Design Assurance Standard

Version 1.0

Adopted March 1st, 2025

Standards Committee

Institute of Operational Privacy Design

R. Jason Cronk President Institute of Operation Privacy Design	Nandita Narla Chair Standards Committee

Members of the Standards Committee Contributing to this Standard:

Kingsley Audu, Jay Averitt, Caroline Carver, Hilary Coote, Tarana Damania, Kayvis Damptey, Peter Duran, Scott Edmiston, Debra Farber, Selin Fidan, Pruthvi Gurram, Steve Hickman, Charlene Hinton, Conor Hogan, Amaka Ibeji, Shalab Jain, Rohit Kumar, Kimberly Lancaster, Dan Lopresto, Rongfei Lu, Carl Mathis, Vamsee Metlapall, Sean Milford, Maaz Bin Musa, Nicole Nguyen, Ralph O'Brien, Vandana Padmanabhan, Mahim Patel, Varun Prasad, Sagar Rahurkar, Smita Rajmohan, Denise Schoeneich, Nabeel Shamsi, Stuart Shapiro, Divya Sharma, Kiran Sharma, Mukta Sharma, Tanusree Sharma, Daniel Smullen, BVJ Srinivas, Akhilesh Srivastava, and Mary Yip.

Table of Contents

Introduction

The Institute of Operational Privacy Design (IOPD) is dedicated to creating and adopting privacy design standards to protect privacy. The IOPD's mission is threefold: to evangelize privacy by design through education and standards, to provide accountability through certification mechanisms, and to publicly recognize good privacy practices by companies around the globe.

The IOPD is a non-profit, membership-based professional organization primarily run by a volunteer Board of Directors. The Board seeks input from advisors who sit across industries and have varied roles within organizations. As advocates for privacy design standards, the IOPD and its members aim to respect the privacy of individuals in all their practices. We hold ourselves to the highest standards, which we expect from all businesses recognized publicly for good privacy practices.

This Design Assurance Standard (the "Assurance Standard") follows two years of effort by the IOPD's Standards Committee after its adoption of the Design Process Standard (the "Process Standard") in January 2023. While the Process Standard details the design process components needed to incorporate privacy considerations and reduce privacy risks, this Assurance Standard confirms an organization's claim that a specific product, service, or business process has been designed, developed, or deployed with privacy aforethought. In other words, the Assurance Standard doesn't apply to an organization but to a specific product, service, or business process. The intent for this conformance standard is for organizations to demonstrate that they have achieved reasonable assurance around claims of "privacy by design and default."

In theory, a product, service, or business process that has been designed, developed, and deployed using the Process Standard should meet the Assurance Standard. In practice, this may not be the case because, for the Process Standard, organizations may select their own risk model, whereas this Assurance Standard uses a defined risk model.[1] The use of a singular defined risk model and an assurance case, more generally, provides a common measure enabling relative comparison of privacy respecting qualities between disparate products, services, and business processes.

In addition to providing a measure for privacy capability comparison, this Assurance Standard provides methodologies to determine whether the evidence supporting a purported privacy claim has been satisfied. Organizations whose products, services, or business processes satisfy these evidentiary burdens can apply for the IOPD's Privacy by Design and Default Trust Mark.

[1] See the formal definition in the Appendix and the section on risk model for this Standard's risk model.

This Standard uses a structured assurance case[2] notation of "claims, arguments, and evidence" (CAE) to validate privacy respecting capabilities of in-scope systems. Although the IOPD's Standards Committee has done most of the hard work constructing the assurance case (i.e., the set of claims, arguments, and necessary evidence), there will be an aspect of customization necessary from organizations utilizing this standard.

The IOPD has crafted the Assurance Standard for organizations to clarify the connections between the evidence and privacy claims. This Assurance Standard approach is novel in comparison to the current landscape where many standards and certifications are limited to variants of the Fair Information Practice Principles (FIPPs)[3] or specify requirements or controls without consideration of whether those controls address risks that are created by the product, service, or business process in question.

Privacy professionals who work in "data protection" rather than privacy may find the language used in the Standard particularly jarring, though some terms (e.g., "proportional") will seem familiar.[4] First and foremost, this Standard addresses the broader domain of privacy, not specifically data protection (or even "information privacy"). Second, many terms are drawn from systems engineering, threat modeling, and other approaches to risk management. Looking under the hood of this Standard, many of the concepts align very closely with those in data protection; one can read the defined term Interactions as equivalent to "data processing," the Applicant is a controller, Threat Actors include controllers, processors, and third parties, such as cybercriminals, many Harms in the Risk Model, such as exclusion[5], mirror data protection rights.

[2] See 'Assurance Cases' section for an explanation of CAE structured notation.

[3] See Federal Privacy Council https://www.fpc.gov/resources/fipps/

[4] After publication of this Standard, the IOPD will begin working on a version for use in complying with Article 25, Data Protection by Design and Default, of the European Union's Regulation 2016/679/EU (i.e., General Data Protection Regulation or GDPR).

[5] Not defined in this Standard, but Exclusion in the Solove Taxonomy is failure to let an individual know about data processing or participate in its use.

Purpose

The intended audiences of this Standard are privacy professionals; organizations designing, developing, configuring, or deploying products, services, and business processes; implementers and Assessors; as well as privacy and data protection regulators. This Standard serves several purposes for these four distinct audiences:

- **For the privacy professional,** the Standard serves to illustrate an aspirational and achievable objective regarding the design, development, or deployment of products, services, and business processes.

- **For organizations**, the Standard represents a way to measure and qualify whether a design is privacy-respecting. This can be for the purpose of internal improvement, brand differentiation, compliance with obligations, and/or satisfying ethical imperatives. For those organizations wanting to assert that they have accomplished “Privacy by Design and Default,” the Standard represents a rigorous set of externally validated criteria to back up that claim.

- **Implementers and Assessors** can utilize the Standard in support of their client engagements to assist those clients in achieving “Privacy by Design and Default” objectives.

- **Government Regulators** can use the Standard as a benchmark for reviewing claims and public statements about “Privacy by Design and Default” in organizations' designs of products, services, or business processes.

Quick Start Guide

This standard can seem daunting at first. To help guide first time implementers, the IOPD includes this guide on how to get started. This guide avoids capitalizing defined terms for ease of reading, whereas the other parts of the standard use defined terms to impart precision. In the event of any conflict, the normative part of the standard governs.

Steps 1 through 4 are about selecting the scope of the target to which the standard is being applied.	
Step 1: Identify who is applying the standard to their product, service, or business process (i.e., the target system). This is either the designer, developer, or deployer and is referred to as the applicant. A designer plans the system, a developer constructs it, and a deployer puts it into functional operation. Sometimes, the applicant can play multiple roles, for instance, being both a designer and a developer.	*For the following example, we will use a large web-based email service provider as an example. This provider is the designer (conceptualizing features) and the developer (coding the features into existence). In some circumstances, they could be the deployer, making the service available to consumers, but in others, such as enterprise clients, those clients may be the ones deploying the service in their environment for the ultimate end user, the enterprise employee.*
Step 2 Specify the product, service, or business process. Be careful; many services may, in fact, be multiple services in one product.	*The service could serve different markets, an electronic mail service, an advertising platform for advertisers, and a company-wide communication management platform for enterprise clients. While it's possible to apply the standard to all, pick one to start with as each has distinct customers, risks, and requirements*
Step 3 Specify the categories of customers. Customers may or may not be consumers of the service. For instance, designers could sell to resellers or deployers who put the service into operation. Many products and services have different market segments and business models demanding different configurations of a product or service. You should look to address all variants of the product or service, but you can also limit yourself to just one or a few.	*Even if the service is narrowed to an electronic mail service, it may still be presented differently to free license clients than to paid clients. Different jurisdictions may also demand different needs and defaults, further segmenting the customer base. Specifying the customers is necessary to identify the default configurations provided to them.*
Step 4 Select the risks of concern. This standard uses a risk model with 64 discrete privacy risks based on threat actors, at-risk parties, and consequences to them. You need not, and most likely will not, address all 64 risks. Pick those that are most pertinent or of concern to the applicant.	*Spam is an obvious concern. You could consider two risks. Non-contracted parties (email senders sending to your clients) put consumers of the service at risk of invasion harms (e.g., intrusion into their personal space) through spamming. You might also consider your customers (contracted parties) spamming others (e.g., bystanders), which would also be an invasion harm.*

Select one from each column in the table below to specify a risk.	

Threat Actors	At Risk Parties	Consequences
The Applicant	Consumers	Information Processing Harms
Contracted Parties	Operators	Information Dissemination Harms
Non-Contracted Parties	Resources	Collection Harms
Other Parties	Bystanders	Invasion Harms

Step 5. **a.** Identify the target system requirements. Requirements can be functional, meaning what the product needs to do to achieve its intended purpose, or they can be quality attributes, meaning characteristics, to evaluate the quality of the product.	*What are the requirements of an email service? People have to send and receive email. They might need to store email for a certain amount of time and be able to delete email. You'd want qualities of resilience, uptime, speed, and more.*
Step 5. b. Identify all relevant interactions between the threat actors and the at-risk parties (or their data) as part of the product or service.	*Interactions include your users sending emails to others, including other users and non-users of your email service, and outsiders sending emails to your users.*
Step 5. c. Show how the interactions are necessary to meet the requirements.	*If you specify that the organization must collect an email address from a person, you must show that this is necessary to send emails (a functional requirement of the service as identified in 5.a).*
Step 6 Specify how the organization's controls reduce each type of risk chosen in Step 4.	*Suppose your risk concerns the receipt of spam by users. In that case, analyzing email for spammy content is a control that will reduce the likelihood that the users will receive unsolicited emails.*
Step 7 Show that documentation exists on how the types of controls the organization uses were verified.	*Your spam filter might shift all spam to a specific folder. Testing documentation might suggest sending spam to an email address and reviewing the particular folder in the account holder's email account to verify that spam was placed in that folder. Testing documentation can be provided externally or produced internally.*
Step 8 Review the controls and demonstrate that the controls meet the requirements from Step 7. This is about performing the tests specified in Step 7 for each control.	*One could conduct the test of the control and demonstrate that spam was shifted to the appropriate folder in a person's mailbox.*

Step 9 Show that controls have objective measures for effectiveness. While Steps 7 and 8 show that a control is operational, Steps 9 and 10 show that it is effective at reducing the applicable risk.	*An objective measure might be that the control is supposed to filter out 90% of spam. However, consumer surveys or focus groups may be necessary for subjective control measures (like the effectiveness of consumer notice).*
Step 10 Demonstrate that the controls in place in the product or service meet the objective measures for effectiveness. You would demonstrate this through tests or assessments. This needs to be done for *each* control in place.	*You could show that 90% of spam is filtered out through actual testing and sampling.*
Step 11 Justify the residual risk. For each interaction between a threat actor and an at-risk party (or their data), state a justification for the specifics of the interaction that accounts for any controls. Note that this is distinct from 5.c., which entails showing that an interaction is necessary to meet the requirements, whereas this step is about justifying the risks. You could have an interaction that is necessary for the proposed product or service but unjustifiable (e.g., a spy camera needs to see people to be a "spy camera," but that doesn't justify spying on people).	*A justification statement might read, "Some spam in the user's inbox is acceptable so as not to block all legitimate email." The justification statement identifies the purpose of allowing interactions (i.e., because we don't want to block all emails) but acknowledges some risk (i.e., some spam will get through).*
Step 12 Identify benefits for each justification. Justification statements may have implied or explicit benefits.	Benefit: *to receive legitimate emails* Beneficiary: *email account holder*
Step 13 Assure that benefits outweigh residual risk. This step is a little different in that the applicant must create an argument showing that benefits outweigh the risks and then support that argument with evidence. Arguments need not be purely utilitarian. Arguments must be reasoned and consistent and supported by objective evidence.	*Argument: Most consumers agree that some spam is acceptable to ensure that legitimate emails are delivered. Evidence: A consumer survey of email account holders showing that as long as 99.99% of legitimate emails get through, they would prefer most spam to be blocked, even if some get through.*
For Step 14, applicants can choose, based on circumstances, which one applies: 14a, 14b, or both.	
Step 14. a. If the product or service is not configurable, the applicant must demonstrate that there is no configurability.	*Most email services will probably have some configurations. However, specific settings may not be configurable. For instance, the service may be hardcoded to receive emails on port 25, which is generally not used. A deployer could demonstrate this through documentation or testing.*

Step 14. b. Document any hidden or exposed configurability and which configurations are provided to which customers. Different customers may be given different default configurations. Exposed configurability may be obvious documented settings. Hidden configurability may include init files or other undocumented settings that only advanced customers can access.	*Screenshots of configuration panels or init files showing that a spam filter could be turned on or off.*
For the final step, applicants may find one or both to be applicable, depending on the design of their product or service.	
Step 15. a. Show that changes to the defaults would increase interactions or reduce the effectiveness of controls. For each default configuration provided to customers, the applicant needs to show that any reasonable changes (e.g., toggling a setting) will result in either more interactions between parties or compromise the operation or effectiveness of controls (e.g., turning off encryption).	*Turning the spam filter off would increase interactions between other parties (those emailing the account holder) and account holders. Another setting might affect the aggressiveness of the filter; thus, turning it down would reduce the effectiveness of the control at reducing spam.*
Step 15. b. Show that changes to the defaults would create an undesirable balance of risk and benefits. Where changes to settings would result in more complex behavior than merely increasing interactions or reducing control effectiveness, the applicant must craft an argument that the balance between risks and benefits would be undesirable. Further, the applicant must, through objective evidence, support the conclusion.	*Argument: Turning the aggressiveness filter up would reduce spam, but it would also produce too many false positives, resulting in undelivered desired correspondence. Evidence: In the default setting, the percentage of wanted to unwanted messages not delivered is 0.01%, but increasing the aggressiveness would result in a 10-fold decrease in effectiveness (i.e., down to 0.1%), which a survey shows would be unacceptable to most consumers.*

