Contribuciones desde la ingeniería de software y sistemas basada en modelos a la privacidad y la protección de datos mediante un enfoque multidisciplinar

Presentado al Premio de Investigación en Protección de Datos Personales «Emilio Aced»

Autor: Yod Samuel Martín García

Contenido

1 Introducción 4
1.1 Propósito y estructura de este documento 4
1.2 Motivación y necesidad del trabajo de investigación 7
1.3 Objetivos 12
1.4 Métodos aplicados 15
2 Estado general del arte y de la práctica 17
2.1 Resumen de conceptos 19
2.2 Definición operativa y modelo de adversario para la privacidad y la protección de datos 24
2.3 Paradigmas dominantes actualmente para abordar la privacidad y protección de datos 28
3 Evaluación de necesidades y análisis de deficiencias 32
3.1 Método para la captura de necesidades y análisis de deficiencias: modelos mentales orientados a tareas 32
3.2 Análisis de deficiencias en el desarrollo de sistemas de los métodos y herramientas de gestión de privacidad y protección de datos 35
4 Ingeniería basada en modelos para la privacidad y la protección de datos 39
4.1 Ingeniería de software y sistemas basada en modelos (MBSSE) 39
4.2 Lenguaje específico de modelado de aspectos de para la privacidad y la protección de datos 42
5 Minimización de datos y métricas de riesgos de privacidad y protección de datos 46
5.1 Método de análisis cuantitativo de contenidos 46
5.2 Definición multidimensional de la minimización en la privacidad y la protección de datos 47
5.3 Métricas de ámbito organizativo, orientadas a riesgos, y centradas en los sujetos interesados, para la privacidad y la protección de datos 48
6 Ingeniería de requisitos de privacidad y protección de datos 53
6.1 La operativización como método para derivar requisitos de fuentes estándar 53
6.2 Enfoque basado en riesgos y enfoque orientado a metas para el análisis de requisitos de privacidad y protección de datos 54
6.3 Operativización de los requisitos del RGPD a través del modelado de la norma ISO/IEC 29100 y la correspondencia entre ambos 56
7 Lenguaje de patrones de diseño para la privacidad y protección de datos 58
7.1 Patrones de diseño: motivación, conceptos y método 58
7.2 Lenguaje de patrones de privacidad 60
8 Modelos normativos reutilizables para el aseguramiento de la privacidad y protección de datos 63
8.1 El potencial del aseguramiento de sistemas para la privacidad y la protección de datos 63
8.2 Escenario de aplicación: método y resultados 65
9 Ingeniería de métodos de privacidad y protección de datos 68
9.1 Construcción de métodos mediante la ingeniería situacional de métodos 68
9.2 Metamodelo de ingeniería de privacidad 69

9.3 Descripción funcional de métodos de ingeniería de privacidad 71
10 Valoración, trabajo futuro y conclusiones 78
10.1 Resumen de resultados obtenidos 78
10.2 Orientación de políticas para la CE (*policy brief*) 80
10.3 Transferencia tecnológica 81
10.4 Agenda de investigación 82
10.5 Conclusión 83
Referencias 86

