# Privacy Impact Assessment: The Foundation for Managing Privacy Risk

By: William Stallings, PhD

Wednesday, March 17, 2021

**PREPARED BY:**

CSIAC Cyber Security & Information Systems Information Analysis Center

*Operated by Quanterion Solutions Incorporated*

266 Genesee Street
Utica, NY 13502

https://www.quanterion.com
qinfo@quanterion.com

CSIAC is sponsored by the Defense Technical Information Center (DTIC) under Contract #FA8075-17-D-0001. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the DTIC.

# Table of Contents

Introduction .... 3
Preparing for the PIA .... 4
Identifying PII and Its Uses .... 4
Privacy Threats .... 5
Privacy Impact .... 6
Likelihood .... 9
Assessing Privacy Risk .... 10
Determine Risk Treatment .... 11
Examples .... 12
References .... 13
About the Author .... 13

# Introduction

A privacy impact assessment (PIA) is an essential element for effective privacy by design. It enables privacy leaders to be assured that the privacy controls implementation satisfies regulations and organizational requirements, and is key to determining what steps must be taken to manage privacy risk for the organization. The standard ISO 29134 (*Guidelines for privacy impact assessment*, June 2017) defines a PIA as: the overall process of identifying, analyzing, evaluating, consulting, communicating and planning the treatment of potential privacy impacts with regard to the processing of personally identifiable information (PII), framed within an organization's broader risk management framework.

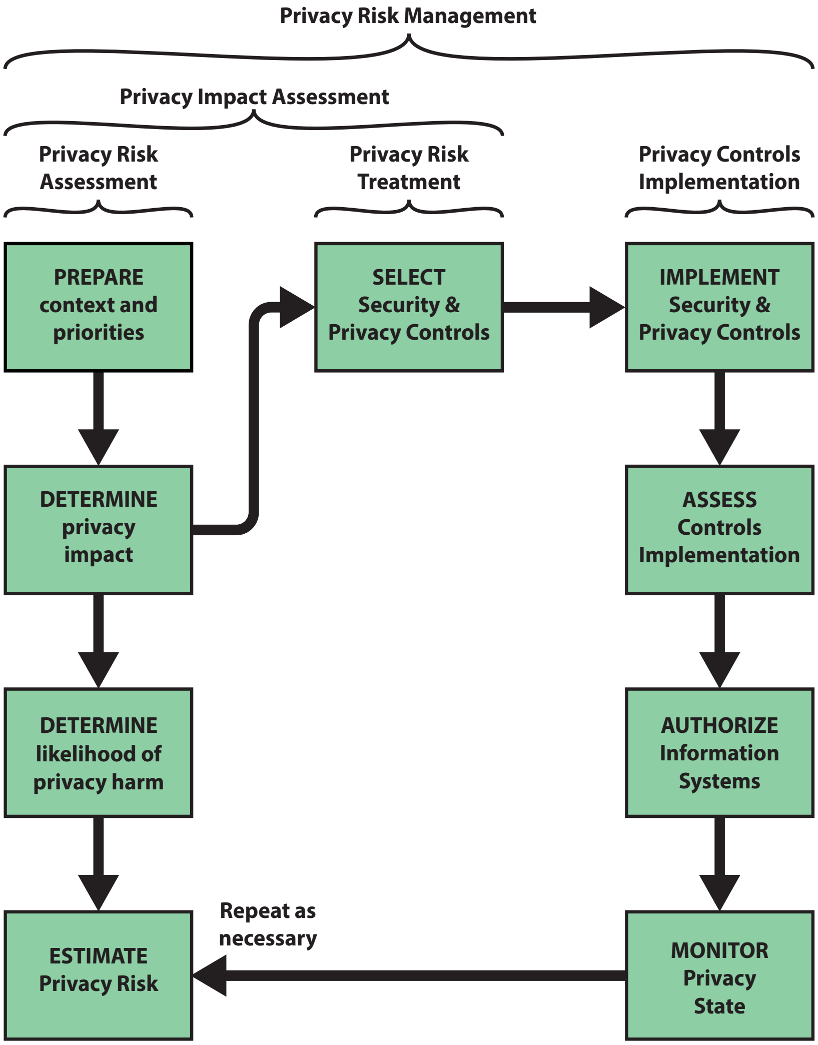

Privacy Risk Management
Privacy Impact Assessment
Privacy Risk Assessment
Privacy Risk Treatment
Privacy Controls Implementation
PREPARE context and priorities
SELECT Security & Privacy Controls
IMPLEMENT Security & Privacy Controls
DETERMINE privacy impact
ASSESS Controls Implementation
DETERMINE likelihood of privacy harm
AUTHORIZE Information Systems
Repeat as necessary
ESTIMATE Privacy Risk
MONITOR Privacy State