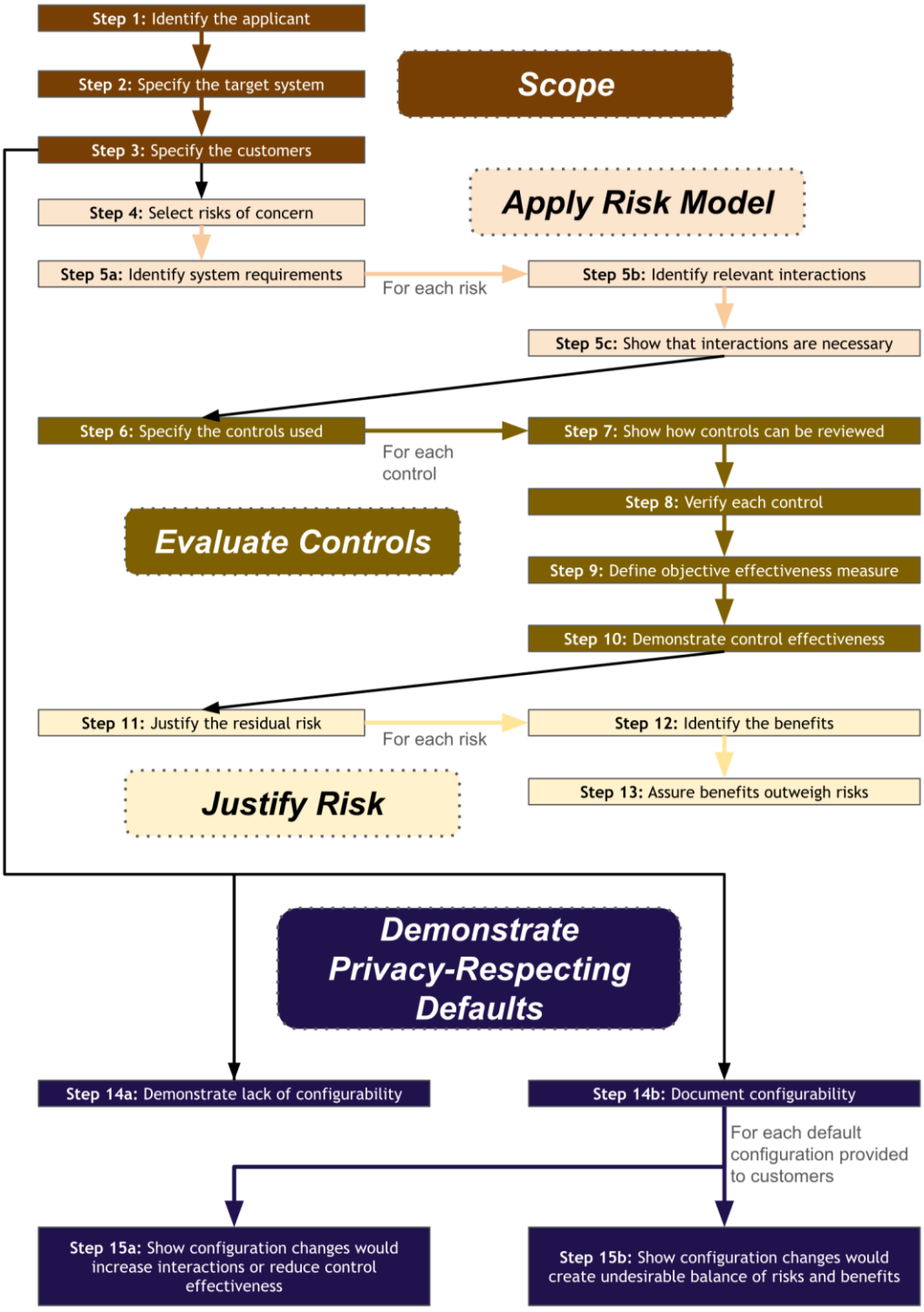

Figure 1 - Flow chart of steps in the Quick Start Guide

Assurance Cases

Assurance cases are a formal approach to establishing confidence in a belief or assertion.[6] Assurance cases explain why a target (system, product, service, or process) is believed to have certain qualities. Historically, the principal property of concern in assurance cases has been safety. Indeed, assurance cases as a generic term emerged out of safety cases[7]. Over time, the approach has been extended to other properties, notably security and privacy. One of the benefits of using an assurance case is flexibility, where a rigid prescriptive requirements-based standard may not be contextually relevant. Given the vagaries of privacy concerns and context, assurance cases seem well suited to privacy.

Assurance cases are based on structured argumentation, a technique that dates back over half a century and, in its original form, is attributed to British philosopher Stephen Toulmin. Structured argumentation consists of decomposing the different elements of an argument and mapping them and their relations to one another. Thus, claims are specified, evidence supporting those claims is described, and the reasoning connecting evidence and claims is articulated.

Assurance cases typically use defined graphical languages to document their arguments. The two most widely used are Goal Structuring Notation (GSN) and Claim, Argument, Evidence (CAE). This standard utilizes CAE to specify its privacy by design and default assurance case. This standard's reliance on assurance cases reflects recognition of the variety of targets to which it might be applied and long-standing deficiencies in the way privacy risk is typically approached. Despite efforts to shift privacy from a checkbox compliance exercise to a risk-based approach, prescriptive controls still dominate, and privacy failures remain common. This trend has only intensified as the socio-technical environment has become more complex.

Using an assurance case helps demonstrate that privacy has been appropriately addressed in the target system. Assurance cases are agnostic regarding the nature of the target of concern, and both enable and compel completely customized explanations in a standard format of how privacy risks are addressed, independent of any prescriptive list of measures. The goal is to ensure that we have confidence in the measures taken to protect privacy based on a clear understanding of the risks involved.

---

[6] Rushby, J.M. (2015) The Interpretation and Evaluation of Assurance Case. Available at https://www.csl.sri.com/~rushby/papers/sri-csl-15-1-assurance-cases.pdf

[7] Ewen Denney and Ganesh Pai, SGT / NASA Ames Research Center, “Towards an Ontological Basis for Aviation Assurance Cases.” Available at https://www.faa.gov/sites/faa.gov/files/air_traffic/technology/swim/governance/Safety%20Cases.pdf

Risk Model

A risk model is a construct that provides the basis for risk assessment of a product, service, or business process. A complete risk model consists of component models for threats, vulnerabilities, and adverse consequences, along with ways of representing likelihood and impact severity. The component models reflect the chain of elements that result in risks, in which threats exploit vulnerabilities, resulting in adverse consequences. Risk models are essential for risk assessment as they specify risks of concern in a given domain and define how they can manifest. Not every threat will be capable of exploiting every vulnerability, nor will every exploitation lead to every possible adverse consequence. The purpose of the risk assessment is to identify alignments of threats, vulnerabilities, and consequences that are viable combinations for the target.

Applicants may leverage a pre-existing model or develop one or more that are tailored to their domains of operation.[8] However, the assurance case at the heart of this standard must be constructed with reference to a specific privacy risk model to standardize the analysis of the Claims, Arguments, and Evidence constituting the case. Customized risk models would require significantly more review and analysis of the case to determine if the risk model sufficiently and completely addresses the risks. Such flexibility would also increase subjectivity and opportunities for gaming the standard, causing confusion in the marketplace and regulators looking for standardization. The standard would be mainly rendered meaningless if each Applicant could utilize a different privacy risk model in their assurance case, and consistently evaluating those cases would become unmanageable. Therefore, this Standard requires Applicants to employ the common privacy risk model defined here. This model is general enough to accommodate all Applicants and Targets. The model leverages Solove's Taxonomy of Privacy[9] problems, a widely held and used model of privacy, and defines different types of vulnerabilities (identified by the kind of interaction), threats (determined by the threat actor and thread action), and consequences (identified by at-risk party and the harm to them).

Explication of likelihood and severity is left to the Applicant. Figure II details this model. The model defines two Vulnerabilities, each of which can be exploited by two actions (resulting in four threat actions). The two Vulnerabilities and corresponding threat actions are:

- **Vulnerability 1:** Threat Actor interacts with At-Risk Parties or data related to At-Risk Parties
  - **Threat action I:** Threat Actor invades personal space or disrespects autonomy
  - **Threat action II:** Threat Actor attempts to collect or solicit information
- **Vulnerability 2:** Threat Actor has control of data related to At-Risk Parties
  - **Threat action III:** Threat Actor processes information
  - **Threat action IV:** Threat Actor disseminates information

[8] The IOPD Design Process Standard v 1.0 allows organizations complete flexibility in defining their risk model and selecting their risks.

[9] https://scholarship.law.upenn.edu/penn_law_review/vol154/iss3/1/

These four threat actions can be taken by any of four types of Threat Actors: the Applicant, Contracted Parties, Non-contracted Parties, and Other Parties. This results in 16 potential Threats (e.g., a potential action by a Threat Actor).

	Threat Action, which can lead to…	**Harm**
I	Threat Actor invades personal space or disrespects autonomy	physical or psychological intrusion or interference with decision-making (i.e., invasion harms)
II	Threat Actor attempts to collect or solicit information	surveillance, interrogation (i.e., collection harms)
III	Threat Actor processes information	aggregation, secondary use, identification, insecurity, and exclusion (i.e., information processing harms)
IV	Threat Actor disseminates information	disclosure, exposure, increased accessibility, distortion, breach of confidentiality/trust (i.e., information dissemination harms)

Table 1 - Four Threats in the Risk Model and corresponding four Harms.

Each of the different threat actions results in exactly one of a set of related Harms. Those Harms follow the categorization of privacy harms under Solove's taxonomy.

The model contains four types of At-Risk Parties: Consumers, Operators, Resources, and Bystanders. Each of these four groups of At-Risk Parties can be impacted by any of the four threat actions (and corresponding harms) perpetrated by any of the four Threat Actors, leading to a total of 64 (4*4*4) potential risks. Not all risks may be pertinent to all Target Systems. Furthermore, the Applicant will scope the Risks they wish to address when applying the Standard, descoping risks due to negligible likelihood or impacts or other considerations, such as analysis costs or market demands.

Example risk: the Applicant (a Threat Actor) processes information (a threat action) of a Consumer (an At-Risk Party) considered a secondary use of data (a Harm).

Example risk: a potential client (a Non-Contracted Party / Threat Actor) of the Applicant asks (a threat action) an employee (an Operator / At-Risk Party) during a marketing call (an Interaction which creates a Vulnerability) being made by the employee a personal question (a Harm, specifically a collection harm).

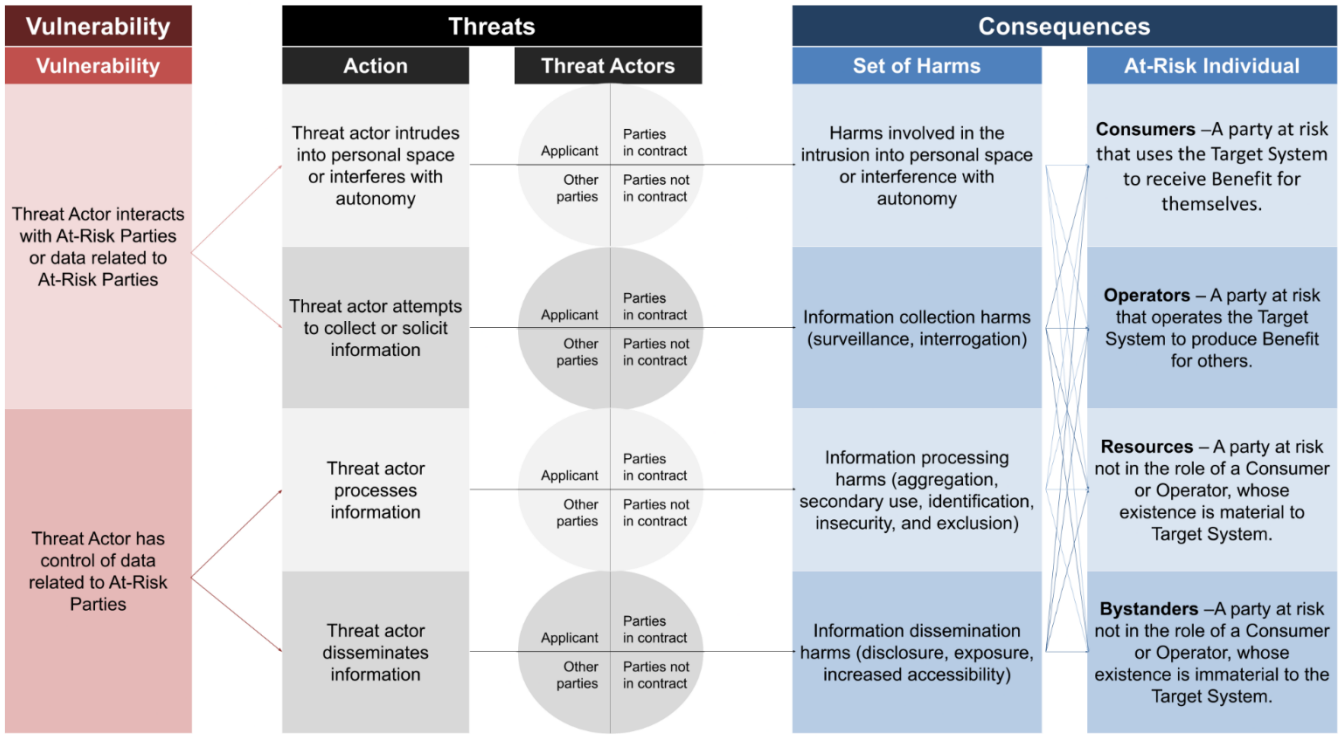

Figure 2 - Risk Model illustrating relationships between Vulnerabilities, Threats, and Consequences in the Risk Model.

