privacy invasive activities
Placement of Adware/Spyware Appropriation Blackmail Breach of Confidentiality Cyber Crime Harming Data Integrity Discrimination Extortion Unwanted Exposure Fraud Identity Theft Intrusion Loss of Control Loss of Data Misuse of Data Phishing Sexual Solicitation Spamming Unauthorized Telemarketing Unauthorized Third Party Sharing

PII controllers should reduce such privacy risks by regularly performing privacy risk assessments. The outcome of those assessments should result in specific risk minimization and mitigation activities including the implementation of adequate privacy and security safeguarding controls. In addition, privacy enhancing technologies should be evaluated and applied where relevant.

Determining PII risk levels should take into account relevant factors. Several important factors that PII controllers should consider are described below. It is important to note that relevant factors should be considered together; one factor by itself might indicate a low risk level, but another factor might indicate a high risk level, and thus override the first factor. Also, the risk levels suggested for these factors are for illustrative purposes; each instance of PII is different, and each PII controller has a unique set of requirements and a different mission. Therefore, PII controllers should determine which factors are specific to their own organization and should create and implement policy and procedures that support these determinations.

Relevant factors to consider when determining PII risk levels include:

a) Distinguishability,

b) Data field combinations,

c) Context of use,

d) Obligation to protect, and

e) Access to and location of the PII

These factors are explained in more detail in sub-clauses of section 7.1 on the classification of PII. Regular privacy risk assessments should be conducted to understand areas in which the PII controller faces particular risks in PII processing.

6.3 Understand business processes

Business processes and the respective PII that is processed need to be understood in order to find an appropriate technical solution model that can fulfil the identified privacy safeguarding requirements. For the purpose of this international standard, the term 'business process' refers to any operational process of a commercial, not-for-profit or state-run organization. Visualizing core business processes and corresponding PII should assist in establishing a common understanding of developing or implementing privacy enhancing ICT systems. The business organization that processes PII should use tools with which they are familiar, e.g., spreadsheets, lists or graphical software to chart processes to encompass all relevant business processes. Formal establishment and maintenance of business and data process models or inventories should conform to local privacy and data protection requirements applicable for that organization. Once such a business process model has been established and it is compared to the data process model, privacy controls such as functionalities for receiving consent, PII categorization and tagging functionalities, audit and logging procedures, retention timeframes, or necessary notices and security alarms can be determined and compared to the necessary privacy safeguarding requirements.

6.4 Address privacy safeguarding requirements

Privacy safeguarding requirements are influenced by the following factors: (1) legal and regulatory factors for the safeguarding of the individual's privacy and the protection of his/her PII, (2) contractual factors such as industry regulations, professional standards, company policies, (3) business factors predetermined by a specific business application or in a specific use case context and (4) other factors that can affect the design of ICT systems and the associated privacy safeguarding requirements. Privacy safeguarding requirements build the basis for an organization's privacy policies and data protection procedures.

Within the privacy reference architecture, organizations need to determine the relevant privacy safeguarding requirements for PII processing for specific individuals and in a specific business or data processing context considering the underlying legal requirements.

The following, general privacy considerations should be contemplated to ensure that PII can be managed and accounted for in ICT applications:

f) It should be possible to pull together definitive records of what PII is known/stored (including PII recorded in logs and backup);

g) It should be possible to identify and describe all PII, no matter how it is collected (including internal generation);

h) A list of individuals who have or may have had access to PII should be maintained;

i) The real need for routine audit logging that includes or creates PII should be assessed as part of system design;

j) The design of privacy controls should include the security of backed up and archived data when it contains PII;

k) The same privacy controls which are used to secure collected PII should be applied to secure any derived PII (e.g. transaction histories in e-retailing applications, behavioural PII used for profiling);

l) Memory dumps and other line or activity traces created by engineers should be subject to security and privacy policies;

m) The supporting data architecture should define the extent to which data assets are shared across entities and a data placement process should govern case-by-case decisions related to data sharing and re-use in the context of the data architecture and an accountability structure.

n) The need and capacity to record privacy inquiries (e.g. to resolve disputes) should be considered in applications based on databases containing PII;

o) Change management procedures for web applications should protect against inadvertent changes that affect privacy, e.g., ensure that changes to web forms do not change the nature or amount of PII collected or ensure that PII must be secured by the destruction of obsolete and defective equipment or media;

p) The design of privacy controls should include the management of ephemeral or incidental PII, e.g. help desk or customer service logs.

7 PII protection mechanisms

7.1 Factors for classifying PII

7.1.1 General

Data processing flow models should be developed as an integral component of a privacy risk assessment and are also the basis for being able to classify PII. The data processing flow cannot only show the areas where PII is collected, transferred, used, stored or disposed of but can also visualize areas where the PII has a certain level of sensitivity or importance and, as a consequence, requires the implementation of stronger safeguarding measures. Classifying data into PII and non-PII is the minimum requirement at this stage but various industries may also require the classification of PII into subsets of categories that need special protection schemes, e.g. certain health data of an individual that requires specific protection.

Possible factors for classifying PII could be the following:

a) Risk level

b) Distinguishability

c) Data filed combinations

d) Context of use

e) Obligation to protect

f) Access to and the location of the PII

These factors and how they can be used to classify PII is described in the following sub-clauses.

7.1.2 Risk level

PII should be handled based on its risk level. Organizations should determine the PII risk level so that additional privacy safeguarding mechanisms can be implemented, if necessary. This sub-clause outlines factors for determining the PII risk level for a particular instance of PII. The PII risk level – low, moderate, or high – indicates the potential harm that could result to the PII principal and the organization if the PII were inappropriately processed. These three risk levels are defined by NIST FIPS 199 Standards for Security Categorization of Federal Information and Information Systems.

The adverse effects of a privacy breach in the form of harm should be considered when attempting to determine which risk level corresponds to a specific set of PII. Harm for the purposes of this document, includes any adverse effects that would be experienced by an individual whose PII was the subject of a privacy breach, as well as any adverse effects experienced by the organization that maintains the PII. Harm to an individual includes any negative or unwanted effects (i.e., that may be embarrassing and also physically or financially damaging). Examples of types of harm to individuals include, but are not limited to, the potential for blackmail, identity theft, discrimination, or emotional distress. Organizations may also experience harm from a privacy breach – including but not limited to administrative burden, financial losses, loss of public reputation and public confidence, and civil liability.

Determining the PII risk level should take into account relevant factors. Several important factors that organizations should consider are described below. It is important to note that relevant factors should be considered together; one factor by itself might indicate a low risk level, but another factor might indicate a high risk level, and thus override the first factor. Also, the risk levels suggested for these factors are for illustrative purposes; each instance of PII is different, and each organization has a unique set of requirements and a different mission. Therefore, organizations should determine which factors, including factors specific to the organization, they should use for determining PII risk levels and should create and implement policy and procedures that support these determinations.