Índice de figuras

Figura 1. Estructura del documento. 5
Figura 2. Anuncios luminosos en el aeropuerto de Bruselas, publicitando las características de gestión de datos personales en los servicios de Google 8
Figura 3. Relevancia a lo largo del tiempo de las consultas realizadas en el buscador de Google en algunos temas relacionados con la privacidad. 9
Figura 4. Collage de diferentes noticias de medios profesionales orientados a la ingeniería que destacan el papel del RGPD (y de la privacidad en general) en las actividades de ingeniería de software y sistemas. 10
Figura 5. Representación gráfica de nuestra misión. 13
Figura 6. Línea temporal de las distintas conceptualizaciones de la privacidad. 18
Figura 7. Modelo de adversario para la ingeniería de privacidad y protección de datos. 26
Figura 8. El ingeniero tras los roles en relación con la privacidad y la protección de datos. 29
Figura 9. La brecha de la privacidad en la ingeniería 30
Figura 10. Notación de los diagramas de Modelos Mentales. 34
Figura 11. Modelo Mental del Soporte al RGPD. 36
Figura 12. Varios modelos utilizados en ingeniería. (de izquierda a derecha y de arriba abajo): representaciones gráfica y formulada del campo eléctrico, modelo polifunicular de la basílica de la Sagrada Familia de Barcelona, modelo en miniatura de una locomotora, y diagrama de clases UML que representa el modelo de datos de un sistema software. 40
Figura 13. Mapa conceptual de la ingeniería de software y sistemas basada en modelos (MBSSE) 41
Figura 14. Dimensiones de la minimización de datos personales. 47
Figura 15. Ilustración del comportamiento contraintuitivo de la valoración del riesgo de privacidad basada en la entropía de la información. 50
Figura 16. Análisis diferenciado de los riesgos extremos en el método PMRM 50
Figura 17. Resultado de la aplicación del AHP para comparar las ofertas de servicio de diferentes proveedores de servicios en la nube respecto a varios criterios de privacidad, y adecuación a los requisitos de los diferentes consumidores de servicios en la nube. 52
Figura 18. Operativización de requisitos de privacidad mediante su descomposición jerárquica y su correspondencia con distintas fuentes de requisitos. 54
Figura 19. Enfoques de análisis de requisitos en el ciclo de vida de desarrollo: orientado a metas y basado en riesgos. 55
Figura 20. Leyenda del libro de códigos para la extracción mediante QCA de requisitos a partir de la norma ISO/IEC 29100. 56
Figura 21. Modelo conceptual de los elementos de la norma ISO/IEC 29100 relacionados con el consentimiento y la elección, se han añadido resaltados en el modelo los argumentos de la tabla anterior. 57
Figura 22. Evolución de la cantidad y madurez de publicaciones académicas sobre elementos reutilizables de diseño para la privacidad [3]. 59
Figura 23. Modelo estructural de los patrones de diseño de ingeniería de privacidad. 60
Figura 24. Capturas de pantalla de los patrones publicados por el autor en el portal privacypatterns.org. 62
Figura 25. Relaciones entre los patrones que abordan la estrategia de “Control” 62
Figura 26. Proceso de alto nivel para el aseguramiento. 64

Figura 27. Representación simplificada de los elementos del metamodelo de los marcos de referencia. 64
Figura 28. Elementos del código para la elaboración de modelos de marco regulatorio y patrones de argumentación 65
Figura 29. Modelo del marco de referencia del Art. 35.1 y 35.2 del RGPD. 66
Figura 30. Patrón de argumentación para el artículo 35.1 del RGPD 67
Figura 31. Modelo del marco de referencia obtenido a partir de la plantilla de EIPDs para red eléctrica inteligente (smart grid). 67
Figura 32. Metamodelo estructural de las extensiones a SEMDM para el marco de ingeniería de privacidad. 70
Figura 33. Rol / Tipo de Rol “Ingeniero de Privacidad”. 70
Figura 34. Diagrama SIPOC de las funciones de gestion de riesgos. 72
Figura 35. Diagrama SIPOC de las funciones de ingeniería de requisitos 73
Figura 36. Diagrama SIPOC de las funciones de diseño dirigido por modelos. 74
Figura 37. Diagrama SIPOC de las actividades de aseguramiento del sistema. 75
Figura 38. Casos de uso abstractos en los métodos de ingeniería de privacidad. 76
Figura 39. Arquitectura propuesta para una cadena de herramientas software para el apoyo a los métodos de ingeniería de privacidad y protección de datos. 77
Figura 40. Transferencia tecnológica de los resultados a través del proyecto PDP4E. 82

Índice de tablas

Tabla 1. Historia de las familias de principios de privacidad. 23
Tabla 2. Definición operativa de la privacidad construida a partir de las dimensiones de la privacidad como concepto esencialmente controvertido. 25
Tabla 3. Modelo de adversario para la ingeniería de privacidad y protección de datos. 27
Tabla 4. Indicaciones para las entrevistas semidirigidas a los ingenieros de software y sistemas acerca de sus tareas en relación con la privacidad y protección de datos. 33
Tabla 5. Correspondencia con el RGPD de las disciplinas de la ingeniería de software. 37
Tabla 6. Estereotipos de privacidad y protección de datos que se pueden adjuntar a los elementos de un modelo de sistema. 43
Tabla 7. Otros puntos de vista del modelado relevantes para la privacidad y la protección de datos. 45
Tabla 8. (Página siguiente) Modelado del principio de consentimiento y elección de la ISO/IEC 29100 (resaltados los argumentos de cada requisito que se deben reemplazar al instanciarlos; texto original en inglés de la fuente analizada). 56
Tabla 9. Relaciones posibles entre los patrones de privacidad. 61
Tabla 10. Contribuciones originales del trabajo de investigación. 78