Scope Selection

Scoping is a critical first step before applying the Standard to a Target System. Scoping includes

- Identifying the intended Customer(s)
- Defining the Target System
- Identifying the classes of Risk treated (including a delineation of the Threat Actors and groups of At-Risk Parties.)

Identifying the Customers

Applicants may have different sales channels, markets, industries, verticals, or other segments. Applicants must identify the distinct types of Customers they serve, at least to the degree that those Customers receive distinct Configurations. Customers may warrant differing Configurations because of their varying risk profiles (e.g., government customer channels will have different risk profiles than private sector customer channels); identifying Customers can be complex for large, multi-channel products and services. Applicants need NOT scope every Customer channel to apply this Standard. For instance, an Applicant may decide only to analyze or certify their consumer market.

IMPORTANT NOTE: Customers should not be confused with Consumers, though they may sometimes be the same. Customers receive the product or service from the Applicant. Consumers use the product or service for some function. For example, Customers for a business process would be the recipient of the output of the process (e.g., the customer of the budgeting process would be the department receiving the budget; the customer of a shipping process would be the fulfillment department; the customer of a marketing campaign development process would be the department whose product is being marketed). A Consumer of a business process is the one who uses the business process to perform its function (e.g., the marketing department uses the marketing campaign development process to develop a marketing campaign). The Operators of a business process are the parties providing the process to the Consumer. In the marketing example, the Operators include the company (who provides the people to complete the process), marketing department management (who provides the steps of the process), and the information technology department (who provides the technology).

Defining the Target System

The Applicant must identify the product, service, or business process the Applicant provides to the intended Customer. This can be done using a plain English description of the product,

service, or business process, including intended purposes, uses, and basic functionality. Where the boundaries are unclear, such as systems that interface with other components or related or consuming systems, the Applicant's responsibilities should be made clear. The Applicant should also make clear whether their role is as a designer (i.e., they make basic design decisions and document the result of that decision), a developer (i.e., they take a design and construct a live functioning instantiation of that design), or the deployer (i.e., they take an instantiation and deploy it in an operational environment) of the Target System. Applicants may take on multiple roles, but the description of the Target System should clarify which roles Applicants play and for which system components.

Applicants must further delineate the Target System with Functional and Non-Functional Requirements. These specific constraints define what the system is supposed to do and how it is supposed to do it, which helps narrow the scope (e.g., defining a Target System as a computer program developed by the Applicant is not as helpful as a computer program which adds numbers with up to 10100 and with a precision to 99 decimal places). Functional and Non-Functional Requirements that have no bearing on the system's Risk need not be included, though the Applicant should analyze to ascertain that. Providing Functional and Non-Functional Requirements is also part of Evidence 2.1.1 in Claim 2.

The Target System should be scoped from the perspective of the Applicant.

For example, the WordPress Foundation designs and develops the software product WordPress.Customers of this organization, such as WordPress.com, deploy the software in various environments. Some of them have customers who may design websites deployed on WordPress instances. The Applicant could be WordPress Foundation, in which case the Target System is the software (i.e., product), Wordpress.com, in which case the Target System is the hosting platform (i.e., service), or the website using wordpress.com, in which case the Target System is the website (i.e., service).

Identifying the Risks

The last part of scoping involves determining in-scope Risk. The Risk Model provides for 64 distinct risks (four Threat Actors x four sets of Harms x four types of At-Risk Parties). It is not expected that Applicants address all 64 risks. Instead, Applicants are encouraged to narrow the scope. There are two standard approaches to selecting risks. Scoping Risk cannot simply reflect whether Controls are in place, ignoring Risk(s) which have not been treated.

Approach One: Risk Subset Selection based on Intended Audience

The Applicant selects Risks relevant to the audience to whom the Standard is being applied. For instance, if the Applicant wishes to showcase to their target market the Applicant's reduction of Risk from particular types of Threat Actors, then the Applicant is free to choose those limited risks. Perhaps the Applicant is only concerned about risks related to data sharing and thus selects harms related to information dissemination. Whatever the scoping decision, the Applicant should have a reasonable justification for the scope.

Approach Two: Risk Subset Selection based on the relevance of Risks

The Applicant selects all 64 risks in the Risk Model. Then, the Applicant systematically reviews each risk and eliminates risks that are 1) not relevant to the Target System or 2) where the risks are negligible.

Some risks may not be relevant to a particular Target System. For instance, the Target System may not involve any Bystanders or Bystander data. Making a note of this could eliminate whole classes of risk. In the example where Bystanders are not relevant, 16 risks are eliminated from consideration.

Some risks may be negligible in either likelihood of occurrence or impact on At-Risk Parties. This would not include Residual Risks, the measure of risk after Controls have been applied, but Risks at the outset. For example, the Applicant might assert that risks related to information sharing harms are negligible because the information is only shared internally[10]. In general, scoping of risks is at the discretion of the Applicant, who must describe the basis for their decisions.

[10] This is not to say this is a valid Justification for dismissing information sharing related harms

Organization of the Assurance Case

The assurance case in this standard follows a Claims, Argument, Evidence (CAE) structure. Claims are subject to Arguments, which are supported by Subclaims or Evidence. Usually, the Applicant need not create a privacy assurance case from scratch - the IOPD Standards Committee has done the preliminary work in drafting this standard. Because the case has been laid out, Applicants are also not free to alter the verbiage of the Claims, Arguments, or Evidence statement, except where latitude is granted to populate the statement with selected options. Applicants are tasked with selecting specific arguments relevant to their Target System. Applicants are also tasked with selectively applying Claims, Arguments, and Evidence to in-scope Risks. Ultimately, the Applicant must provide the Evidence based on their selections. There are two deviations from this that Applicants should be aware of.

For Claim 6, “Benefits Outweigh Residual Risk”, the Applicant must, for each in-scope Residual Risk, construct Argument 6.1 for their particular Target System and context in which that Target System is designed, developed, or deployed. Depending on the Argument, the Applicant will then need to construct Subclaims, additional Arguments, and Evidence to support the Claim that Benefits outweigh Residual Risk.

Similarly, in the Privacy by Default prong of the case, Applicants must construct an argument for Claim 9: “Changes to the Configuration(s) as delivered to the Customer(s) would create an undesirable balance of Benefits and Risks”. Applicants may use the same argument structure for each Configuration delivered, though supported by differing Evidence, or Applicants may provide distinct arguments for different Configuration(s). Regardless, each Configuration delivered must be supported by the Claim that alterations would be undesirable.

Mandatory Claims, Arguments, and Evidence are indicated in the following case descriptions. They are illustrated as solid borders in the diagram in the following section. Selective Claims, Arguments, and Evidence (i.e., those where the Applicant must make a selection from in-scope risks or identified controls) are indicated as such in the case descriptions and illustrated with dashed borders in the diagram.

Case: Privacy by Design and Default

The Privacy by Design and Default Case includes one root Claim, three supporting Claims, and 10 Subclaims. Each Claim is supported by arguments that are further supported by Subclaims or Evidence. Figure 3 contains the complete case flow from Evidence to the root Claim (indicated in green). Shapes in blue indicate where the Applicant must complete the argument with their own construction. Figures 4 and 5 detail the Privacy by Design and Privacy by Default prongs, respectively, including the Claim, Argument, and Evidence statements.