7.1.3 Distinguishability

Organizations should evaluate how the PII can be used easily to distinguish particular individuals. For example, PII composed of individuals' names, fingerprints, and social security numbers uniquely identify individuals, whereas PII composed of individuals' phone numbers only would require the use of additional data sources (e.g., phone directories), and would only allow some unique individuals to be identified (e.g., unique identification might not be possible if multiple individuals share a phone or if a phone number is unlisted). PII composed of only an individual's area code and gender would not generally allow any unique individual to be identified dependent on the uniqueness of the value in the dataset. PII that is easily distinguishable may merit a higher risk level than PII that cannot be used to distinguish individuals without unusually extensive efforts.

Organizations may also choose to consider how many individuals can be distinguished from the PII. Breaches of 25 records and 25 million records may have different impacts, not only in terms of the collective harm to individuals but also in terms of harm to the organization's reputation and the cost to the organization in addressing the breach. For this reason, organizations should carefully determine if the amount or the type of PII that is processed represents a particularly high risk level.

7.1.4 Data field combinations

Organizations should evaluate the sensitivity of each individual PII, as well as the sensitivity of a set of PII together. For example, an individual's financial account number is generally more sensitive than an individual's phone number or zip code, and the combination of an individual's name and social security number is less sensitive than the combination of an individual's name, social security number, date-of-birth, mother's maiden name, and credit card number. Organizations may also consider certain combinations of PII to be more sensitive, such as name and credit card number, than each data field would be considered without the existence of the others.

7.1.5 Context of use

Context of use is defined as the purpose for which the PII is processed as well as how that PII is used or could potentially be used. When ICT designers consider the specification of data required for a business process, the risks of its use in another context should be considered. These risks could result from loosely specified access privileges or changes in governance rules. Examples for the context of use include, but are not limited to, performing a statistical analysis, determining the eligibility for benefits or the administration of benefits, research, tax administration or law enforcement. Organizations should assess the context of use because it is important to understand how the disclosure of data elements can potentially harm individuals and the organization. Organizations should consider what harm is likely to be caused if the PII is disclosed (either intentionally or accidentally) or if the mere fact that the PII is being collected or used is disclosed could cause harm to the organization or individual. For example, law enforcement investigations could be compromised if the mere fact that information is being collected about a particular individual is disclosed.

The context of use may cause multiple instances of the same type of PII to be assigned different PII risk levels. For example, suppose that an organization has three lists that contain the same PII (e.g., name, address, phone number). The first list is people who subscribe to a general-interest newsletter produced by the organization. The second list is people who have filed for retirement benefits, and the third list is individuals who work undercover in law enforcement. The potential impacts to the affected individuals and to the organization are significantly different for each of the three lists. Based on context of use only, the three lists are likely to merit risk levels of low, moderate, and high, respectively.

Examples of topics that are relevant to the context of use as a factor for determining PII risk levels are the following: abortion; use of alcohol, drugs or other addictive products; illegal conduct; illegal immigration status; information that could be damaging to the financial standing, employability or reputation of an individual; information leading to social stigmatization or discrimination; political opinions or religious beliefs; psychological well-being or mental health; and other information due to specific cultural or societal factors.

7.1.6 Obligation to protect

Some organizations are subject to special laws or regulations that govern the obligation to protect PII (e.g., financial services or insurance firms). Other organizations may be obliged to protect PII by their own policies, standards, or management directives. These obligations may impact the determination of PII risk levels.

7.1.7 Access to and location of the PII

Organizations should take the nature of authorized access to the PII into consideration. The more often PII is accessed and the more people have access rights to PII, the more likely are privacy breaches. Copying of PII should be minimised to ensure all copies are current and correct. Where multiple copies exist, the number of people authorised to access the PII and the frequency of access will be the aggregate of the figures for each copy. Another factor to be considered is the level of direct control an organization has over the PII being processed. If the PII, for example, is accessed via teleworkers' systems or mobile systems outside the control of the organization, an organization should assign a higher risk level to the PII.

Additionally, organizations should take into consideration whether PII that is stored or regularly transported off-site by employees should be assigned a higher risk level. For example, surveyors, researchers, and other field employees often need to store PII on laptops or removable media as part of their jobs. PII located offsite is more vulnerable to unauthorized access or disclosure because it is more likely to be lost or stolen than PII stored within the physical boundaries of the organization.

7.2 Privacy controls in the data processing life cycle

7.2.1 General

One important element of this International Standard is the integration of privacy controls in every step of the data processing life cycle. The safeguarding of an individual ICT system user's privacy is often seen as a secondary element in designing ICT systems and gets added later, for example, by simply adding a security or data protection scheme, protecting the individual's PII from unauthorized use. However, for assuring the privacy of their customers and system users, businesses should implement privacy controls as a primary element in every phase of the data processing life cycle when setting up systems.

In order to implement effective privacy controls in an organization, transparency on the data processing flows and the respective PII that is processed need to be created. Data processing flow diagrams help to describe the PII in business processes, e.g. human resources, production, marketing, and sales. Data processing flow diagrams are a graphical representation of the "flow" of PII through the ICT system and between the different actors. If the organization transfers PII to external entities of the business or to external service partners, e.g., to processing agencies, the data processing flow diagram should include those data flows. Data processing flow diagrams need to be set up when PII is exchanged and collaborative processes occur with third parties (i.e. external partners). In that case, these external parties that have access to PII should be made aware of their obligations in a formalized manner, e.g. by setting up third party agreements. One possible format for a data processing flow diagram can be to construct a data flow table. This diagram or table follows the collection, transfer, use, storage or disposal of PII and includes information such as, e.g., the type of PII, the PII risk level, who the PII was collected by, the purpose for processing, to whom the

PII is going to be transferred the receipt of consent by the PII principal, the period for which it is obtained and processed and at which location it will be stored.

The steady increase of data availability and the rapid decline in the cost of storage have created a situation in which control over data – and in this case specifically personally identifiable information – becomes difficult to achieve. Without control over data, it will be difficult to fulfil the expectations that ICT users have when it comes to the handling of their PII and it will be impossible to comply with laws and regulations that require more control. Therefore, it is very important to understand privacy within each phase of the data processing life cycle.