***Figure 1. Privacy Risk Management Framework. Note. Reprinted from Information Privacy Engineering and Privacy by Design (p. 355), by William Stallings,2020.***

Figure 1 indicates the scope of PIA. Note that the term *privacy impact assessment* is a misnomer on two counts. First, PIA does not simply assess impact but assesses privacy risk, as explained subsequently. Second, PIA is not limited to assessment but includes selecting controls for privacy risk treatment.

The remainder of this paper looks at the principal tasks illustrated in the two left-hand columns of Figure 1.

## Preparing for the PIA

Preparing for a PIA should be part of strategic security and privacy planning. The important components of PIA preparation are the following: describe the system or project that is the subject of the PIA, identify potential user behavior that could impact privacy, and determine the relevant privacy safeguarding requirements.

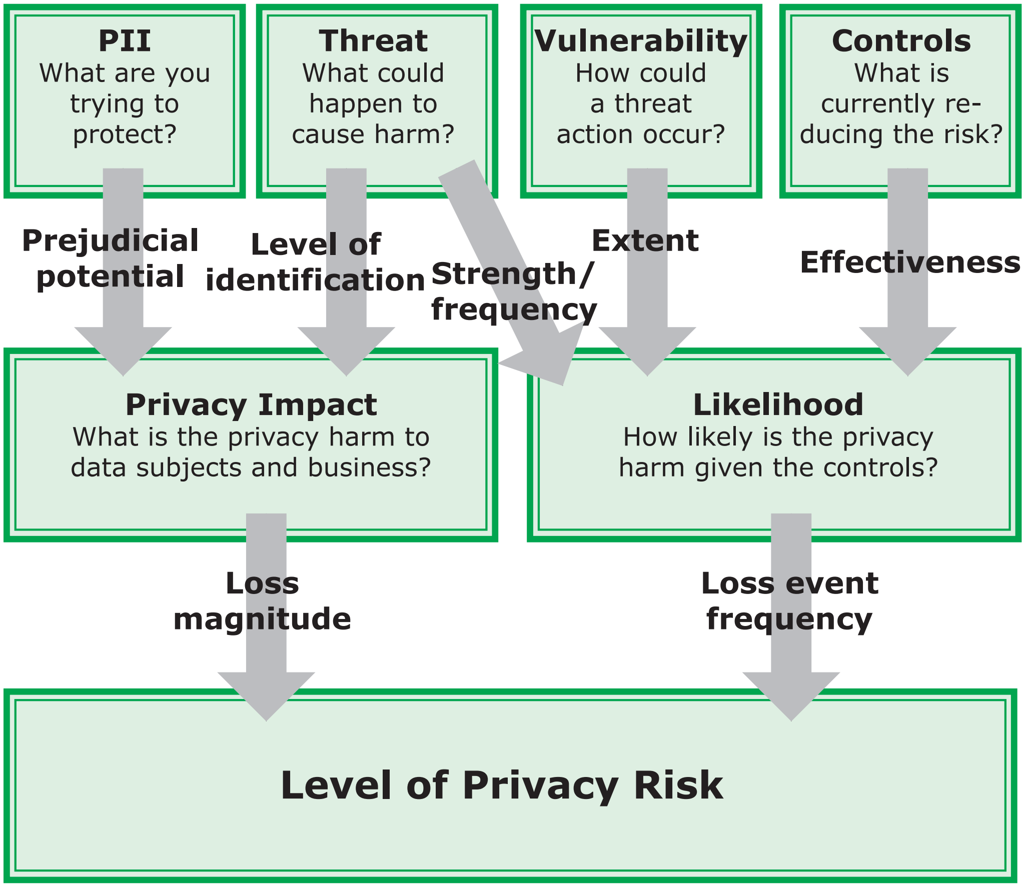

PII
What are you trying to protect?
Threat
What could happen to cause harm?
Vulnerability
How could a threat action occur?
Controls
What is currently reducing the risk?
Prejudicial potential
Level of identification
Strength/ frequency
Extent
Effectiveness
Privacy Impact
What is the privacy harm to data subjects and business?
Likelihood
How likely is the privacy harm given the controls?
Loss magnitude
Loss event frequency
Level of Privacy Risk

*Figure 2. Information Privacy Engineering and Privacy by Design: Understanding Privacy Threats, Technology, and Regulations Based on Standards and Best Practices, by William Stallings, 2020.*

## Identifying PII and Its Uses

The remaining three steps in the first column of Figure 1 are the heart of privacy risk assessment. Figure 2 expands these steps to show the overall structure of privacy risk assessment, which is essentially the same structure used for information security risk assessment, but with a focus on privacy. The left-hand side of Figure 2 deals with determining privacy impacts. The analysis involves determining what PII is part of the system as well as determining what data actions could constitute a threat. The results of threats applied to PII are privacy impacts, or equivalently, privacy harms. The right-hand side of Figure 2 indicates that the likelihood of privacy harm is derived from an assessment of the strength and frequency of threats on the one hand and the vulnerabilities of the system plus any existing controls on the other hand. Finally, a privacy risk can be estimated as a function of privacy impact and the likelihood of privacy harm.