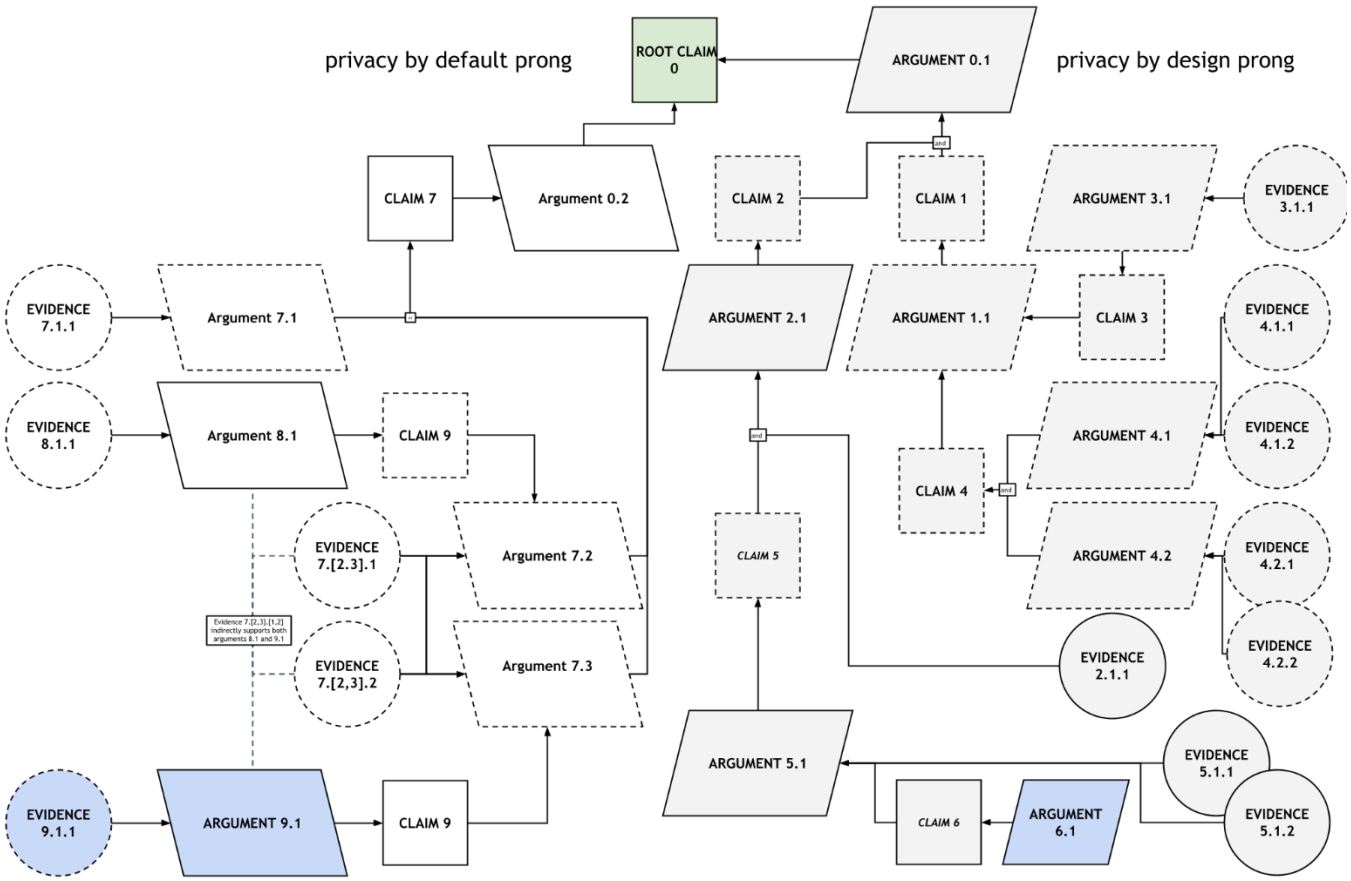

Figure 3 - Privacy by Design and Default Case Structure

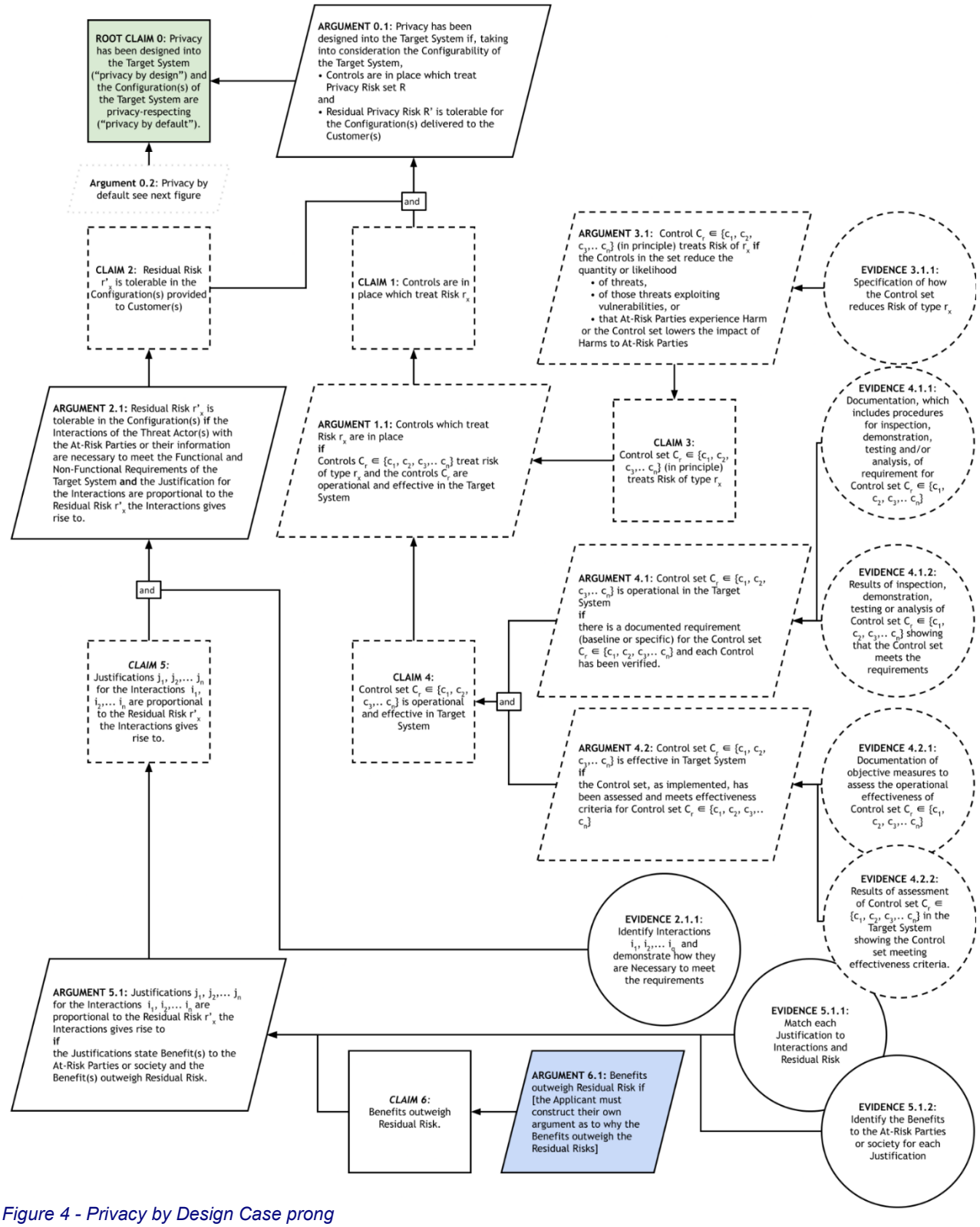

Figure 4 - Privacy by Design Case prong

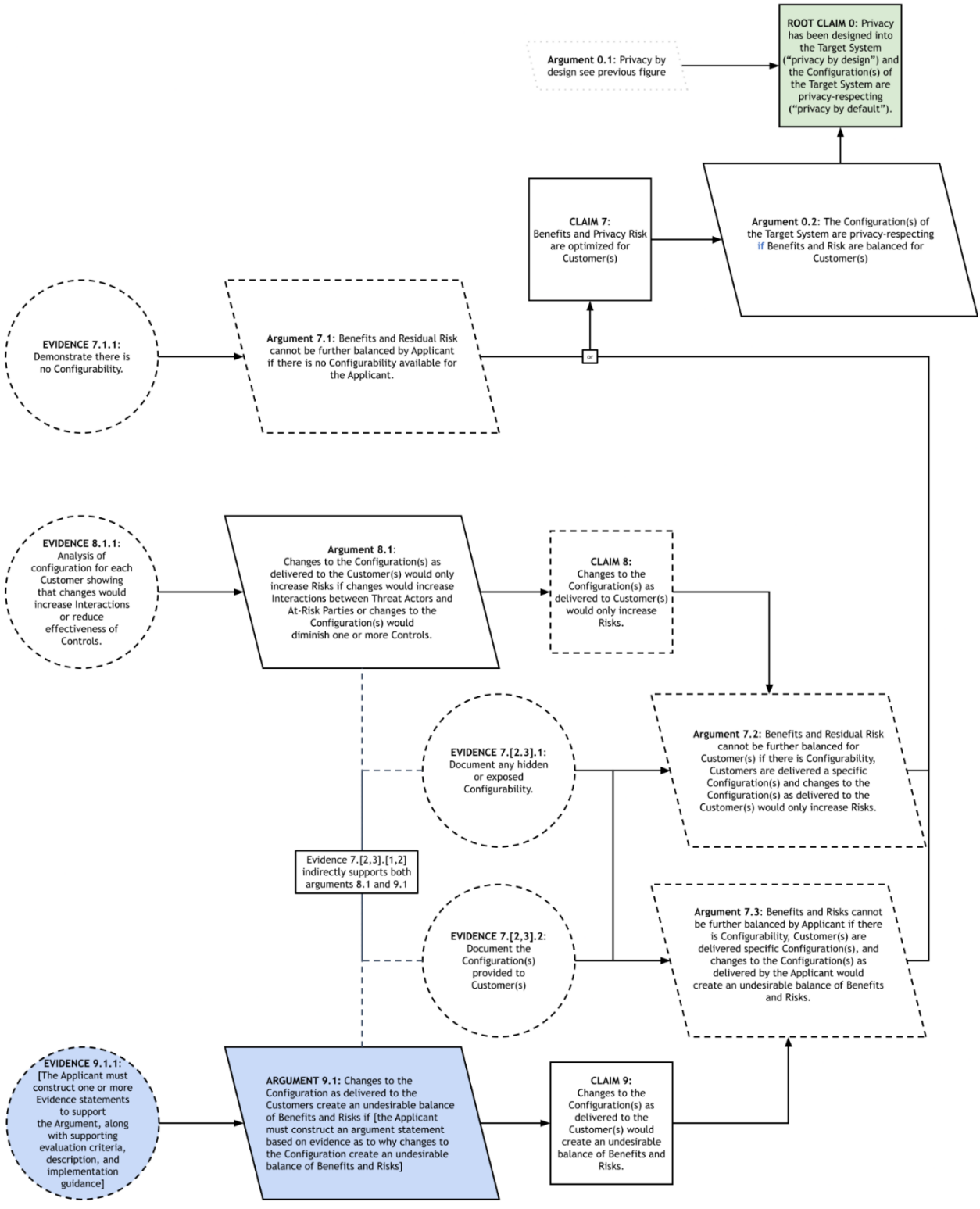

Figure 5 - Privacy by Default Case prong

Typographical Conventions Used in the Case

The following constructs are defined here:

R = the set of in-scope Risks, selected by the Applicant from the Risk Model
R' = the set of Residual Risks after Controls are applied
r_x = a specific Risk in R
r'_x = a specific Residual Risk after Controls are applied

C = the set of Controls in operation in the Target System
c_y = a specific Control in C

I = the set of Interactions in the Target System between Threat Actors and At-Risk Parties, determined by R.
i_z = a specific Interaction in I

J = the set of Justifications for the Interactions in I
j_z = a specific Justification in J for i_z

Italics are for non-normative text, typically used for examples or supplementary information, such as analogies or references to concepts in common understanding.

Root Claim

Claim 0	Privacy has been designed into the Target System ("Privacy by Design"), and the Configuration(s) of the Target System are privacy-respecting ("Privacy by Default").
Mandatory	**Description:** The root claim is essentially a restatement of the concept of privacy by design and default. There are two parts to the claim. First, that privacy has been designed into the Target System (i.e., the object of evaluation). When applied to the design, "privacy" can be thought of as a quality of the system. In other words, there is a thoughtfulness in the design that addresses privacy concerns. Applicants may note that, though the claim is about the design of the system, the subsequent arguments, sub-claims, and evidence are not about the process of the design but rather the end results. The IOPD's Design Process Standard covers the design, development, and deployment of systems with privacy taken into consideration at the beginning instead of being bolted on after. The second part of this claim restates what is meant for a Target System to exhibit "Privacy by Default." Privacy by Default is more complex than Privacy by Design. It essentially means that, as delivered to Customers of the Applicant, the Target System's settings strike a balance between Benefits and Risks.
Argument 0.1 Reasoning Step[11]	Privacy has been designed into the Target System if, taking into consideration its Configurability, Controls are in place that treat Privacy. Risk set R and Residual Risk set R' is tolerable for the Configuration(s) delivered to Customer(s).
	Description: What does it mean to design privacy into a system to address privacy concerns? First, there is an understanding by the Applicant designing, developing, or deploying the Target System that systems create a set of Risks for parties, denoted R in the Argument statement. In other words, there is a chance that some Interaction(s) within the system will occur

[11] An Argument is a reasoning step if a Claim can be deduced from a set of Subclaims. *Example: the animal can reach tree- tops if the animal is a giraffe or the animal is a flying bird*

	that will negatively impact a party and broadly that these Interactions fall under the umbrella of what's considered a Privacy Harm. Once Risks are understood, the Applicant seeks to address those risks through Controls. Controls rarely eliminate Risks but are designed to reduce them. What's left is the set of Residual Risks, denoted R' in the Argument statement. Where Risks cannot be eliminated, the Residual Risks must be tolerable (see Claim 2). R denotes the set of risks selected by the Applicant in scoping. R' denotes the set of risks in R after controls have been applied. Individual risks within these sets are denoted r_x and r'_x, respectively. The references to Configurability and Configuration in the Argument statement take into consideration that some controls or functionality of the system may be enabled or disabled by default and enabled or disabled by the Customer or others. Controls need not be enabled in the Configuration to treat risks by default. They could be disabled because, for instance, the control limits the system's functionality. **Subclaims:** Claim 1 Controls are in place which treat Risk r_x Claim 2 Residual Risk r'_x is tolerable
Argument 0.2 Tautological Step[12]	The Configuration(s) of the Target System are privacy-respecting if Benefits and Risks are balanced for Customer(s).
	Description: Privacy-respecting is not an absolute. There is a balance between Benefits (to the Applicant, to Customers, to At-Risk Parties, to society, and to other stakeholders) and Risks to parties. To respect privacy, the Applicant must balance these Benefits and Risks when delivering the Target System for use by the Customer(s). This is done within the confines of the Configurability of the Target System, giving Customers a Configuration to meet their needs without creating undue Risks for At-Risk Parties.

[12] An Argument is a tautological step if a Subclaim simply defines or restates a Claim. *Example: the zoo has an aviary if the zoo has a place to keep birds.*

	Subclaims: Claim 7 Benefits and Risks are balanced for Customers.

Claim 1 Controls are in place which treat Risk r_x

Claim 1	Controls are in place which treat Risk r_x
Selective	**Description:** Controls are actions that reduce risk. The crux of this claim is that the controls are in place in the Target System.
Argument 1.1 Reasoning Step	Controls that treat Risk r_x are in place if Controls $C_r \in \{c_1, c_2, c_3,..c_n\}$ treat Risks of type r_x, and the Controls C_r are operational and effective in the Target System.
	Description: To be in place, Controls must be designed and operating effectively to reduce or eliminate risks to parties appropriately. Not every Control treats every risk type. Appendix II provides a noncomprehensive mapping of common privacy controls and the manner in which they address risk in this standard's Risk Model. Assuming a Control treats a type of risk, it must be implemented, functional, and functioning effectively in the Target System. If all of these are true, then the claim can be justified. Note: Having a Control in place does not mean that a Control must actively prevent risk. A Configuration may enable or disable a Control. The claim here is that the controls, if enabled, will treat risk. The determination of whether a Control needs to be enabled by default is made as part of Claim 2. Residual risk is tolerable in the Configuration(s) provided to Customers. The Control may address risks not present in the default, may only be triggered if a risk materializes, or may address risks for particularly risk-averse parties. "In place" merely means it is functionally available should it be needed. **Subclaims:** Claim 3 Control $C_r \in \{c_1, c_2, c_3,.. c_n\}$ (in principle) treats Risks of type r_x Claim 4 Control c_y is operational and effective in Target System.

Claim 2 Residual Risk r'_x is tolerable in the Configuration(s) provided to Customers(s)

Claim 2	Residual Risk r'_x is tolerable in the Configuration(s) provided to Customer(s)
Selective	**Description:** Each Residual Risk remaining after application of Controls must be tolerable. The inclusion of Configuration(s) recognizes that not all risks may be present in the default due to some functionality being absent and that not all Controls need to be active if the risks they treat are not present. Customers are free to change the Configuration, subject to the Configurability of the system, to enable functionality or disable Controls to meet their own needs and risk tolerance. But, the Configuration(s) as delivered to Customer(s) must be tolerable "out of the box."
Argument 2.1 Reasoning Step	Residual Risk r'_x is tolerable in the Configuration(s) provided to Customer(s) if the Interactions of the Threat Actor(s) with the At-Risk Parties or their information are Necessary to meet the Functional and Non-Functional Requirements of the Target System and the Justification for the Interactions are proportional to the Residual Risk r'_x the Interactions give rise to.
	Description: Interactions resulting from the Configuration(s) must include only those Necessary to meet the requirements of the Target System. While additional Interactions, which introduce additional risk, may be enabled, the concern here is the Interactions contemplated while the Target System is in the specific Configuration. Necessity is the key to this part of the argument. If it is not Necessary, it should be left to subsequent configuration rather than enabled. For each of those Interactions, there must be a Justification, beyond its necessity to meet system requirements. Without a Justification, an unfounded system requirement could be established, necessitating an Interaction (e.g. Requirement: collect email addresses). Justification provides the reasoning behind the requirement and, ultimately, the Interaction (e.g., justification: to communicate with the user about their account). Further, Justification for those Interactions, both individually and collectively, must be Proportional to the Residual Risk(s) resulting from those Interactions. The proportionality of each Residual Risk is measured in Claim 5.