7.2.2 Phases of the data processing life cycle

7.2.2.1 Collection

Many entities collect information from individuals. When collecting PII, entities should do so in a manner that respects the privacy preferences and legal rights of the PII principal and privacy safeguarding requirements as stated by applicable law. Information collected should be recorded as associated with the individual in order to ensure that the information belongs to them, including where necessary obtaining third party corroboration. Consent should be obtained from the PII principal to collect the information in association with the terms provided to the individual, including providing a means of contacting the PII controller or processor to request further explanation about any activities that the individual is unclear about. Consent to store and process PII should be obtained from the individual via a prominent notice that discloses how the information will be used. Prominent notice and consent are particularly important when the individual might not expect their information to be used in a particular manner, such as when data the individual has chosen to store with the PII controller will be scanned and used to target advertising. The unauthorized collection of PII without legal justification could be harmful as it invades the privacy of the individual (e.g., by means of surveillance technology).

In addition, data collection processes should be designed to only collect PII that is necessary for the respective transaction and should not collect the PII in an unsolicited form (e.g., web application forms are sometimes designed without such mechanisms). The entry of unsolicited PII can be minimized by context specific display of input fields reducing or eliminating areas in the web form where this information could be entered (e.g., removing unnecessary check boxes and free text fields). In addition, the use of fields with predefined entries (e.g., list boxes and drop-down lists), containing non-PII options, should be considered. When a text field is necessary, the User Interface (UI) should discourage the PII principal from entering PII in scenarios where the recipient intends to collect only non-PII.

7.2.2.2 Transfer

Transferring, disseminating, or releasing PII to others, sometimes also referred to as disclosure, means that data is moving further away from the control of the PII principal. Accountability and responsibility for the transferred data should be agreed upon and maintained by each party involved in_the data processing. This agreement should be in writing where required by applicable law. The transfer of sensitive PII should be avoided unless it is necessary to provide a service that the individual has requested, or unless it is required by law. Some jurisdictions have instituted laws that specifically require formal contractual agreements that include all privacy safeguarding requirements between the involved parties when PII is transferred outside the jurisdiction that has a prescribed level of privacy protection. Cross-border transfers are very common in operating information communication systems nowadays. For that reason, more attention should be given to measures controlling the transferred data even though it might not be in the hands of the original PII recipient anymore.

7.2.2.3 Use

Using data means any form of PII processing that does not include “collect” “transfer” “store” “archive” or “dispose”. The privacy principles of ISO/IEC 29100 limit the processing of PII in ways that are incompatible with the originally defined purposes; applicable law may do so as well. Where such processing is considered necessary, the consent of the PII principal should be obtained unless otherwise allowed by law. In some cases, the PII recipient should give notice to the PII principal about the specific use of the data.

The processing and modifying of PII for secondary purposes is a critical issue. Such secondary uses could include using the PII for marketing purposes. The ISO/IEC 29100 Privacy Principles (and some data protection and privacy legislation) state that PII may not be processed in a way that is incompatible with the originally identified purposes without the consent of the PII principal. Organizations wishing to process PII in such a manner should provide prominent notice that such processing is intended and should obtain the specific consent of the affected PII principals for the new purposes.

7.2.2.4 Storage

Data is stored in many different forms and in many different places. In order to fulfil privacy safeguarding requirements, it may be necessary to store data in such a way that it can be identified as PII. Therefore, documentation should be associated with the data, such as specific tags that mark, for example, the purpose that it can be used for, the consent given and any specific sensitivities that should be observed (e.g., certain data categories should be encrypted or deleted after a certain period of time). While tagging PII provides the ability to automate policy enforcement, it could also create the risk of unauthorized harvesting of this information through automated means. Therefore, safeguarding controls should be implemented wherever data is tagged as PII or wherever PII is marked with additional information concerning the PII principal. Given these risks, alternative means for documenting PII other than tags should also be considered.

Storing sensitive PII on a system should be avoided when not absolutely necessary. When it is necessary to store sensitive PII, the specific and unambiguous consent of the PII principal should be collected taking into account specific measures where required by applicable law. In this case, the data should be stored only for the shortest amount of time necessary to achieve the specific business purpose. Sensitive PII should be stored with appropriate controls and mechanisms to prevent unauthorized access, modification, destruction, removal, or other unauthorized use.

7.2.2.5 Disposal

In the final stage of the data processing life cycle, data gets deleted, anonymized, archived, destroyed, returned or disposed of in some other way. Specific data within data records might get locked from unauthorized use by marking it for disposal. It should be noted that deleting data does not necessarily mean that the data is ultimately disposed of because data deleted in information systems can often be recovered. Although it might seem to be an obvious task in data handling, procedures concerning disposal of PII sometimes do not comply with privacy safeguarding requirements. If data contains PII and it is locked from unauthorized use by marking it for disposal, the PII should be anonymized before it is locked. Specifications given by the PII principal (e.g., usage purpose) or specifications given by legislation (e.g., expiration date for specific PII) should be considered before PII is disposed of. The PII controller should implement controls in storage systems to dispose of PII when it expires or when the purpose for the storing or processing of the data is no longer valid. If applicable laws or other obligations require the retention of PII for a defined period of time, the relevant data should be locked and safely stored to protect it against further use apart from the circumstances defined by the underlying law, or other legal obligation such as one specified in a binding legal agreement.

Archived data needs careful attention especially when PII is involved. The privacy principles state that PII should be retained only as long as necessary to fulfill the stated purposes, and then be securely destroyed or anonymized. However, if the PII recipient is required by applicable law to retain PII after the other purposes has expired, the data should be locked, (i.e. archived and be exempted from further usage). The primary considerations in archiving PII are to ensure that the appropriate data protection mechanisms are in place, including access management solutions that provide access to archived PII only to authorized individuals.

7.3 Implementing privacy management systems

The use of a privacy management system enables a PII controller and PII processor to more effectively meet its privacy safeguarding requirements using a structured approach. This structured approach also provides a PII controller and PII processor the ability to measure outcomes and continuously improve its privacy effectiveness. An effective privacy management system impacts people, processes and technology, is part of the internal control program and risk mitigation strategy of an organization, and its implementation helps to satisfy compliance with data protection and privacy regulations.