First, consider the identification of PII and its uses. The assessor, or assessment team, needs to create an inventory that identifies all the PII in the system and documents its collection, processing, storage, and/or transmission to third parties.

Essential to understanding the privacy risks involved in the use of PII in a given system or project is a complete description of how PII flows into, though, and out of the system. A workflow diagram is a useful tool for assuring that all aspects of PII handling are documented. The assessor should provide a commentary on and indicate the potential privacy impacts for each point in the workflow.

## Privacy Threats

For a PIA, the primary focus is the impact of privacy violations on the individual. The threats can be classified as follows:

- **Appropriation:** PII is used in ways that exceed an individual’s expectation or authorization. Appropriation occurs when personal information is used in ways that an individual would object to or would have expected additional value for.
- **Distortion:** The use or dissemination of inaccurate or misleadingly incomplete personal information. Distortion can present users in an inaccurate, unflattering or disparaging manner.
- **Induced Disclosure:** Pressure to divulge personal information. Induced disclosure can occur when users feel compelled to provide information disproportionate to the purpose or outcome of the transaction. Induced disclosure can include leveraging access or privilege to an essential (or perceived essential) service.
- **Insecurity:** Insecurity refers to the improper protection and handling of PII. Identity theft is one potential consequence. Another possible consequence is the dissemination of false information about a person, by altering that person's record.
- **Surveillance:** Tracking or monitoring of personal information that is disproportionate to the purpose or outcome of the service. The difference between the data action of monitoring and the problematic data action of surveillance can be very narrow. Tracking user behavior, transactions, or personal information may be conducted for operational purposes such as protection from cyber threats or to provide better services, but it becomes surveillance when it leads to privacy harms.
- **Unanticipated Revelation:** Non-contextual use of data reveals or exposes an individual or facets of an individual in unexpected ways. Unanticipated revelation can arise from aggregation and analysis of large and/or diverse data sets.
- **Unwarranted Restriction:** Unwarranted restriction to PII includes not only blocking tangible access to PII, but also limiting awareness of the existence of the information within the system or the uses of such information.