Evidence 2.1.1	**Subclaims:** Claim 5 Justifications j_1, j_2,... j_p for the Interactions i_1, i_2,... i_q are Proportional to the Residual Risk r'_x the Interactions give rise to. **Evidence**: Evidence 2.1.1 Identify Interactions and demonstrate how they are Necessary to meet the requirements
	Identify Interactions i_1, i_2,... i_q and demonstrate how they are Necessary to meet the requirements.
	Description: For each Interaction by each Threat Actor, in all Configuration(s) delivered to Customer(s), the Applicant must: 1. document how a Functional or Non-Functional Requirement is directly dependent on that interaction and 2. demonstrate that the same Functional or Non-Functional Requirement cannot be achieved without that interaction. **Evaluation Criteria:** The Assessor must review, on a pass or fail basis, whether the Applicant has 1. explicitly and comprehensively defined and documented each Interaction, 2. directly linked every Interaction in every Configuration delivered to Customer(s) to one or more Functional or Non-Functional Requirement, and 3. demonstrated how the removal of each Interaction results in at least one Functional or Non-functional Requirement being unachievable or severely impaired. **Implementing Guidance:** During design, development, and deployment, the Applicant should review any interactions to ensure they support one or more Functional or Non-Functional Requirements. Additionally, Applicants should complete the following steps: 1. Document each Interaction. 2. Document each unique type of Threat Actor and each of their potential Interactions with the At-Risk Parties or their data. Note that

[redacted]	types of Threat Actors are more granular than the four classes of Threat Actors in the Risk Model used. A type of Threat Actor is a grouping of unique Threat Actors that shares a common profile of Interactions (e.g., customer service agents or call centers, where the Applicant employs more than one call center). 3. Directly link each Interaction with the At-Risk Parties or their data by each group type of Threat Actor to at least one requirement. 4. For each Interaction, document how removing that Interaction results in at least one Requirement being unachievable.

Claim 3 Control set C_r treats Risk of type R_x

Claim 3	Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ (in principle) treats Risk of type r_x
Selective	**Description:** The intent of implementing controls is to reduce risk. This claim concerns a particular Control set treating a particular type of risk. This claim must be made for each Control set claimed to be in place by the Applicant and must address each Risk from the Risk Model that the Applicant has selected as in scope. While treatment of risk generally offers some broader means (e.g., transferring, accepting), the treatment in this standard is limited to the application of Controls, which reduces risk per the factors in the argument below. This Claim is selective in that Applicants must itemize the Controls they have in place in the Target System, map them to each of the Risks they have determined to be in scope, and select this Claim for each set of those Controls.
Argument 3.1 Tautological Step	Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ treats risk of type r_x if the Controls in the set reduces the quantity or likelihood • of threats, • of those threats exploiting vulnerabilities, or • that At-Risk Parties experience Harm or the Control set lowers the impact of Harm to At-Risk Parties.
	Description: Risk treatment (i.e., reduction) can occur through one of four means: <u>Reducing Threats</u>: The Control set could reduce the opportunity of the Threat Actors (e.g., the Applicant, Contracted Party, Non-contracted Party, or Other Party) to act. Without opportunity, there is no threat. *If the associated risk is an action on data, deleting the data removes the Threat Actor's ability to act on that data (i.e., no data, no opportunity).* <u>Reducing Exploitation</u>: The Control set could reduce the motivation of the Threat Actor (i.e., the likelihood that the threat exploits vulnerabilities). *A contract clause to terminate the contract in case of breach of terms will disincentivize a Contracted Party to take advantage of data they may have on*

	a party. Reducing Harms: The Control set could reduce the likelihood an At-Risk Party (e.g., Consumers, Operators, Resources, and Bystanders) experience Harms. This factor concerns threat materialization (i.e., threat materializing into a Harm). This occurs where there is some impediment (or difficulty) of the potential threat to materialize into something that impacts the party. *A Threat Actor may have data about a party and want to do something with it (i.e., the threat and the desire to exploit their possession), but if the data is encrypted, the Threat Actor will have a harder time; thus, the party is less likely to experience any Harm from the Threat Actor's actions.* Reducing Impacts: The Control set could lower the impact, not just the likelihood of an impact. *Reducing the specificity of a medical record from "the patient visited an HIV specialist" to "the patient visited a doctor" may reduce the tangible impact should that information be shared with someone.* **Evidence:** Evidence 3.1.1 Specification of how the Control set reduces the risk of type r_x
Evidence 3.1.1	Specification of how the Control set reduces Risk of type r_x
	Description: This evidence is about providing a defensible statement that each Control in the set reduces risk through one of the four means listed in Argument 3.1. The Applicant need not provide proof nor real-world evidence of risk reduction but must provide a reasonable argument that the Control addresses the Risk in the way specified. See the description of Argument 3.1 for examples. **Evaluation Criteria:** The Assessor will review each statement to determine how each Control reduces corresponding in scope Risks. Each statement must: • identify the Control, • identify the Risk, including the Threat Actor, At-Risk Party, and Consequence, • identify the means of risk reduction (threats, exploitations, harms, or impact), • state how that Control achieves the means.

	Example: *"Deleting data about Consumers held by the Applicant reduces the quantity of Information Processing harms because future information processing cannot occur without data" includes the Control, the Risk (Applicant, Consumer, Information Processing harms), means (reduction of threats) and achievement of means ("future information processing cannot occur without data").* The Assessor will further review the soundness of the achievement of means and reject those that are logically flawed or not based on available evidence. **Implementing Guidance**: Applicants should think about how Controls reduce risk during the selection (i.e., requirements phase) and design of Controls, but ultimately, the statement construction may occur solely for the benefit of an Assessor. The Applicant should review the available literature (e.g., requirements documentation, external sources about control effectiveness) to construct the statement for the benefit of post-hoc review.

Claim 4 Control set C_r is operational and effective in the Target System

Claim 4	Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ is operational and effective in the Target System
Selective	**Description:** Claim 3 covers whether there are Controls that treat Risk, but to have those risks addressed in the operation of a particular Target System, those Controls must be operational, meaning they work (*e.g., no good having a lock on a door that's broken*), and they are effective, meaning they actually reduce the risk they are meant to reduce (~~*e.g.,the*~~*e.g., the working lock isn't easily bypassed by strong push*). One important note is that a Control need not be enabled (*e.g., the door can presently be unlocked*). This is covered by the default state (*i.e., Configuration*) in which the Target System is delivered to Customers. *Continuing with the analogy, an organization could deliver a building with the door locked or unlocked, per the Customer's needs, this claim is about whether that lock works and works effectively at preventing people without keys from entering.* This Claim is selective in that Applicants must itemize the Controls they have in place in the Target System and select this Claim for each Control set. This Claim has two parts: operationality and effectiveness. Each part is supported by a separate Argument. Both Arguments must be made to establish the Claim.
Argument 4.1 Evidentiary Step[13]	Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ is operational in the Target System if there is a documented requirement (baseline or specific) for the Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ and each Control has been verified.
	Description: To claim operationality, there must be both a requirement for a Control set and verification that the Control set, as implemented in the Target System, meets the requirement. A requirement can come in the form

[13] An Argument is a evidentiary step if Evidence makes the Claim more likely than not, based on inductive reasoning. Evidentiary steps may provide a method to measure the confidence of the claim. *Example: the animal is a flying bird if (Evidence it is a bird) and (Evidence of it flying).*

	of a baseline requirement that supports all systems of the Applicant or a specific requirement in the context of the Target System. To support the existence of the requirement, the requirement must be documented, and evidence (e.g., Evidence 4.1.1) of that documentation (e.g., entries in a supporting tool, written requirements) must be presented. Further, to be considered operational, each Control must be verified through inspection, demonstration, testing, or by whatever means the requirement document or test case documentation specifies to evidence that the Control works. Verification should show that the Control is operational as designed. Effectiveness is addressed in Argument 4.2. **Evidence:** Evidence 4.1.1 Documentation, which includes procedures for inspection, demonstration, testing, and/ or analysis of the requirement for Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ Evidence 4.1.2 Results of inspection, demonstration, testing, or analysis of Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ showing that the Control set meets the requirements
Evidence 4.1.1	Documentation, which includes procedures for inspection, demonstration, testing, and/or analysis of requirements for Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$
	Description: To claim a requirement exists for a particular Control in the Target System, there must be documentation that records this requirement. The Applicant must have a copy of or reference to available documentation. Further, merely having a requirement doesn't mean that the requirement was implemented; hence, there is a need to verify that the requirement has been met and the Control set has been implemented. Verification of Control sets usually involves gathering evidence, validating evidence, analyzing evidence, and concluding whether it's operational. Evidence can be gathered through inspection, demonstration, testing, and/ or analysis. The process of verification must also be documented. **Evaluation Criteria:** To sufficiently evidence the existence of a requirement, the presented documentation must contain the following: • A description, with enough specificity to facilitate implementation, of

	the Control set; and • One or more methodologies for verifying the implementation of the Control set in a system. The methodologies for inspection, demonstration, testing, and/or analysis must be written with enough specificity to provide an objective conclusion as to whether the Control set has been properly implemented. An Assessor should review the documentation for each Control set presented to assess the sufficiency of the documentation in meeting the above criteria. Results are rendered as sufficient or insufficient. **Implementing Guidance:** Many organizations have undocumented requirements, especially when it comes to Non-Functional Requirements. While it is important to document requirements, it's imperative when designing for privacy. Being able to clearly explain why system components affecting Risk are in place is central to the claim that Residual Risk is tolerable (see Claim 2). Documentation can come in the form of some baseline system requirements policy or standard (*e.g., "all systems must be resilient and have 99.99% uptime"*). System-specific Functional Requirements are typically found in system design documents or product/sprint backlogs in Agile development. Requirements may not be formal but must be recorded in a form accessible to the designers, developers, and deployers. Quality attributes (i.e., Non-functional Requirements) are commonly found in baseline requirements documentation, such as a corporate system standards document, including any external standards the Applicant applies. For applicable external standards, there should be some documentation or evidence, such as a policy document, that demonstrates the Applicant actually uses the standard.
Evidence 4.1.2	Results of inspection, demonstration, testing, or analysis of Control set $C_r \in \{c_1, c_2, c_3, .. c_n\}$ showing that the Control set meets the requirements.
	Description: Control set requirements or test documentation must include a method of assessing the operation of the Control set in a particular system. This evidence is about showing that such an assessment took place and that the results of the assessment show that the Control set is operational. The method of assessment is left to the Applicant, the Control designer, or an

	independent body. **Evaluation Criteria:** The methodology of the Control set assessment must identify the required evidence to support evaluating Control set operationality. To determine if this evidence is sufficient, • The Control set assessment must match the assessment methodology, and • a conclusion must be rendered (and supported by evidence) that the Control set is operational In the event of a large number of Control sets, it is sufficient for the Assessor to randomly select a sufficient number of sample Control set assessments to give a 95% confidence level that all the Control sets meet the evidence criteria. Any assessment that uses statistical sampling must include the methodology and resulting confidence level. **Implementing Guidance:** Because of the vagaries of assessment of Control set operationality, it is important to have a central repository and a process for assessment performance. This process should include timely re-reviews and version control in the event of potential material changes to the Target System. The repository should note the date the assessment was performed and any relevant Target System version, as well as maintain copies of the assessment evidence and a conclusory statement. In the event of a conclusion of non-operationality of a Control set, the Control set should be reassessed after completion of any remedial actions undertaken.
Argument 4.2 Evidentiary Step	Control set $C_r \in \{c_1, c_2, c_3,.. c_n\}$ is effective in the Target System if the Control set, as implemented, has been assessed and meets effectiveness criteria for Control set $C_r \in \{c_1, c_2, c_3,.. c_n\}$
	Description: The operational effectiveness of a Control set describes how well an implemented Control set is functioning in order to mitigate specific Risks the Control set intends to treat. Effectiveness must be measured against an objective standard (i.e., Evidence 4.2.2). **Evidence:** Evidence 4.2.1 Documentation of objective measures to assess the effectiveness of Control set $C_r \in \{c_1, c_2, c_3,.. c_n\}$

Evidence 4.2.1	Evidence 4.2.2 Results of assessment of Control set $C_r \in \{c_1, c_2, c_3,.. c_n\}$ in the Target System showing the Control meeting effectiveness criteria
	Documentation of objective measures to assess the operational effectiveness of Control set $C_r \in \{c_1, c_2, c_3,.. c_n\}$

	Description: The measures of operational effectiveness of a Control set describe whether the Control set operates consistently to a specified degree. Ideally, these measures should be objective, independently available, such as through a recognized standard, and independently verifiable. The Applicant may provide their own objective measures, though this may result in enhanced scrutiny by an Assessor. This evidence is about providing either external support for the objective measures used or providing internal documentation as to how the measures work, how they are objectively measured, and how they assure operation effectiveness.
	
	Evaluation Criteria: For each Control set, a methodology must be presented to assess the operational effectiveness of the Control set. This methodology may be internal or external (*e.g., NIST 800-53A Rev.5 Assessing Security and Privacy Controls*). More scrutiny should be given to internal assessment methodologies that have not undergone peer review.
	