The following are key considerations when establishing a privacy management system:

Table 3 – Key considerations for establishing privacy management systems

Key factors for establishing privacy management systems	**Description**
Policies	The PII recipient should set a clear policy concerning the processing of PII, aimed at maintaining internal and external requirements, that is binding for every supervisor and every employee handling PII. Every procedure with respect to these tasks is to be evaluated for its compliance with the defined policy. The PII recipient should also set clear policy and procedures concerning the collection, transfer, usage, storage, archiving and disposal of data.
Inventory	The PII recipient should establish a process to categorize any incoming data as regular data, PII data or sensitive PII data, thus allowing for a separation of PII data at the earliest possible time. The PII recipient should also create and maintain an inventory of all PII processes. By this means the appropriate handling of PII can be ensured from the start.
Procedures and controls	Technical and organizational guidelines for privacy should be developed and implemented with regard to the aforementioned privacy principles.
Governance	In order to ensure that privacy principles are adhered to and controls satisfy the specified privacy safeguarding requirements, a governance and internal control authority should be established. The internal control authority should establish regular privacy risk assessments as well as key privacy areas to be audited regularly within the organization's internal and external audits to ensure compliance with privacy rules and regulations and, if included in the overall governance scheme, compliance with this International Standard. The person responsible for all processes involving the handling of PII could either take on the internal control authority him/herself or assign the authority to another function such as the internal audit department or an external auditor. Even if the internal control authority is assigned to an external auditor, the organization remains accountable for the proper safeguarding of the PII.
Compliance	Entities that receive and process PII should develop and maintain privacy risk assessments: • to evaluate compliance with data protection and privacy legal and regulatory requirements, • to evaluate relevance and effectiveness of internal policies, procedures and control, • to resolve privacy issues that could be of potential concern.

Documentation	Records are needed for dispute resolution, auditing, and other privacy management purposes. Records should be generated when consent is obtained from PII principals, when PII is received and/or altered, and when PII is passed to another PII recipient. Records should be protected against unauthorized access, alteration, and deletion (ISO/IEC 27002 provides relevant guidance). Additional records could be needed to meet local legal or auditing requirements. The need for retaining records should be balanced against any additional privacy risk retaining such records may create.
Training and awareness	Organizations should ensure their employees, vendors and other individuals are aware of their responsibilities when processing PII.

Privacy education, training, and proactive communication on the respective legal and regulatory, contractual, and business requirements that influence the privacy safeguarding requirements should be part of a broad and consistent privacy management program implementation. Those communication mechanisms should also address regular updates in the organization’s internal privacy policies and procedures. This includes publishing easily accessible and simple-to-use complaint procedures and conducting regular audits and privacy risk assessments to assure consistent safeguarding of PII.

8 Privacy enhancing technologies

Once the particular risks to the privacy of the PII principal are known and the Data processing flows are determined, privacy enhancing technologies (PET) can support the protection of the individual's privacy by eliminating or reducing PII or by preventing unnecessary and/or undesired processing of PII.

PET are a major contribution in addressing and solving some of the privacy safeguarding requirements identified in a business process that involves PII. PET have been defined as a coherent system of ICT measures that protect privacy by eliminating or reducing PII or by preventing unnecessary and/or undesired processing of PII; all without losing the functionality of the data system. A number of physical, personnel, technical, and administrative mechanisms can be called PET and they should be considered when implementing privacy controls in ICT systems.

Examples of PET that could be used include the following:

- techniques to anonymize, pseudonymize, or unlink PII from the otherwise identifiable PII principal;
- technical means to classify data as PII and to associate the PII with specified purposes;
- protecting PII in transit and storage using a variety of encryption techniques;
- automated notifications on privacy risks (e.g., before PII is submitted through unprotected communication channels);
- tools to prevent the recording of data traces (e.g., when surfing particular web sites);
- tools to support transparency and the individual's participation or that in other ways help the individual to exercise his/her legal and/or contractual rights;
- privacy-friendly default settings to allow users to voluntarily and specifically consent to the processing of PII (e.g., in social networking services);
- using specific privacy seals or data audit certifications to assess and denote the appropriate protection levels;
- automated mechanisms to communicate privacy policies and have the PII principal match them with their own privacy preferences;
- the collection and/or storage of PII in encrypted or secret-shared forms that allow the use of privacy-preserving data mining techniques such as secure multi-party computation; and
- the use of specialized cryptographic techniques or protocols such as biometric encryption and limited show blind signatures to add privacy to more complex protocols.

For a better understanding, PET can be classified according to their main objectives, e.g., minimizing PII, minimizing PII processing, empowering control for the PII principal and applying secure methods for data access control, storage and processing.

8.1 PII minimization techniques

8.1.1 Pseudonymization

Pseudonymized or de-identified records are those that have had enough PII removed or obscured such that the remaining information does not directly identify an individual and where there is no reasonable basis to believe that the remaining information could be used to identify an individual. Pseudonymized records can be re-identified and made personally identifiable by using a code, algorithm or pseudonym that is assigned to each record. A common technique for pseudonymization is to use a one-way cryptographic function such as a hash function.

To be an effective mechanism for minimizing privacy risk, the re-identification code, algorithm or pseudonym should be maintained in a separate system, with appropriate controls in place to prevent unauthorized access to the re-identification information. Additionally, the data elements in the pseudonymized records should not be linkable via public records or other reasonably available external records in order to avoid rendering the records personally identifiable.

For example, if an organization wishes to analyze trends in the purchasing habits of its customers, it may pseudonymize the records in its database prior to performing the analysis in order to reduce the privacy risks associated with handling the data. If the database contains names, credit card numbers, items purchased and dates of purchase, a copy of the database could be made where the names have been deleted and a keyed one-way function or a random function has been used on each credit card number. The resulting data set can be passed to the individuals performing the data analysis as the risk of identifying the credit card holders is significantly reduced. Note that the key of the one-way function or the mapping used in the random function must be stored securely.

Another example is using health care test results in research analysis. All of the distinguishable PII fields can be removed, and the patient ID numbers can be obscured using pseudo-random data that is linked to a cross-reference table located in a separate system. Access to the cross-reference table, the only means by which the original (complete) PII records can be reconstructed, should be limited to authorized individuals only.

Additionally, pseudonymous or de-identified data can be aggregated for the purposes of statistical analysis, such as making comparisons, analyzing trends, or identifying patterns. An example is the aggregation and use of multiple sets of de-identified data for evaluating several different types of education loan programs. The data describes characteristics of loan holders, such as age, gender, region, and outstanding loan balances. With this dataset, an analyst could draw statistics showing that a particular number of women within a certain age range have outstanding loan balances greater than a specified amount. Although the original data sets contained distinguishable identities for each individual and is considered to be PII, the de-identified and aggregated dataset would not contain linked or readily distinguishable data for any individual. The ability to link an individual to such distinguishable data is also sometimes called linkability. Linkability describes a situation in which an attacker can sufficiently distinguish whether two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) within the system (comprising these and possibly other items) are related or not.

It must be noted that it may be possible to infer the identities associated with individual records by combining other attributes. This attack is not easy to automate and the attacker may require additional knowledge about the PII principals whose data is in the database. Pseudonymized data should be evaluated and if the risk of identification through inference is considered high then the data should not be published. In these cases, the use of more secure PII processing techniques may be justified. Pseudonymization and secure multi-party computation may be considered.