## Privacy Impact

An organization determines potential privacy impact, or magnitude of harm, by developing a profile of PII that is stored and/or flows through the organization's information systems and cross-referencing that to the potential threats. The impacts can be classified as follows:

- **Loss of Self-Determination**: The loss of an individual's personal sovereignty or ability to freely make choices. This includes the following categories:
  - **Loss of autonomy**: Includes needless changes in behavior, including self-imposed restrictions on freedom of expression or assembly.
  - **Exclusion**: Lack of knowledge about or access to PII. When individuals do not know what information an entity collects or can make use of, or they do not have the opportunity to participate in such decision-making, it diminishes accountability as to whether the information is appropriate for the entity to possess or the information will be used in a fair or equitable manner.
  - **Loss of liberty**: Improper exposure to arrest or detainment. Even in democratic societies, incomplete or inaccurate information can lead to arrest, or improper exposure or use of information can contribute to instances of abuse of governmental power. More life-threatening situations can arise in non-democratic societies.
  - **Physical harm**: Actual physical harm to a person. For example, if an individual's PII is used to locate and gain access to cyber-physical systems that interact with the individual, harms may include the generation of inaccurate medical device sensor readings, the automated delivery of incorrect medication dosages via a compromised insulin pump, or the malfunctioning of critical smart car controls, such as braking and acceleration.
- **Discrimination**: The unfair or unequal treatment of individuals. This includes the following categories:
  - **Stigmatization**: PII is linked to an actual identity in such a way as to create a stigma that can cause embarrassment, emotional distress, or discrimination. For example, sensitive information such as health data or criminal records, or merely accessing certain services such as food stamps or unemployment benefits may attach to individuals creating inferences about them.
  - **Power Imbalance**: Acquisition of PII that creates an inappropriate power imbalance, or takes unfair advantage of or abuses a power imbalance between acquirer and the individual. For example, collection of attributes or analysis of behavior or transactions about individuals can lead to various forms of discrimination or disparate impact, including differential pricing or redlining.
- **Loss of Trust**: The breach of implicit or explicit expectations or agreements about the handling of personal information. For example, the disclosure of personal or other sensitive data to an entity is accompanied by a number of expectations for how that data is used, secured, transmitted, shared, and so on. Breaches can leave individuals reluctant to engage in further transactions.
- **Economic Loss**: Economic loss can include direct financial losses as the result of identity theft, as well as the failure to receive fair value in a transaction involving personal information.

The organization must assess the relative magnitude of each of the potential harms, in terms of the amount of harm to the individual.

Assessing privacy impacts is challenging, because it is individuals, not organizations, the directly experience a privacy harm. Assigning a magnitude to a privacy impact is challenging because there may be significant variability in the harm perceived by individuals, especially for embarrassment or other psychologically-based harms.

With respect to organizational losses, NIST IR 8062 (*An Introduction to Privacy Engineering and Risk Management in Federal Systems*) suggests that organizations may be able to use other costs as proxies to help account for individual impact, including:

- Legal compliance costs arising from the problems created for individuals,
- Mission failure costs such as reluctance to use the system or service,
- Reputational costs leading to loss of trust, and
- Internal culture costs which impact morale or mission productivity as employees assess their general mission to serve the public good against the problem's individuals may experience.

A typical approach to estimating privacy impact is to look at the two factors that contribute to the impact (Figure 2):

- **Prejudicial potential**: An estimation of how much damage would be caused by all the potential consequences of a threat.
- **Level of identification**: An estimation of how easy it is to identify data subjects with the available data processed by the available software.