	To determine if the evidence is sufficient, the presented methodology must:
	• be based on objective criteria,
	• be deemed applicable by an internal or external Assessor,
	• be documented in a way that is understandable and self-contained,
	• be performable in a finite amount of time,
	• state the following:
	  ○ the type of evidence to be gathered
	  ○ the methods for gathering the evidence
	  ○ the criteria for evaluating the reliability and sufficiency of the evidence
	  ○ the process for assessing the Control set based on the evidence
	  ○ the dependencies on which the assessment may rely on
	  ○ the measure for which the Control set can be deemed effective or not effective

	In the event the provided methodology presents a way to measure operational effectiveness on a spectrum but without criteria to state whether the Control set is effective or not (e.g., "the Control set is x% effective"), the documentation must demonstrate the Applicant's determination of effectiveness with sufficient justification, in context, of why that determination was made. **Implementing Guidance:** Control effectiveness assessment methodologies should ideally be externally provided, either through a recognized standards body or from an independent entity with subject matter expertise in the Control's functions. Similarly, for Control Set effectiveness assessment methodologies, though this may be less likely, because of the combinatorial explosion possible for sets of controls. Having said that, a standard approach (e.g., Boolean algebra, 'but for' analysis, etc.) may be applicable in general, obviating the need for custom predefined assessment methodologies for specific sets of controls. Independent parties need not be the ones conducting the assessment, though this provides stronger evidence that the results (evidence 4.2.2) are unbiased.
Evidence 4.2.2	Results of assessment of Control set $C_r \in \{c_1, c_2, c_3,.. c_n\}$ in the Target System showing the Control set meeting effectiveness criteria.
	Description: This evidence is about showing that an assessment of operational effectiveness has been conducted for each Control and the set as a whole and that the results of the assessment confirm that the Control set is holistically effective. The method of assessment must match the documentation per Evidence 4.2.1. **Evaluation Criteria:** The methodology used to assess the effectiveness of the Control set must identify the required evidence to evaluate whether the Control set is effective or not. To determine if this evidence is sufficient, • the Control set assessment must match the assessment methodology, and • a conclusion must be rendered (and supported by evidence) that the Control set is effective. While the Control set assessment need not be conducted by an independent party, the relationship between the parties conducting the individual Control

	and/or Control set assessment and the operation of the Control set should be taken into account. Results should be scrutinized for any potential bias. In the event of a large number of Control sets, it is sufficient for the Assessor to randomly select a sufficient number of sample Control sets assessments to give a 95% confidence level that all the Control sets meet the evidence criteria. **Implementing Guidance:** Because of the vagaries of assessment of Control set effectiveness, it is important the Applicant maintain a central repository and a process for assessment performance. The repository should note the date the assessment was performed, maintain copies of the assessment evidence, and include a conclusory statement. In the event of a conclusion of non-effectiveness of a Control set, the Control set should be reassessed after completion of any remedial actions undertaken.

Claim 5 Justifications for the Interactions are proportional to the Residual Risk r'_x the Interactions give rise to

Claim 5	Justifications $j_1, j_2, \ldots j_p$ for the Interactions $i_1, i_2, \ldots i_q$ are proportional to the Residual Risk r'_x the Interactions give rise to.
Selective	**Description:** There is a complex relationship between Interactions and Risks. One Interaction can lead to multiple related Risks. Similarly, multiple Interactions may contribute to one Risk. For each in-scope Risk, denoted r_x, the set of Interactions (between a potential Threat Actor and At-Risk Party) that contribute to that Risk must be justified by the Applicant. A Justification is a statement indicating the reason the Interaction takes place in the Target System. Furthermore, each Justification must be Proportionate to the Residual Risk (i.e., the Risk remaining after Controls have been applied). In other words, the greater the Residual Risk, the stronger the Justification required.
Argument 5.1 Evidentiary Step	Justifications $j_1, j_2, \ldots j_p$ for the Interactions $i_1, i_2, \ldots i_q$ are proportional to the Residual Risk r'_x the Interactions gives rise to if the Justifications state Benefit(s) to the At-Risk Parties or society and the Benefit(s) outweigh(s) the Residual Risk.
	Description: The key element of this argument is that proportionality hinges on benefits to the At-Risk Party or society and that those benefits outweigh the Residual Risk. The Justification statements must have an explicit or implied Benefit. This Benefit is either directly to the At-Risk Party or society. Benefits to the Applicant cannot be considered here. Any Benefits to the Applicant come through meeting requirements, as shown in Evidence 2.1.1. Outweighing the Benefits need not be strictly based on utilitarian comparisons but may include ethical considerations, as described in Claim 10. Further, Benefits need not be siloed, and consideration of multiple Justifications may be pooled to judge the proportionality of Interaction(s) and Residual Risk(s). **Evidence:** Evidence 5.1.1 Match each Justification to Interactions and Residual Risk.

	Evidence 5.1.2 Identify the Benefits to the At-Risk Parties or society for each Justification. **Subclaims:** Claim 6: Benefits outweigh Residual Risk.
Evidence 5.1.1	Match each Justification to Interaction(s) and Residual Risk(s).
	Description: The Applicant must perform a matching exercise to ensure that every Interaction in the Target System and every in-scope Risk has at least one Justification attached. **Evaluation Criteria:** For each in-scope Risk put forth by the Applicant, the Applicant must identify the Interaction(s) in the Target System that give rise to that Risk. Further, the Applicant must link the sets of Interaction(s) and Risk(s) to one or more Justification statements. **Implementing Guidance:** Ideally, Risks are identified, mitigated, and justified as part of a risk management process during the design, development, and deployment of systems. Justifications for Interactions can generally be identified early on. The Applicant should have a good sense of the Justification prior to creating the Target System but should employ, as part of its risk management practices, a procedure to ensure that Justification continues to align with and outweigh Residual Risks.
Evidence 5.1.2	Identify the Benefits to the At-Risk Parties or society for each Justification.
	Description: Each Justification statement must explicitly or implicitly include a Benefit to At-Risk Parties or society. **Evaluation Criteria:** The Assessor will review Justification statements to ensure they: • include one or more Benefits (If they do not include an explicit Benefit, the Benefit should be reasonably obvious to the Assessor.), • clearly identify the beneficiary or have a reasonably obvious beneficiary that is apparent to the Assessor and

- social benefits must provide external validation (e.g., law, policy paper, advocacy group)

Implementing Guidance: The Applicant should compose a list of Justification statements. For each Justification statement, the list should explicitly identify each Benefit and beneficiary. An example list is provided below.

Justification Statement	***Benefit***	***Beneficiary***
The interactions allowed for a personalized shopping experience.	*Personalization, reduction in time spent finding items of interest (explicit).*	*Consumers ("Shoppers")*
The interactions disincentivize shoplifting.	*Reduction in law enforcement expenditures to investigate crime.* *Enforcement of social contract to pay for goods and services (implicit).*	*Society*

Table 2 - Table 2 Example of justifications, benefits, and beneficiaries

Claim 6 Benefits Outweigh Residual Risks

Claim 6	Benefits outweigh Residual Risks
Mandatory	**Description:** The Benefits of the product, service, or business outweigh the Residual Risks to the At-Risk Party. The measures of benefits and risks need not be done on a purely utilitarian scale but may include ethical, societal, and/or other considerations.
Argument 6.1 Evidentiary Step	Benefits outweigh Residual Risk if [the Applicant must construct their own argument as to why the Benefits outweigh the Residual Risks]
	Description: The Applicant must construct an argument as to why the benefits outweigh any risks remaining (i.e., Residual Risks) after all controls have been applied. Only benefits to At-Risk Parties or society may be considered. Benefits to the organization are tied to the necessity of Functional and Non-Functional Requirements found in **Evidence 2.1.1**. Such an argument may include a balance of interests, policy concerns, ethical factors, and opinions of the stakeholders. There need not be one argument for all benefits and risks, but these can be classified and grouped, and the Applicant may provide multiple arguments to cover the entire range of benefits and risks in the Target System. Arguments should be written in the abstract and not include specific activities and risks. Evidence as to whether specific activities and risks have benefits that outweigh those risks should be documented in 6.1.[]. The argument may also be supported by subclaims. In this case, the sub-claims should be intuitively obvious, and they need no additional supporting arguments or evidence. **Evaluation Criteria:** While normally reserved for evidence, evaluation criteria are provided here because Assessors will be tasked with evaluating the Argument statements provided by Applicants. Arguments must: • Use inductive or deductive reasoning as to why benefits outweigh risks. Such reasoning must be logically consistent and based on available evidence.

Evidence 6.1.[ ]	• Include objective evidence to support the conclusions **Implementing Guidance:** Applicants should begin with a well-founded rationale of why they feel the benefits outweigh the risks. This rationale should be formalized into a logical argument that can be objectively validated. For instance, the argument might be that beneficiaries judge the benefits to them as worth the risk. This argument would be supported by evidence of the individual choice and their informed adjudication of this choice. Note this may be a high bar. While obvious in many voluntary activities (*e.g., a party chooses to rock climb knowing the risks of death or injury*), many risks may not be so obvious, intuitive, or reasonably explained to affected parties. Particular consideration should be given to power imbalance and vulnerable groups (e.g., children, ethnic minorities, LGBTQ+, those with disabilities, or those with no choices, or who cannot stop or remove such a service, such as when dealing with the public sector, or who have to take on the service based on their Postal Code). **Evidence:** Evidence 6.1.[ ] [Evidence statement to be provided by the Applicant consistent with their Argument]
	[The Applicant must provide evidence statements to support their Arguments]
	Description: The details of the evidence will depend on the arguments provided by the Applicant in Argument 7.1. The Applicant may provide multiple evidence statements in support of their Argument. **Evaluation Criteria:** The Applicant must provide Evaluation Criteria upon which the Assessor must review the evidence. The Assessor will review two items. First, they must review whether the evidence statement logically supports the Arguments. Second, they must evaluate whether the evaluation criteria assess the sufficiency of the evidence enough to support the Arguments. Additionally, once the evidence statement and evaluation criteria are assessed, the evidence needs to be assessed against the evaluation criteria provided by the Applicant.

	Implementing Guidance: The Evidence statement should be a direct restatement of the elements supporting the Applicant's constructed argument. Evaluation criteria should be written in plain language such that an external Assessor can evaluate the sufficiency of the evidence to support the Argument and, ultimately, the claim. Evaluation criteria can consider the existence of discrete elements or an analysis of elements resulting in some level of confidence as to the truth of the elements. Applicants need not supply Implementation Guidance to themselves, though such guidance may be helpful to standardize processes related to meeting the Standard in the future, especially where such evaluation must be applied to multiple Target Systems and for multiple risks that may evolve over time. An output similar to the ICO's Legitimate Interest Assessment[14] balancing test should be considered.

[14] https://ico.org.uk/media/about-the-ico/disclosure-log/4017958/ic-109330-z1w4-attachment-2.pdf

Claim 7 Benefits and Residual Risks cannot be further balanced by Applicant

Claim 7	Benefits and Residual Risks cannot be further balanced by Applicant
Selective	**Description:** In the context of this prong of the assurance case ("privacy by default"), balancing Risks and Benefits for Customers takes into consideration the Configuration handed to Customers and the further Configurability of the Target System. Claim 7 in the "privacy design" chain covers Benefits outweighing Residual Risk, so the assumption here is that has occurred. Therefore, this claim covers whether any system Configurability provided to Customers requires them to affirmatively make changes in the Configuration to increase risk, but presumably with commensurate Benefits. In other words, the design as delivered balances risk and benefits but allows for changes to that balance. This claim is supported by one or more of these Arguments. Applicants may choose any or all but must make at least one Argument. The Applicant could, for instance, say there is no configurability in this feature of the Target System; there is Configurability in this feature, but changes would only increase risk; or there is Configurability in the remainder of the Target System, but as delivered risks and benefits are balanced. Applicants should note the use of the term Customers. Customers are the recipients of the Target System and may or may not be Consumers. For instance, if the Applicant's Customers are business entities that then provide a service to Consumers, the Customers are not the Consumers of the Applicant's Target System. See definitions to determine overlap.
Argument 7.1 Evidentiary Step	Benefits and Residual Risk cannot be further balanced by Applicant if there is no Configurability available for the Applicant.
	Description: In situations where there is no Configurability (i.e., no changes which could affect Risk), any Configuration delivered to Customer(s) may be considered balanced within the scope of Target System's design, developed version or deployed version. The reason this is considered balanced is because this is a comparative analysis, and if there are no other options to compare against (because there is no Configurability), changes to balance

	are moot. Design, development, or deployment changes **cannot** be considered at this stage in the privacy by default analysis. The Target System version is set in the privacy by design analysis, and any questions of reduced risks in the version (rather than the Configuration) are addressed in Claim 2: Residual Risk r'_x is tolerable in the Configuration(s) provided to Customer(s). **Evidence:** Evidence 7.1.1 Demonstrate there is no configurability.
Evidence 7.1.1	Demonstrate there is no Configurability.
	Description: Changes to balance are moot where there is no Configurability (and the design addresses Risks; see Claim 2), but the Applicant must demonstrate that there is no Configurability. This may be a high burden since many systems have internal settings that may be adjusted (variables and such). The question becomes which of those settings have been exposed to the Applicant. No ability to (reasonably) change settings equates to no Configurability. This does not include the ever present ability of an Applicant to reengineer a system to alter its behavior. It is sufficient if settings are not exposed in a way that the Applicant would normally engage. Designers generally have much more leeway, developers a little less so, and deployers of systems generally have the least ability to configure. Due to the vagaries of systems, demonstration may come in many forms (e.g., screenshots of non-adjustable settings, operations manuals). The Applicant should pick a method of demonstration that reasonably conveys a lack of Configurability. **Evaluation Criteria:** The Assessor should review the provided evidence and make an inference that it reasonably conveys that the Applicant has no available settings at their disposal. The Applicant does not need to provide incontrovertible proof, nor do they need to demonstrate that they cannot, through extraordinary means, alter the behavior of the system. *For example, a non-technical operator of a website need not consider the ability to alter the code running the website (or inject unexpected commands through a webform). The same would not be the case for the developer building a web application.* **Implementing Guidance:** Proper demonstration of no configurability will depend on the context of the Target System and the relationship of the