8.1.2 Anonymization

Anonymized data is separate and distinct from pseudonymized data with respect to re-identifiability. When PII is anonymized, a re-identification algorithm, code or pseudonym does not exist or has been removed such that it is no longer available. In addition, data is typically anonymized by using statistical disclosure limitation techniques to ensure the data cannot be re-identified. Examples of these techniques include:

- generalizing the datamaking information less precise, such as grouping continuous values or replacing categorical values with broader terms;
- suppressing the data—deleting an entire record or certain parts of records that would render it identifiable;
- introducing noise into the data—adding small amounts of variation into selected data such as weight, height or age;

- swapping the data—exchanging certain data fields of one record with the same data fields of another similar record (i.e., swapping the postal codes of two records);
- replacing data with the average value—replacing each value of data within clustered records with the average value for the entire group of data.

Using these techniques, the information is no longer PII, but it can remain useful. For example, anonymized information can be used for system testing without requiring all of the safeguards necessary when PII is used for testing purposes.

An anonymization service takes PII (i.e., data that includes fields which name or unambiguously describe an identifiable individual), removes all personal identifiers, and ensures that the resulting anonymized data is computationally hard to reverse through a de-identification process.

The following issues should be carefully considered to ensure that anonymization services operate securely and effectively:

a) the mechanisms for processing PII fields to achieve anonymization. This includes both the deletion of a PII principal's name fields as well as the processing of fields such as birth date or postal code to replace the content of such fields by mapping field content to a narrow range of values such as broad age categories (in the case of birth date) or broad geographic areas where many individuals live (in the case of postal codes).

b) assessment of the risk of identification through inference[1];

c) techniques and methodologies for mitigating risk of re-identification in small samples;

d) trusted third party involvement in anonymization;

e) controlled re-identification practices where policy permits (or legislation requires) such practices.

Privacy preferences may require PII to be anonymized to the greatest extent possible to serve an identified purpose (i.e. identifiable information should not be used when anonymous information will serve the purpose). In addition, PII should be retained only as long as necessary to fulfill the identified purpose, and then securely destroyed or anonymized. The anonymization service should be used to enhance the individual's privacy for the purpose of secondary uses and disclosures of PII, such as in research data, in public health monitoring and assessments that draw their data from the individual's health record, or for some marketing purposes.

8.1.3 Unobservable data management

Unobservability ensures that a system user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used. Unobservability requires that users and/or subjects cannot determine whether an operation is being performed. Unobservability is like a real time equivalent of unlinkability but the difference lies in the fact that the objective of unobservability is to hide an entity's use of a resource, rather than the entity's identity.

In the context of PII processing, it is appropriate to use unobservable data management to create data analysis systems where the PII processor is not fully trusted. The PII controller can set up a data analysis system that allows the PII processor to perform aggregations and other data mining tasks without access to the individual records in the database.

[1] Such an assessment is usually done by calculating the so-called cell size of each record in the resulting anonymized or pseudonymized data set

The primary goal of such a system is to protect the PII principal by not disclosing identifiable information to the PII processor. This can be achieved by running a “sandbox-like” system where the PII processor sends queries and only the final answer is disclosed to the processor. The queries must be restricted to make sure that the minimal amount of information is disclosed. The sandbox may be implemented as a database system with a limited query interface. More security can be added by using secure multi-party computation. Multi-party computation may also be required to achieve another security goal that is relevant in research settings. The PII processor may not want the PII controller to observe which filtering parameters are used in the aggregations as this may leak scientific results.

8.2 Minimization of PII processing

Query restriction is a technique concerned with the protection of PII processed in data mining. The technique facilitates the provision of PII to third parties or public organizations for data mining purposes while not risking the misuse of PII and not jeopardizing the precision of the data mining algorithms.

Methods of query restriction include restricting the size of query results, controlling the overlap amongst successive queries, keeping audit trails of all answered queries and constantly checking for possible compromises, suppression of data cells of small sizes, and clustering entities into mutually exclusive atomic populations.

8.3 Empower control for the PII principal

Consent is a pivotal area for the privacy reference architecture. Consent is provided by the PII principal and needs to be informed, explicit, non-repudiable and auditable. There are various methods of obtaining consent and some of them are briefly discussed here in order to provide context for the consent services that might encompass them:

a) Blanket consent: consent which is asked for as a blanket license to distribute personal data to multiple agencies or destinations may not fall within the “informed consent” guidelines. Blanket consent is not a valid method to obtain consent and can result in a serious privacy breaches. Additionally, blanket consent may imply consent for processes or data destinations not envisioned by the PII principal.

b) Pre-consent: consent typically through an intermediary. This is consent granted where another system, not the owner of the data, is interacting with the user to obtain their consent for a data transfer. This is acceptable, so long as the intermediary is not the ultimate data recipient. So the intermediary may hold the data temporarily of the purposes of consent.

c) Post consent: consent after a transfer has taken place, to enable the transaction to be finalised and validated. Not an acceptable consent mechanism.

d) Consent granter by an intermediary: This is problematic, as the agency with stewardship of the data must trust the intermediary to correctly transfer the consent form the user, and the user must also trust the intermediary to pass on the consent correctly. In some cases the user must consent also for the intermediary to view the data, even if the data is simply passing through the system and will not be stored. Alternatively, the data can be securely encrypted so that only the source and destination parties can view the data.

The two typical design patterns for consent are a simple consent service and a consent intermediary/privacy broker. In a simple consent service, the PII principle gives consent to the PII controller to undertake some activity with the PII belonging to the principle. In a consent intermediary/privacy broker service, the act of consent is delegated or relayed through a third party.

8.4 Secure methods for data access control, storage and processing

8.4.1 Biometric encryption

Biometric encryption allows the use of a biometric sample to encrypt or code some other information, like a PIN or account number, or cryptographic key, and then store the biometrically encrypted code, not the biometric sample itself. The biometric encryption process removes the need for public or private sector organizations to collect and store actual biometric information in their database. Thus, privacy and security concerns associated with the creation of centralized databases can be addressed. Biometric encryption allows an individual's biometric data to be transformed into multiple and varied identifiers for different purposes so that these identifiers cannot be correlated with one another. In addition, if a biometric identifier is compromised, a completely new one may be generated from the same finger or iris of an individual.

8.4.2 Secret sharing

Secret sharing refers to a method for distributing a secret amongst a group of participants, where the secret can only be reconstructed when the predetermined combination of participants co-operate. Typically, in a secret sharing scheme there is one dealer, n players, and a secret that needs to be shared and distributed among those n players. A specific combination of participants who is eligible to recover a secret is set. A dealer generates n shares from the secret and distributes them between the participants. The shares from the predetermined combination of participants induce the secret. A typical condition for this predetermined combination is any k participants from n can recover the secret, and such a scheme is called a k-out-of n secret sharing scheme.