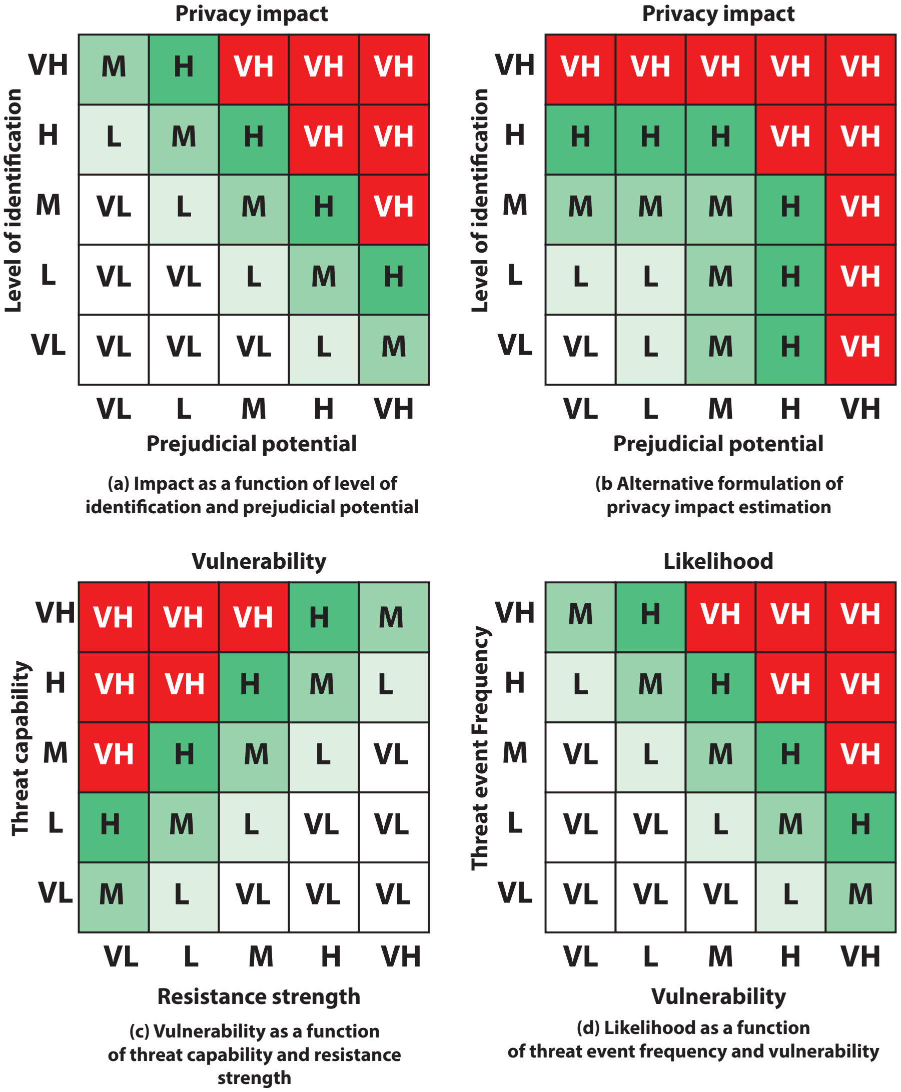
A typical approach of characterizing prejudicial potential is to use five levels, such as very low, low, moderate, high, and very high. An analyst could take into consideration to the type and amount of PII that is to be protected and the relevant threats that could violate the privacy of the PII principals (data subjects). Other factors include sensitivity of the PII breached, numbers of PII principals affected, and level of organizational impact. For individuals, prejudicial impact should consider economic loss, reputational loss, and other personal factors. For organizations, factors to consider are reputational impact, economic loss, and whether to what degree regulations or laws were violated.

Similarly, five levels can be used to characterize the levels of identification. For example, very low corresponds to the case where, with the information available to an adversary, it is virtually impossible to identify individuals; and very high corresponds to it being very easy to identify individuals. For a given threat category, the PIA analyst can estimate how easy is it to identify an individual should the threat gain access to the information residing on an organization asset.

A common method of developing a risk assessment, both in information security and information privacy, is the use of matrices, as illustrated in Figure 3. Figure 3.a depicts the estimation of privacy impact. As discussed, the privacy impact is a function of prejudicial potential and level of identification. Thus, for a given type of threat to a given information asset, the matrix indicates the estimated privacy impact.

Note that the both the levels assigned to the two factors, and the values assigned to each cell in the matrix are a matter of judgement on the part of the analyst or assessor, based on context and the privacy requirements and privacy policy of the organization. This is true throughout the risk assessment process. For example, Figure 3.b shows a privacy impact matrix that is more conservative than that of Figure 3.a.