	Applicant. Since designers have the broadest leeway, it will be difficult for them to argue that the design constraints limit the configuration of the design. Developers may be able to argue that design requirements limit their developers. Deployers will have the easiest time demonstrating that the developed system they are provided by the developers provides no Configurability. *A designer of a striking device (e.g., a hammer, mallet, etc.) is limited by the materials and material sciences. They may further be limited by more specific goals (be able to strike a steel spike into concrete without breaking). Otherwise, the design is highly configurable. The designer may pass on some configurability to the engineer (i.e., the development task will bring the design to life). The engineer will still have some development options to consider. The end-user deploying the striking device at a worksite may be left with few configuration options, as the striking device is a static tool. As a counter-example, the design could allow for different headpieces depending on what's being struck, allowing for configuration by the end user.*
Argument 7.2 Evidentiary Step	Benefits and Residual Risk cannot be further balanced for Customer(s) if there is Configurability, Customers are delivered a specific Configuration(s), and changes to the Configuration(s) as delivered to the Customer(s) would only increase Risks.
	Description: If the system allows for the settings to be changed (i.e., the system has Configurability), changes will only increase the Risks, with no commensurate Benefit. For instance, a setting may turn off a control with no upside. A setting could also create a Threat where one didn't exist (such as collecting data). If there is a change in Benefits, Applicants must look to Argument 7.3 to argue that any changes would upset the balance in an undesirable way. **Evidence:** Evidence 7.2.1 Document any hidden or exposed Configurability Evidence 7.2.2 Document the Configuration(s) provided to Customer(s) **Subclaims:** Claim 8 Changes to Configuration(s) as delivered to the Customer(s) would

	only increase Risks.
Argument 7.3 Evidentiary Step	Benefits and Risks cannot be further balanced by Applicant if there is Configurability, Customer(s) are delivered specific Configuration(s), and changes to the Configuration(s) as delivered by the Applicant would create an undesirable balance of Benefits and Risks.
	Description: This argument considers the balance between Benefits and Risks with the specific Configuration provided to specific Customers. Context may vary depending on the market (e.g., business to business, business to consumer, business to government), industry, vertical, or other factors that may adjust the types of Threats, Vulnerabilities, Consequences, Threat Actors, or At-Risk Parties. Benefits need not necessarily outweigh Risk as other factors, such as ethics, fairness, or social policy, may contribute to the analysis. **Evidence:** Evidence 7.3.1 Document any Configurability hidden or exposed to Customers Evidence 7.3.2 Document the Configuration(s) provided to Customer(s) **Subclaims:** Claim 9 Changes to the configuration as delivered to the Customers would create an undesirable balance of Benefits and Risks
Evidence 7.[2,3].1	Document any hidden or exposed Configurability.
	Description: To assess the balance of a Configuration, Configurability (i.e., available settings) must be identified. This includes settings visible to Customers, Consumers, or others, whether or not they can easily access or modify those settings. It also covers any hidden options requiring advanced configuration or developer tools that may be available and utilized by the Applicant to change the default Configuration(s) for Customer(s). Each configurable element must be explained along with its options, function(s), and effect(s) within the Target System. The Applicant may also provide a Justification for why some available settings are included in the documentation, such as they are not intended to be accessible settings or

are beyond their normal skills and activities (*e.g., configuration files or changing values in code where the Applicant is not intended to be making such alterations*).

Evaluation Criteria: The Assessor will review the documentation to determine that each setting is described in sufficient detail to ascertain:

- how that setting is set
- what options are available
- to which Functional or Non-functional Requirement the setting relates
- what effect(s) does the setting have on the Target System (e.g., turn on or off functionality, security or privacy controls).

The Applicant need not address Configurability that has a negligible effect on Risks (*e.g., changing the color mode, unless changing the color mode has an effect on Threat Actor, making it more difficult for them to collect data*).

The Assessor may make an independent review of the Target System, including any environments into which the system is deployed, to ensure completeness of the documentation. This is especially important if the documentation provided by the Applicant contains noticeable gaps, lacks specificity, or has contradictory information. Design and development decisions and settings that are not visible or readily accessible may also need to be addressed; therefore Assessors should be familiar with the design and development process to determine if decisions have an impact on the Configurability (*e.g., the decision to develop a feature for iPhone and not Android distinguishes Configurations between two market segments*).

Implementing Guidance: The nature of this documentation will depend on whether the Applicant is a designer, developer, or deployer of the Target System. Deployers should look for settings provided in the Target System by the developer, either those clearly available (*e.g., an administrator's dashboard*) or described in documentation (*e.g., a configuration file*). The deployer should also consider the environment into which the system is deployed and whether settings in the broader environment might also be considered part of the Configuration of the Target System (*e.g., deploying on various databases where configuration of the database will also impact the risks of the data stored here*).

While not part of this evidence, but rather Evidence 8.[2,3],2, deployers

	must document not only the Configurability but also the options chosen. Developers, having much more leeway, need to consider not only their decisions to include Configurability into the components they develop but how that decision may also be a Configuration option. Of course, if the decision is made solely by the designers and the developers have no authority to make decisions, there is no need to document that as part of the Configurability. It becomes extremely important that developers document decisions because some of those decisions may have a significant impact.
Evidence 7.[2,3].2	Document the Configuration(s) provided to Customer(s)
	Description: This evidence requires specifying the specific Configuration(s) provided to specific types of Customer(s), detailing setting options chosen. This list should mirror that provided in 8.[2,3].1 and extends that list to include the specific options chosen and match those options to the type of Customer(s) for which that option was selected. **Evaluation Criteria:** The Assessor will review the documentation for completeness with the following considerations: • has the Applicant identified all of the Customer segments to which different Configurations are applied, and • has the Applicant identified all of the selected settings consistent with the Configurability of the Target System as described in 7.[2,3].1. **Implementing Guidance**: While documentation after the fact may be done (for instance, in a retroactive analysis of the Target System for conformance with this Standard), it is recommended that first, the Applicant keeps a running record of the Configurability of the Target System, including effects, for Evidence 7.[2,3].1, and second, has a configuration repository to retain Configuration by Customer segment. This can be automated, in part, for large, diverse deployments. An even more robust documentation system could include Justifications for those Configuration selections to avoid post-hoc Justification being provided to satisfy Claim 2.

Claim 8 Changes to Configuration(s) as delivered would only increase Risks

Claim 8	Changes to Configuration(s) as delivered to Customer(s) would only increase Risks
Mandatory	**Description:** The Target System, in the Configuration(s) that it is delivered to Customers, is set to minimize Risks and, therefore, cannot be changed to reduce said risks further. Any change(s) made to the Configuration will raise risk, which could either increase the likelihood of occurrence or impact should risk materialize.
Argument 8.1 Evidentiary Step	Changes to the Configuration(s) as delivered to the Customer(s) would only increase Risks if changes would increase Interactions between Threat Actors and At-Risk Parties or changes to the Configuration(s) would diminish one or more Controls.
	Description: Any alteration(s) to the Configuration(s) as delivered to Customer(s) increases the likelihood of Harm if said alteration(s) enables one or more Threat Actors to more easily or effectively engage with an At-Risk Party or their proxy, such as data, or such alteration(s) disables or reduces the effectiveness of one or more Controls. **Evidence:** Evidence 8.1.1 Analysis of configuration for each Customer showing that changes would increase Interactions or reduce the effectiveness of Controls (Evidence 7.2.1) and (Evidence 7.2.2). As part of the argument preceding Claim 8, the Applicant must provide Evidence 7.2.1 and Evidence 7.2.2, which are incorporated to support this Argument.
Evidence 8.1.1	Analysis of Configuration for each Customer showing that changes would increase Interactions or reduce the effectiveness of Controls
	Description: Applicant provides evidence in the form of analysis, screenshots, source code with visual output, or other attestation that

	reasonably supports the Configuration provides effective Controls. Additionally, they must demonstrate that any alteration to said Configuration will weaken Controls or raise the likelihood that one or more Threat Actors will effectively engage with At-Risk Parties. **Evaluation Criteria:** The Assessor shall adjudicate whether the evidence supplied by Applicant reasonably supports the conclusion that changes to the Configuration for each Customer would increase Interaction or weaken Controls. This is with respect to proving that Configuration for each Customer is set to minimize Threat Actor engagement with At-Risk Parties, and any alteration to said Configuration weakens Controls. **Implementing Guidance:** For each possible setting change, the Applicants need to review how that change would affect Interactions between Threat Actors and At-Risk Parties. Applicants are not required to consider every possible Configuration but should reasonably anticipate where individual changes to settings do not increase Interactions, but where multiple setting changes could have such an effect. The Applicant needs to also consider the effects of setting changes on Controls. Similarly, the primary focus is on individual settings disabling or directly weakening Controls. However, it is important to reasonably consider that the cumulative effect of multiple settings changes could weaken a Control where individual settings may not.

Claim 9 Changes to Config(s) would create undesirable balance of Benefits and Risks

Claim 9	Changes to the Configuration(s) as delivered to the Customer(s) would create an undesirable balance of Benefits and Risks
Mandatory	**Description:** Settings are features of the Target System that can be enabled or disabled. The state of these settings is a Configuration. Different Configurations may be delivered to different Customers. Whether the position of a setting enhances or diminishes Benefits or increases or decreases Risk depends on the specific setting's effect in the context of the design. This claim statement makes the assertion that changes to settings would be undesirable when considering the Benefits and Risks involved. Note that this is not a one-size-fits-all for every Configuration delivered to every type of Customer. Different Customers operate in different markets, industries, and verticals, and thus engender different risks, with varying likelihoods and impacts to At-Risk Parties. Similarly, Benefits may be heavily dependent on these contextual factors as well; thus, each Configuration should be viewed in light of the particular context in which it is deployed.
Argument 9.1 Evidentiary Step	Changes to the Configuration as delivered to the Customers create an undesirable balance of Benefits and Risks if [the Applicant must construct an argument statement based on evidence as to why changes to the Configuration create an undesirable balance of Benefits and Risks]
	Description: The Applicant must construct an argument as to why the balance between Benefits and Risk would be undesirable if the Configuration delivered to the Customer were altered. Only Benefits to At-Risk Parties or society may be considered. Benefits to the organization are tied to the necessity of Functional and Non-Functional Requirements found in Claim 2. Such an argument may include a balance of interests, policy concerns, ethical factors, and opinions of the stakeholders. There need not be one Argument for all Configurations, Benefits, and Risks, but these can be classified and grouped, and the Applicant may provide multiple Arguments to cover the entire range of Benefits and Risks in the delivered Configurations. The Argument may also be supported by subclaims. In this case, the sub-

claims should be intuitively obvious, and they need no additional supporting Arguments or evidence.

Evaluation Criteria: While normally reserved for evidence, evaluation criteria are provided here because Assessors will be tasked with evaluating the Argument statements provided by Applicants. Arguments must:

- use inductive or deductive reasoning as to why the balance between Benefits and Risks would be undesirable. Such reasoning must be logically consistent and based on available evidence.
- include objective evidence statements to support the conclusions

The question of desirability need not be made from any party's perspective. In other words, the argument does not need to consider the subjective desires of any one party. The argument should appeal to normative ethical principles, legal or moral obligations, considerations of fairness, equity, and justice, and utilitarian weighing of Benefits and Risks. There is currently no agreed-upon the construction of an objective argument for the undesirability of a resulting Configuration change, thus it is up to the Applicant to demonstrate that they have thought about it and the Configuration was not the result of an accident, ignorance or unsavory motivations, but rather careful deliberation.

Implement Guidance: Applicants should begin with a well-founded rationale of why they believe changes to the Configuration would be undesirable from a Benefits and Risk perspective. This rationale should be formalized into a logical argument that can be objectively validated. For instance, one argument might be that the affected parties judged the benefits to them as worth the risk. This Argument would be supported by evidence of the parties' choice and their informed adjudication of this choice. Note this may be a high bar. While obvious in many voluntary activities (e.g., an individual chooses to rock climb knowing the risks of death or injury), many risks may not be so obvious, intuitive, or reasonably explained to individuals.