A secret sharing scheme can secure a secret over multiple servers and remain recoverable despite multiple server failures. Assume a system administrator or a software component wants to recover the secret even when there are up to k failures among n servers. Then he generates n shares using an (n-k)-out-of-n secret sharing scheme, and stores each share on one server. This way he can recover the secret even when there are failures in k servers. At least (n-k) servers have to cooperate to recover the secret value. For a lesser number of servers this task is infeasible.

Secret sharing can be used to store and process PII with excellent privacy guarantees. Assume that a number of organizations take the role of the players in the secret sharing scheme. Then the data can be collected directly in secret shared form and processed using secure multi-party computation. Such a system can have strong provable security guarantees that are hard to achieve with other techniques. For an example on how to apply secret sharing and secure multi-party computation in practice refer to the Annex.

8.4.3 Secure multi-party computation

Secure multi-party computation (MPC) is a private computation technique for evaluating functions on secret inputs. Amongst the existing MPC techniques, two-party protocols and multi-party protocols can be distinguished.

a) Two-party protocols: There are various cryptographic methods for performing two-party computations. Cryptocomputing, for example, performs computations on ciphertexts using homomorphic encryption. Other techniques can use different cryptographic primitives. Two-party protocols are suited for client-server systems. For example, a client encrypts a secret query and sends it to a server system that performs computations and returns a result without learning the contents of this query.

b) Multi-party protocols: In some cases, computations involve more than two parties. In this case, a well-studied solution is to use secret sharing to distribute the input secrets the parties. After the secrets are distributed, the parties engage in cryptographic protocols to compute output secrets based on the inputs while preserving the privacy of input data. Shares representing the outputs are published at the end of computation so that the result can be learned.

Secure multi-party computation fits well in scenarios where multiple parties are involved in the processing of information. The technique is especially suitable for aggregating data from various sources and building shared private data warehouses with data mining capabilities.

Note that different MPC techniques require different security guarantees. Protocols that are secure in the passive model provide privacy when the parties do not deviate from the protocols. However, protocols that are secure in the presence of an active adversary can detect deviations. Also, the number of parties involved can vary. Since a large number of parties increase the communication complexity of the system, selecting a smaller number of representatives to carry out the computation is preferable.

Annex A
(informative)
Use Cases

The following use case examples should illustrate how the concepts discussed throughout this International Standard can be applied.

A.1 User-centric information sharing service for government agencies

Government agencies spend considerable resources in collecting, verifying, and maintaining information that is already authoritatively known to other parts of government. Multiple investments across government entities and a reliance on secondary documents to move information between agencies via individual citizens are the result. For both parties, the government on the one hand, and for individual citizens on the other, inconvenience, extra costs, and delays in government services is the consequence. From a risk perspective, these multiple data processes also represent problems to assuring the privacy for each individual citizen.

Applying the privacy reference architecture for this use case, it first should be determined what PII principles are involved, e.g. individual citizens requesting government services and government agency employees using specific databases. The resulting privacy safeguarding requirements need to be articulated and converted into functional requirements that can generate a high-level design of the use case. The following questions need to be answered:

a) Who are the parties involved?

b) What are the roles of the different accessing parties and what access rights to they need for what type of data?

c) What are the roles and functions the government agencies or government individuals have?

d) What approach should be implemented to share PII between government agencies?

The following figure shows a high-level design for the privacy reference architecture suggested to be extended to the corresponding privacy controls to be implemented at each system component.

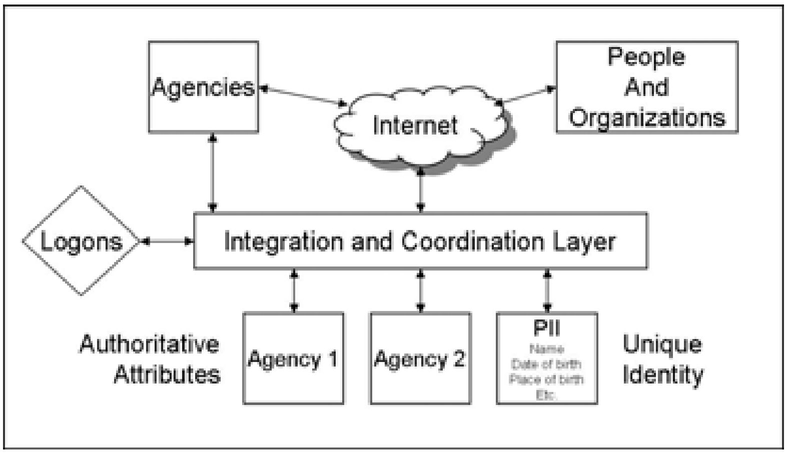

Figure 2 – High-level design "User-centric information sharing service for government agencies"

A.2 User-centric information sharing in online communities

The popularity of social networks is due in part to the convenient and easy way in which individuals can interact with one another. Trust is established over time by openly display information about one another. This type of information sharing, however, raises concerns about the safeguarding of the individual's privacy, especially, because social networks do not usually offer a rich set of privacy controls to protect the individual.

In a social network setting, the stakeholders are the individual members, although in the case of a centrally hosted community, the community operator has legal obligations that can require them to have access to members' personal information. The architecture addresses various privacy problem domains which collectively give individuals a higher level of control and, therefore, confidence (trust) over visibility and use of their PII.

The architecture is based on operating principles that govern key areas, i.e., law, trust, privacy, control and identity, which together represent the core values of the architecture. Each principle is influenced by the chosen trust model and the choice of the model strongly directs the topology of the solution, e.g., centrally or peer administered. The choice of trust model is based on the agreed level of trust between individuals and/or between individual and the community operator. In some situations, it may be appropriate to host part of the community functionality on a personal and trusted platform that is owned and controlled by the individual, e.g., in the case of a self-managing or peer-based community.

Architecture features are included to promote trust-building (e.g., reputation) and enhance awareness (e.g., privacy advisor). Users are provided with the ability to be open or private, using partial identities as a means of identification where appropriate. These features distinguish the architecture from those employed by existing online communities.

A.3 Example architecture for privacy enhanced community services

The example architecture described here is designed for maximum flexibility and is, therefore, essentially topology-agnostic. It is a service-orientated design with services targeting the community member in the first instance but supporting inter-member relationships and community management.

Achievement of privacy is dominated by the use of a concept called partial identities. Partial identities provide community members with the ability to operate anonymously while at the same time ensuring that other community members and the community operator (which in a peer-to-peer configuration may simply be a collection of other members) are confident (have trust) in the integrity of others and can fulfil legislative requirements. Partial identities offer conditional anonymity but support law enforcement and a desire for enhanced trust and openness between members. The condition element of conditional anonymity is governed by a trusted authority which can be an external independent body, a trusted collection of members or an outsourced hosting entity.