The following suggested methodology for combining prejudicial potential and level of identification is based on (E.U. Smart Grid Task Force, 2018). The methodology includes the following steps:



Privacy impact

| Level of identification | VL | L | M | H | VH |
|---|---|---|---|---|---|
| VH | M | H | VH | VH | VH |
| H | L | M | H | VH | VH |
| M | VL | L | M | H | VH |
| L | VL | VL | L | M | H |
| VL | VL | VL | VL | L | M |

Prejudicial potential

(a) Impact as a function of level of identification and prejudicial potential

Privacy impact

| Level of identification | VL | L | M | H | VH |
|---|---|---|---|---|---|
| VH | VH | VH | VH | VH | VH |
| H | H | H | H | VH | VH |
| M | M | M | M | H | VH |
| L | L | L | M | H | VH |
| VL | VL | L | M | H | VH |

Prejudicial potential

(b Alternative formulation of privacy impact estimation

Vulnerability

| Threat capability | VL | L | M | H | VH |
|---|---|---|---|---|---|
| VH | VH | VH | VH | H | M |
| H | VH | VH | H | M | L |
| M | VH | H | M | L | VL |
| L | H | M | L | VL | VL |
| VL | M | L | VL | VL | VL |

Resistance strength

(c) Vulnerability as a function of threat capability and resistance strength

Likelihood

| Threat event Frequency | VL | L | M | H | VH |
|---|---|---|---|---|---|
| VH | M | H | VH | VH | VH |
| H | L | M | H | VH | VH |
| M | VL | L | M | H | VH |
| L | VL | VL | L | M | H |
| VL | VL | VL | VL | L | M |

Vulnerability

(d) Likelihood as a function of threat event frequency and vulnerability

*Figure 3. Elements of Risk Determination, by William Stallings, 2020.*

1. Identify a list of privacy threat categories relevant to the organization's environment.
2. Identify the primary assets associated to each threat category. A primary asset is a set of one or more pieces of PII allocated on a specific IT system that requires protection.
3. For each primary asset, identify the relevant privacy harms that may occur.
4. For each potential harm resulting from a particular threat, associate the value of the level (very low up to very high) that best matches the prejudicial effect.

5. For each primary asset determine the prejudicial effect as the maximum value of prejudicial effect for a potential harm. This is a level from very low to very high.
6. For each primary asset of each threat category, use a matrix such as Figure 3a or 3b to estimate the privacy impact, or severity, for the primary asset.
7. The severity of a given threat category is the maximum value of severity for all associated primary assets.

# Likelihood

IR 8062 defines likelihood in privacy risk assessment as the estimated probability that a threat action will occur and be problematic for a representative or typical individual whose PII is processed by the system. This is a complex issue that involves assessing the following factors:

- The probability that a threat action is attempted, either intentionally or unintentionally.
- The vulnerabilities in the system that would enable the attempted threat action to occur. This in turn is function of the capability, or strength, of the threat and the resistance strength of a system or asset to that particular threat.
- The reduction in probability due to the effectiveness of existing or planned security and privacy controls.

As with impact assessment, the risk analyst can use a five-level scale for each of these factors. As an example consider the following ranges. The capability of a threat source ranges from the source has no special capabilities to carry out a threat (very low) to the source has malicious intent and considerable expertise to carry out a threat (very high). Threat event frequency ranges from less than once very ten years (very low) to over 100 times per year (very high). Resistance strength ranges from privacy violation very difficult, requiring sustained effort and specialized expertise (very high) to no special technical knowledge is needed, or can be a result of careless usage (very low). However, the analyst needs to modified these resistance strength estimates by considering any controls already designed or planned for the system that provide protection to the primary assets (PII data). Finally, control effectiveness ranges from only protecting against 2% of an average threat population (very low) to protecting against 98% of an average threat population (very high).

Privacy vulnerabilities can be classified as:

- **Technical vulnerabilities**: flaws in the design, implementation and/or configuration of software and/or hardware components, including application software, system software, communications software, computing equipment, communications equipment, and embedded devices.
- **Human resource vulnerabilities**: key person dependencies, gaps in awareness and training, gaps in discipline, improper termination of access.
- **Physical and environmental vulnerabilities**: insufficient physical access controls, poor siting of equipment, inadequate temperature/humidity controls, inadequately conditioned electrical power.