Evidence

Evidence 9.1.[ ] [Applicant must provide Evidence Statement, Description, Evaluation Criteria and Implementation Guidance]

Evidence 9.1.[ ]	Evidence 7.3.1 and Evidence 7.3.2 As part of the Argument preceding Claim 9, the Applicant must provide Evidence 7.3.1 and Evidence 7.3.2, which are incorporated to support this Argument.
	[The Applicant must construct one or more Evidence Statements to support the Argument, along with supporting evaluation criteria, description, and implementation guidance]
	Description: The details of the evidence will depend on the arguments provided by the Applicant in Argument 10.1. The Applicant may provide multiple evidence statements in support of their Argument. **Evaluation Criteria:** The Applicant must provide evaluation criteria upon which the Assessor must review the evidence. The evaluation criteria supporting this Argument should be substantially the same as the evaluation criteria provided in Evidence 6.1[ ]. The Assessor will review two items. First, they must review whether the evidence statement logically supports the Arguments. Second, they must evaluate whether the evaluation criteria assess the sufficiency of the evidence enough to support the Arguments. Additionally, once the evidence statement and evaluation criteria are assessed, the actual evidence needs to be assessed against the evaluation criteria provided by the Applicant. **Implementing Guidance:** The Evidence statement should be a direct restatement of the elements supporting the Applicant's constructed argument. Evaluation criteria should be written in plain English so that an external Assessor can evaluate the sufficiency of the evidence to support the Argument and, ultimately, the claim. Evaluation criteria can consider the existence of discrete elements or an analysis of elements resulting in some level of confidence as to the truth of the elements. Applicants need not supply implementation guidance to themselves, though such guidance may be helpful to standardize processes related to meeting the Standard in the future, especially where such evaluation must be applied to multiple Target Systems and for multiple risks that may evolve over time.

Appendix I Definitions

Where definitions come from external sources, those sources are referenced in footnotes.

Applicant

The Role that applies the standard to the Target System. The Applicant designs, develops, or deploys the Target System.

Argument

Reasoning that provides the bridge between what is known or is assumed (Subclaims, Evidence) and the *Claim* being asserted. Note that "Argument" is an overloaded word. It is used with a specific meaning here.[15]

Assessor

The party that evaluates an Applicant's conformance to the standard. Assessors may be internal (a department or individual employed by the Applicant) or external (a party contracted by the Applicant to review their conformance).

At-Risk Party

A Role impacted by a Harm because of their Role in the Target System. While generally an individual (*i.e., natural person*) is at risk, the term here is not limited and may be used, in context, by the Applicant to refer to a non-natural person, such as a business, that can be impacted by a Harm.

Benefit

A desired consequence of an Interaction.

Bystander

A Role whose existence is immaterial to the operation of the Target System. *An example would be a person in the background of a photograph.*

[15] https://claimsargumentsevidence.org/notations/claims-arguments-evidence-cae/

Claim

An assertion about a property of the Target System. A Claim is **Mandatory** if it is required by this Standard. A Claim is **Selective** if required by this Standard but it allows the Applicant to select the specifics of the Claim. A **Subclaim** is a Claim that is made as part of an Argument supporting another Claim.

Configurability

The ability to change settings in the Target System. *In the illustration below, a lowered switch cover prevents changes to the configuration of the switches. As used in this standard, an Applicant chooses the configurability of the system (which switch covers are up or down) to enable the Customer to configure the system (turn switches on and off). The Applicant also chooses the configuration (the initial setting of switches to on and off) of the system as delivered to the Customer. Note, some settings may be set by the Applicant that are not configurable by the Customer (the switch is on or off, but the cover is closed)*

Figure 6 - A physical set of switches that can be enabled or disabled. The settings of those switches represent a potential configuration of the system. The switch covers represent the configurability of the system. Lowering a switch cover is analogous to removing the configurability of a particular setting (the switch is set to whatever setting was made before the switch cover was closed).

Configuration

System settings choices are made regardless of Configurability. *In Figure 7, the switch settings (whether they are on or off) represent the Configuration, regardless of whether it is further configurable (which is dependent on the position of the switch cover).*

Consequence

A desired (Benefit) or undesired (Harm) result of an Interaction.

Consumer

A Role that receives Benefit from the Target System (i.e., they consume the output of the Target System).

Contracted Party

A Role in contract (directly or indirectly) with the Applicant. *Contracted Parties include vendors, clients, partners, employees, contractors, and their vendors, clients, employees, and contractors, and others. Contracted parties are sometimes referred to as third parties (or fourth, fifth, etc. parties), but this term can be ambiguous in certain contexts, so it is avoided in this Standard.*

Control

An action taken by the Applicant to reduce Risk. Controls are organized into two types, System Controls and Environmental Controls, though some actions may satisfy both types. For the purposes of this Standard, Controls are limited to System Controls, and any reference to Controls means System Controls. Environmental controls are often implemented by the Applicant, Customer, or Other Parties and affect the environment in which the Target System operates. *For instance, a Control that restricts the sale of the Target System in repressive regimes would be an Environmental Control. Other Environmental Controls would be a process to conduct a risk assessment of the Target System or an enterprise access management policy. System Controls are implemented within a Target System. Examples include probabilistic, risk-based, access controls, or a contract dictating terms with a vendor within the Target System. As systems may be sociotechnical, System Controls can but need not be technical.*

Customer

A Role that receives, from the Applicant, a Configuration of the Target System. The term 'receives' includes license, lease, purchase, and other forms of procurement and does not require payment. Customers need not operate the Target System. They may be a distributor, resellers, installers, or otherwise repurpose the Target System for further delivery.

Evidence

An artifact that establishes facts that can be trusted and lend confidence to the truth of a *Claim*. In projects, there can be many sources of information, but what makes this evidence is the support or rebuttal it gives to a Claim.[16]

Functional Requirement

A defined constraint on a system that affects the system's environment outside the system boundary.

Harm

An undesired privacy-related Consequence of an Interaction with an individual. While group and societal harms are possible consequences, this Standard focuses on the Harms specified in the Risk Model.

Interaction

An action between a Threat Actor and At-Risk Parties or their proxies (e.g., data related to those At-Risk Parties).

Justification

A statement supplied by the Applicant as to why an Interaction should be allowed in light of the potential Risks. *In GDPR parlance, Justifications are a combination of purposes of processing activities and legal bases.*

Necessary

A characteristic of a proposed Interaction based on the need to meet Functional Requirements or Non-Functional Requirements of the Target System.

Non-Functional Requirement

A defined constraint on a system that relates to a desirable property or quality attribute within the system boundary. *Non-Functional Requirements describe desired characteristics rather*

[16] Ibid

than desired functions. An example is accessibility, where an organization may want its product or service accessible to those with temporary or permanent capability losses.

Non-contracted Party

A Role not in contract with the Applicant but contemplated as part of the Target System. *An internet service provider (ISP) for an internet-connected device would be a Non-contracted party, contemplated as needed by the Target System, but not in contract with the Applicant designer.*

Other Party

Any Role not performing a function in the Target System but that could interact with an At-Risk Party or their proxy, such as data related to the party, by virtue of the Target System's operation.

Proportionate

The notion that a measure of the Applicant's supplied Justification exceeds a measure of the Risk against which it is compared. Note that measures need not be a simple risk calculation but may include factors of equity, fairness, or other ethical concerns.

Operator

A Role that operates the Target System to produce Benefit for others. Conventionally, this is a worker whose labor is used to produce the output of the Target System.

Resource

A party whose existence is material to the Target System. *An example would be a data subject of a data brokerage service*.

Risk

A measure of likelihood and severity of Harm using the Risk Factors under the Risk Model.[17] For the purposes of this standard, Risk refers only to privacy-related risks.

[17] Design Process Standard, v 1.0 - available at https://instituteofprivacydesign.org/certification-standard/

Residual Risk

A measure of Risk remaining after a change in the context, such as applying Controls.

Risk Factor

A characteristic used in a Risk Model as an input to determining the level of risk in a risk assessment.[18]

Risk Model

A representation that elaborates key terms and abstract factors that contribute to or negate Harms (see NIST definition).[7] This standard uses a specified Risk Model (see section 6).

Role

A party's relationship to the Target System. A party may play multiple roles within one Target System.

Target System

The system designed, developed, or deployed and scoped for evaluation under this Standard.

Threat Actor

A Role whose action could result in a Harm to an At-Risk Party. This standard defines four categories of Threat Actors: Applicant, Contracted party, Non-Contracted Party, and Other Party. Threat Actors may be further classified by the Interactions they engage in (*e.g., Contracted party call centers*). *Threat actors in the privacy context are not the hackers and cybercriminals found in a cybersecurity context. A threat actor in a privacy context is any party that could interact with an individual or their data, resulting in harm to that individual.*

Threat

A potential action by a Threat Actor that, if realized, could result in Harm(s) to At-Risk Parties. For the purpose of the Risk Model used in this Standard, Threats are implied from the category of potential Harm: processing of data, dissemination of data, attempted collection of data, and invasions into personal space or autonomy.

[18] https://csrc.nist.gov/glossary/term/risk_factor

Vulnerability

A condition or state that puts a party at risk of experiencing Harm. For the purposes of the Risk Model used in this standard, there are two vulnerabilities that arise from Interactions: (1) Threat Actor interacts with a party or their data, and (2) Threat Actor has control, though not necessarily possession, of a party's data. *An at-risk party's vulnerability to a threat actor, such as because the threat actor has data about the at-risk party, should not be conflated with system vulnerabilities (i.e., weaknesses).*

Appendix II Entity Relationship Guide

Figure 8 is provided to help readers visualize[19] the relationships between various defined entities and attributes. The primary relationship is between the Applicant and the Target System (shown in bold in the figure), of which there is only one of each for the purposes of this Standard.

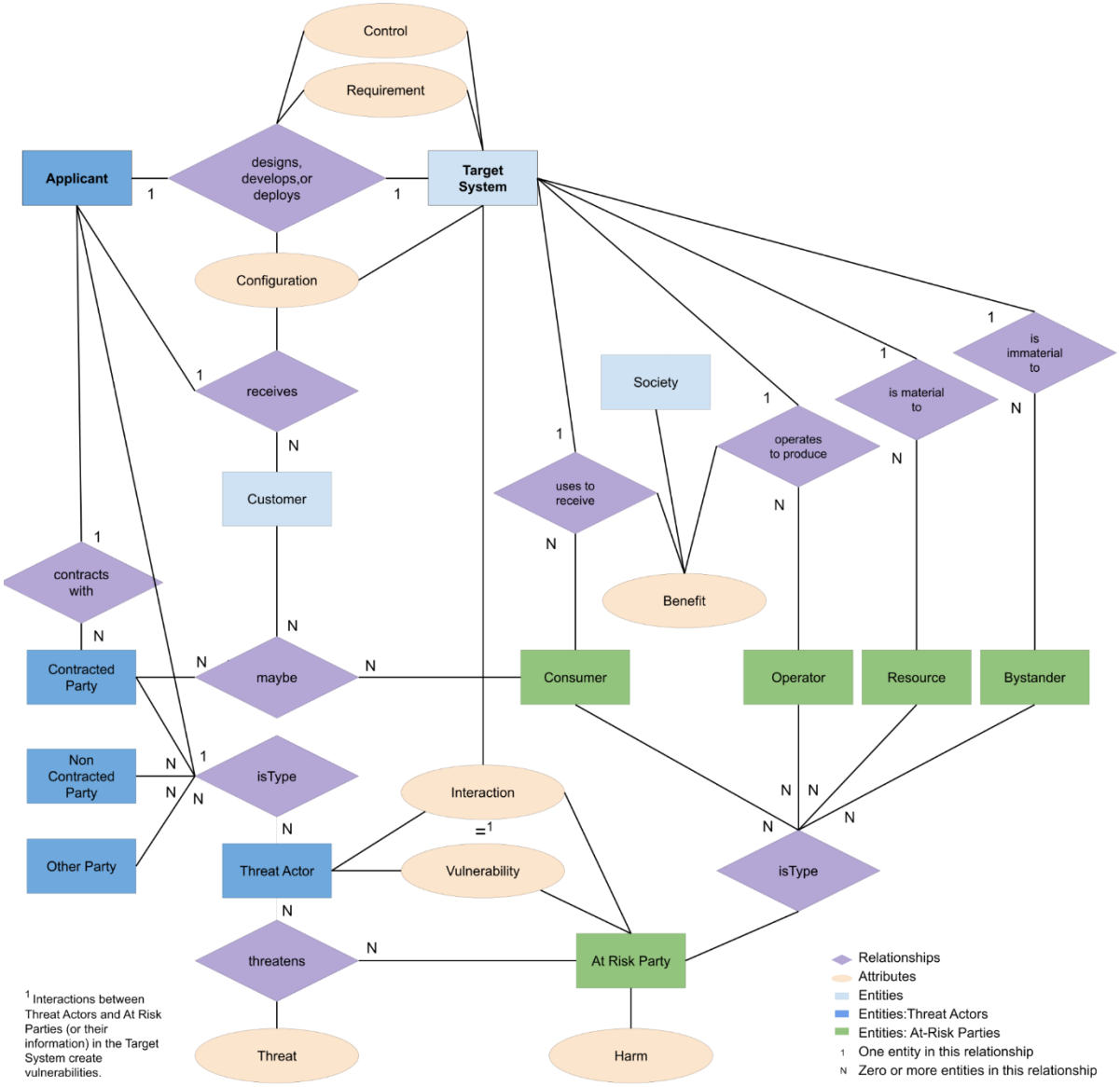

Figure 7 - Figure 8 Entity Relationship diagram showing the relationship between entities and their attributes

[19] See Wikipedia entry on Entity Relationship Diagrams, available at https://en.wikipedia.org/wiki/Entity%E2%80%93relationship_model