Individuals mainly establish trust using the specially designed reputation mechanisms, although the openness and informative style of the architecture also help. These same mechanisms help the whole community understand any risks associated with sharing personal information and as such raise the level of trust throughout the community and between individuals. Privacy respecting reputation ensures that despite members being allowed to have multiple (partial) identities, they remain accountable through a single private overarching identity. The reputation management features of the architecture satisfy the subjective nature of reputation. They provide a reputation defining mechanisms but do not set thresholds for trustworthy member behaviour.

A privacy advisor further enhances member privacy by checking member activities in real-time, by (1) looking for evidence of activities that may undermine the member’s attempt to remain private, and (2) by educating the member when the subtitles of a member’s actions may expose sensitive personal information. The privacy advisor acts solely on behalf of and is loyal to a community member, except where member actions are not in the best interest of the community as a whole, or where the action may be illegal within the jurisdiction(s) where the community operates. In this respect, the privacy advisor is truly personal.

Members preserve privacy by interacting with one another in private ‘rooms’ (they can also interact in a more ad-hoc public manner but must acknowledge some loss of privacy). Private rooms allow members to share content in a controlled way, restricting readership with regard to the reputation and privacy concerns of targeted members. Content can be anything from simple messages through to multi-media attachments. The potential for the member to accidentally reveal private information about themselves during this type of exchange is minimised by the use of the privacy advisor.

The architecture considers the full lifecycle of membership activity, from registration with the community, interaction with other members, use of shared facilities, and ultimately concerns that arise when a member terminates membership of a community but leaves personal artefacts behind.

The example architecture was build upon a set of principles which in turn defined features that are derived form typical requirements for community services including mobile communities. Features are translated into system components that operate at varying levels of abstraction, and which together form a simple component hierarchy. The components can be grouped into modules covering the functionality of the community services in 10 modules: registration, multiple (partial) identities, access management, privacy advisor, social presence, reputation, revocation, external services, sub-communities and content sharing.

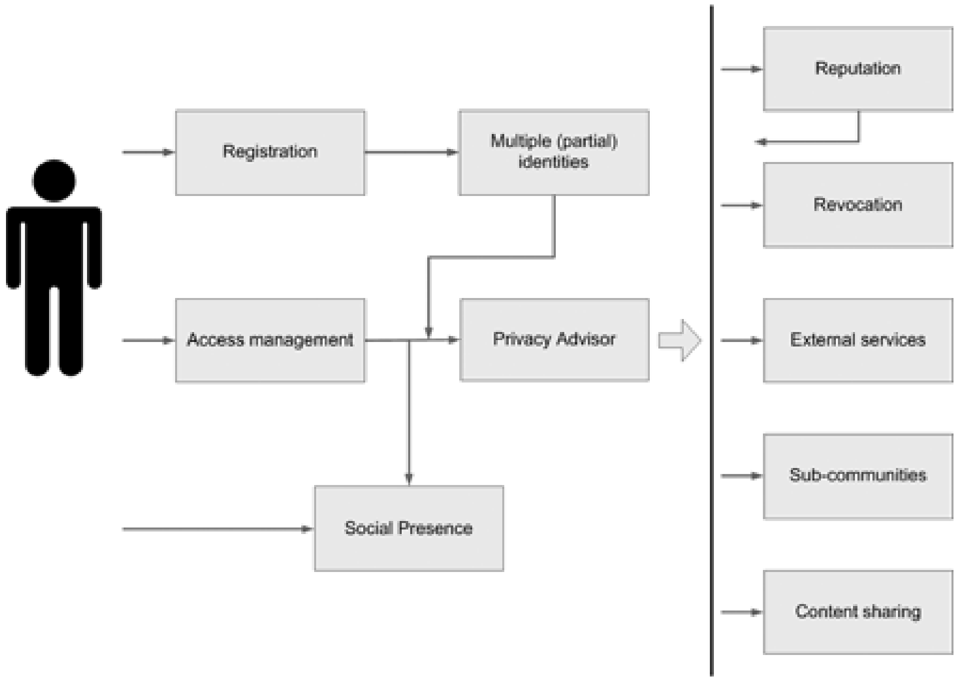

Figure 3 – The example architecture for privacy enhanced community services.

The overall architecture is shown in the figure 3 with the following elements:

- The social presence module consists of the location and event logging components and cooperates with the profile management component.
- The privacy advisor module is based on the service selection component and receives also input from the consent management and reputation management components.
- The registration module receives input from the user and is interacting with the policy management, profile management and consent management components.
- The multiple (partial) identities module consists of the partial identity management and anonymisation components and closely interacts with the profile management, reputation management and event logging components.
- The access management module covers the authentication and authorisation components and cooperates with the profile management, social presence, consent management and event logging components.
- The reputation module is based on reputation management, feedback management and partial identity management components and it interacts with the personal profile management and event logging components.
- The revocation module consists of the revocation, policy management and personal profile management components. The module also notifies the event logging, reputation management and content sharing components.
- The external services module uses the partial identity management and external service delivery components to provide for external services. Events are logged via the event logging component.

- The sub-communities module is based on the profile management, sub-community management, delegation and consent management components. Events are logged via the event logging component.

- The content sharing module is built on the import/export, secure repository and content sharing components with the help of notification, event logging and personal profile management components.

A.4 The hybrid architecture for privacy-preserving applications

The flow of information within an application may contain both PII and non-identifying (public) data. If no privacy enhancing technologies are used, both kinds of data are stored and processed in the same computation environment. In such a scenario, individuals with access to the computing hardware may also be able to access the PII. Therefore this environment is called the public computation environment. PII and public information interact freely in this environment. Figure 4 illustrates such an application.

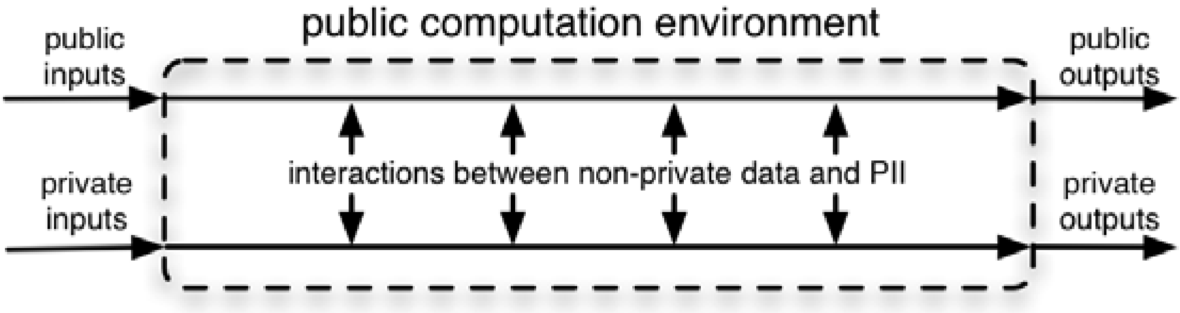
public computation environment
public inputs
private inputs
interactions between non-private data and PII
public outputs
private outputs

Figure 4 – Processing private data in a public computation environment

The hybrid architecture requires applications to process PII and public data in separate computation environments. This also means that the interactions between PII and public data have to be more strictly controlled. Private inputs can affect public outputs only in such ways that do not allow PII to be reconstructed from these outputs. Figure 5 illustrates an application with a hybrid architecture.