- **Operational vulnerabilities**: lack of change management, inadequate separation of duties, lack of control over software installation, lack of control over media handling and storage, lack of control over system communications, inadequate access control or weaknesses in access control procedures, inadequate recording and/or review of system activity records, inadequate control over encryption keys, inadequate reporting, handling and/or resolution of security incidents, inadequate monitoring and evaluation of the effectiveness of security controls.
- **Business continuity and compliance vulnerabilities**: misplaced, missing or inadequate processes for appropriate management of business risks; inadequate business continuity/contingency planning; inadequate monitoring and evaluation for compliance with governing policies and regulations.
- **Policy and procedure vulnerabilities**: privacy policies and procedures that are inadequate to fully protect PII, including conformance with FIPPs.
- **Dataset vulnerabilities**: weakness in de-identification measures; inadequate masking of PII in statistical datasets, inadequate protection against discovery of PII by analysis of multiple datasets.

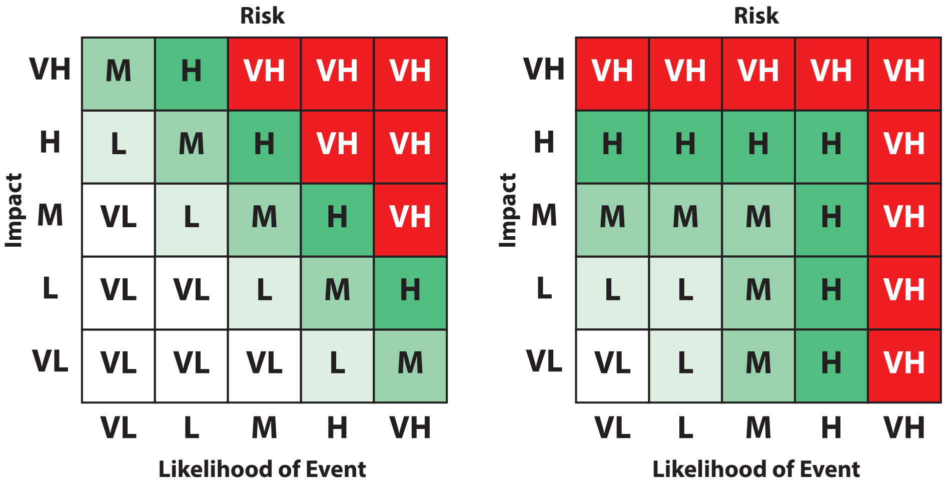
The result of this analysis is a residual ease of privacy breach. Then the analyst could use Figure 3.c to estimate the extent of vulnerability. Finally, the analyst could use a matrix similar to Figure 3d to estimate likelihood as a function of threat event frequency and the extent of vulnerability.

As with impact assessment, the analyst needs to perform the likelihood analysis for each primary asset for each threat category.

## Assessing Privacy Risk

Privacy risk assessment is based on an estimate of impact and of likelihood. The organization should carry out this assessment for each primary asset (PII stored on a system) and each threat category.

Figure 4 gives two examples of a qualitative risk assessment matrix. The matrix defines a risk level for each impact level and likelihood level. The structure of the matrix on the left in the figure is commonly used. For example, ISO 29134 uses a 4×4 matrix with this structure. The structure on the right would be suitable for a more conservative or risk-averse organization.



Risk

| Impact \ Likelihood of Event | VL | L | M | H | VH |
|---|---|---|---|---|---|
| VH | M | H | VH | VH | VH |
| H | L | M | H | VH | VH |
| M | VL | L | M | H | VH |
| L | VL | VL | L | M | H |
| VL | VL | VL | VL | L | M |

Risk

| Impact \ Likelihood of Event | VL | L | M | H | VH |
|---|---|---|---|---|---|
| VH | VH | VH | VH | VH | VH |
| H | H | H | H | H | VH |
| M | M | M | M | H | VH |
| L | L | L | M | H | VH |
| VL | VL | L | M | H | VH |

*Figure 4. Two Example of Privacy Risk Matrices, by William Stallings, 2020.*

For each primary asset and threat category, the organization should set a priority, based on where the risk is located in the matrix and on the risk criteria used by the organization. The following are typical guidelines for the five risk levels:

- **Very high**: These risks must be absolutely avoided or significantly reduced by implementing controls that reduce both their impact and likelihood. (E.U. Smart Grid Task Force, 2018) recommends that the organization implement independent controls of prevention (actions taken prior to a damaging event), protection (actions taken during a damaging event) and recovery (actions taken after a damaging event).
- **High**: This risks should be avoided or reduced by implementing controls that reduce the impact and/or likelihood, as appropriate. For example, the matrix on the left in Figure 4 has a high risk entry for a very high impact and low likelihood; in this case the emphasis is on reducing impact. The emphasis for these risks should be on prevention if the impact is relatively high and the likelihood is relatively low, and on recovery if the impact is relatively low and the likelihood relatively high.
- **Moderate**: The approach for moderate risk is essentially the same as for high risk. The difference is that moderate risks are of lesser priority and the organization may choose to devote less resources to addressing them.
- **Low**: The organization may be willing to accept these risks without further control implementation, especially if the treatment of other security or privacy risks also reduce this risk.
- **Very low**: The organization may be willing to accept these risks because further attempts at reduction are not cost effective.

## Determine Risk Treatment

The final step of the PIA is to determine what risk treatments will be applied to the identified risks. This involves three tasks:

- Choose treatment options
- Determine controls
- Create risk treatment plans

The risk treatment options are the following:

- **Risk reduction or mitigation**: Actions taken to lessen the probability and/or negative consequences associated with a risk. Typically, an organization achieves risk reduction by selecting additional privacy controls.
- **Risk retention**: Acceptance of the cost from a risk.
- **Risk avoidance**: Decision not to become involved in, or action to withdraw from, a risk situation.
- **Risk transfer or sharing**: Sharing with another party the burden of loss from a risk.

The assessor must recommend the most appropriate treatment option for each identified privacy risk. The treatment option decision for each risk will depend on balancing a number of factors, notably the cost of each option to the organization and the organization's obligation to protect the privacy of the PII.

For each risk for which the assessor chooses risk reduction, the responsible privacy personnel need to select the appropriate combination of security and privacy controls to mitigate or eliminate the risk. The privacy leader may perform this task, or it may be assigned to one or more other privacy personnel, including the privacy assessor. The preference should be to use industry-standard controls, such as those defined in ISO 29151 (*Code of Practice for Personally Identifiable Information Protection*, August 2017) and NIST SP 800-53.

The assessor should develop a risk treatment plan for each identified risk, to include the following information:

- Rationale for chosen treatment option
- Specific controls to be implemented
- Residual risk after treatment
- Result of a cost/benefit analysis
- Person responsible for implementation
- Schedule and resources
- How the implementation will be monitored and evaluated.

# Examples

The reader may find it useful to examine real-world examples of PIAs. Two worthwhile documents are (E.U. Smart Grid Task Force, 2018), which deals with a smart grid application, and [SNZ12], which covers the New Zealand Integrated Data Infrastructure. In addition, the U.S. Department of Homeland Security maintains an inventory of publicly-available PIAs produced by various government agencies (https://www.dhs.gov/privacy-impact-assessments). Although these examples are all from government agencies, they provide useful insights for businesses as well.

# References

E.U. Smart Grid Task Force. *Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems.* September 2018. https://ec.europa.eu/energy/en/topics/markets-and-consumers/smart-grids-and-meters/smart-grids-task-force/data-protection-impact-assessment-smart-grid-and-smart-metering-environment

Statistics New Zealand. Privacy impact assessment for the Integrated Data Infrastructure. 2012. http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/keep-data-safe/privacy-impact-assessments/privacy-impact-assessment-for-the-idi.aspx

Stallings, W. (2020). Information Privacy Engineering and Privacy by Design: Understanding Privacy Threats, Technology, and Regulations Based on Standards and Best Practices, by William Stallings, 2020.

# About the Author

**Dr. William Stallings** holds a PhD from M.I.T. in Computer Science. He is an independent consultant and author of numerous textbooks on cybersecurity, computer networking, and computer architecture. He has twelve times received the award for the Best Computer Science and Engineering Textbook of the Year from the Textbook and Academic Authors Association. His most recent book is *Information Privacy Engineering and Privacy by Design* (Pearson, 2020). He is also author of *Effective Cybersecurity: A Guide to Using Best Practices and Standards* (Pearson, 2019), and *Cryptography and Network Security, Principles and Practice* (Pearson, 2020). Dr. Stallings is on the editorial board of Cryptologia, a scholarly journal devoted to all aspects of cryptology.