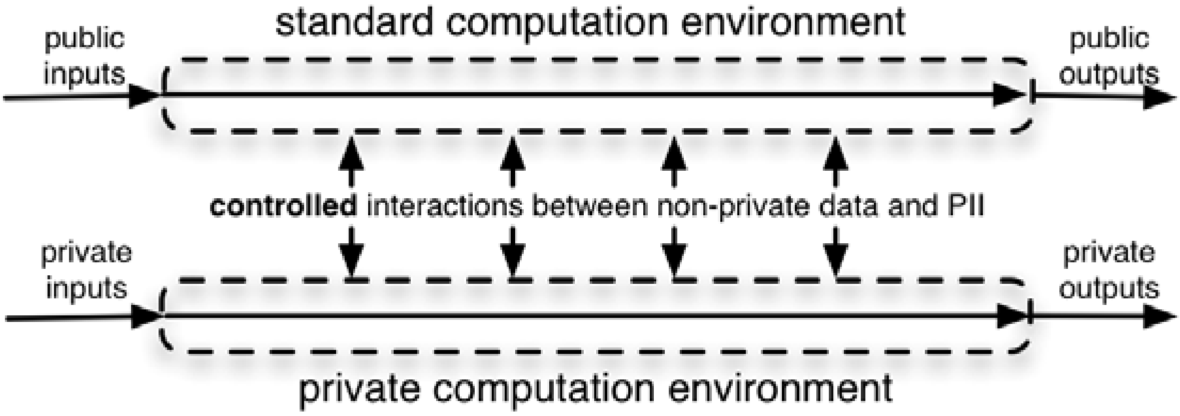
standard computation environment
public inputs
public outputs
controlled interactions between non-private data and PII
private inputs
private outputs
private computation environment

Figure 5 – Processing private data in an application using the hybrid architecture

Both the public and private computation environments have to be defined with regard to specific PII. If the organization processing the PII is authorized to process the data, but not to share it with others, then its local servers are a private computation environment but rented servers (e.g., in a computing cloud) are a public computation environment. On the other hand, if the organization should not witness the PII, its local servers are a public computation environment and a private computation environment has to be implemented using privacy enhancing technologies like anonymization or secure multi-party computation.

A.5 A data aggregation system based on share computing

This architecture presents a pattern for applying share computing to create data aggregation systems like survey systems, reporting tools and data warehouses. This architecture is suitable under the following conditions:

a) PII is collected from several individuals;

b) the output of PII processing is a report of aggregated values and trends;

c) at least three organizations can be identified who are interested in the results;

d) for performance reasons it is preferable if the input dataset does not contain more than a hundred thousand records.

Note, that the last assumption may be void by the emergence of more powerful share computing systems.

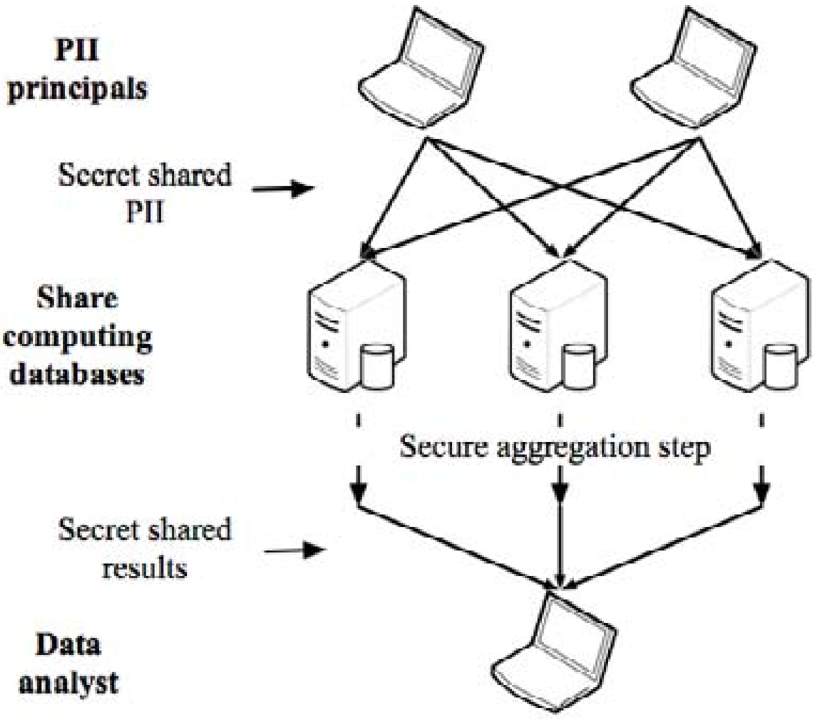

Figure 6 – Example of a data aggregation system based on share computing

The organizations who participate in the computation will run a database system with secret shared storage and share computing capabilities. The organizations will be compelled not disclose their database contents to anyone. This becomes easier if the organizations value the privacy of PII principals.

An electronic data entry system can be constructed that performs the secret sharing step under the control of PII principal. The data will be distributed into shares and the shares will be securely transferred to the organizations running the share computing databases. This technique empowers the PII principal by providing a guarantee that unless the majority of the organizations in the system disclose their shares, nobody will be able to see the values provided by the individual.

The secret shared PII can now be processed using secure multi-party computation techniques. The computation results will be in secret-shared form and the computation process will not leak anything about the PII. The shares of the aggregated results can be published to the data analyst who reconstructs them into analysis results. Figure 6 gives an overview of the architecture of such a system.

This system has the following security features that are hard to ensure with other techniques.

a) The individual values of PII will not be witnessed by anyone except the PII principal.

b) The organizations hosting the share computing databases have a significantly reduced risk of insider attacks as the share databases reveal nothing about the individual PII values.

c) Should the share database of any given party become compromised or stolen, a special multi-party computation procedure called resharing can be used to compute new shares of the PII so that even when more parties become compromised, the risk to PII is minimal.

If the proper protocols are used, the system will be able to perform unobservable data management with provable security guarantees.