1 Introducción

1.1 Propósito y estructura de este documento

Este documento presenta un resumen extendido y traducido al castellano de la tesis doctoral [1] titulada *Contributions to Privacy and Data Protection from a Multi-Disciplinary Model-Based Software and Systems Engineering Approach* (Contribuciones a la privacidad y la protección de datos desde un enfoque multidisciplinar y de ingeniería de software y sistemas basada en modelos), para su presentación al Premio de Investigación en Protección de Datos Personales Emilio Aced 2023, convocado por la Agencia Española de Protección de Datos (AEPD). Parte de los contenidos aquí presentados se han publicado en distintos artículos en revistas de impacto y ponencias en conferencias internacionales [2]–[16].

El trabajo de investigación que presentamos ofrece una serie de contribuciones en el campo de la ingeniería de privacidad y protección de datos, motivadas por la promulgación del Reglamento General de Protección de Datos de la UE (RGPD). A pesar de las relevantes implicaciones que esta norma acarrea en el desarrollo de software y sistemas, no ha venido acompañada de herramientas para que los ingenieros que crean estos sistemas apliquen en la práctica de su trabajo los principios de responsabilidad proactiva y de protección de datos desde el diseño y por defecto. Por ello, planteamos la necesidad de **introducir métodos, herramientas y técnicas en la práctica cotidiana de la ingeniería de software y sistemas** que faciliten el abordaje de las cuestiones de privacidad y protección de datos desde el inicio del desarrollo de un sistema, en lugar de dejarlas de lado hasta el momento del despliegue y la puesta en operación de un servicio. En particular, nos hemos centrado en integrar la privacidad y la protección de datos en **cinco disciplinas** de la ingeniería de software y sistemas: *gestión de riesgos, ingeniería de requisitos, diseño, aseguramiento de sistemas e ingeniería de métodos*. Para ello, aprovechamos el enfoque de la **ingeniería de software y sistemas basada en modelos** (MBSSE – Model-Based Software and Systems Engineering), que ha demostrado su adecuación para organizar sistemáticamente el conocimiento del sistema y de su contexto, integrar puntos de vista propios del sistema con otros orientados a aspectos transversales, y facilitar el procesamiento humano o automatizado de los modelos.

Esta investigación se inspira en la *necesidad de crear puentes entre la comunidad legal y la tecnológica*. Hemos mencionado que los ingenieros de software y sistemas se ven fuertemente afectados por el RGPD, al ser quienes desarrollan las soluciones de las organizaciones que llevarán a cabo tratamientos de datos personales. Sin embargo, existe una **brecha epistémica** entre su forma de trabajar y la del personal jurídico, dado que los ingenieros están acostumbrados a trabajar con conceptos tales como *bases de datos*, *casos de uso*, *requisitos*, *amenazas*, *patrones*, o *argumentaciones lógicas*, que distan mucho de lo que pueden encontrar en un texto legal como el RGPD. Además, las contribuciones a la privacidad desde el mundo de la informática y la ingeniería se suelen dirigir a los usuarios de los servicios (mediante las tecnologías de mejora de la privacidad o PETs - Privacy Enhancing Technologies), pero olvidan las necesidades de quienes están creando esos servicios. En nuestro caso, por el contrario, consideramos que *se debe involucrar completamente a los ingenieros en el cumplimiento de la privacidad y la protección de datos*, para que creen sistemas que cumplan con el RGPD, se adhieran a los principios de protección de datos y protejan los derechos de los interesados.

Así, nuestro trabajo asume la **misión** de trasladar a la práctica general de la ingeniería de software y sistemas los conocimientos asentados en privacidad y protección de datos, proporcionando a los ingenieros métodos y herramientas que operativicen los principios y la regulación de la protección de datos, y que se integren con los que ya usan habitualmente en las actividades que tienen lugar a lo largo de las etapas del SDLC (Ciclo de Vida de Desarrollo del Sistema), por lo tanto, llevando a la realidad los principios de responsabilidad proactiva y privacidad y protección de datos desde el diseño.

Nuestra investigación ha producido como **resultados** una serie de *modelos y métodos, propios del ámbito de la ingeniería de software y sistemas*, que introducen la privacidad y la protección de datos en las mejores prácticas de MBSSE. No obstante, más allá de ello, el esfuerzo por *acercar las dos 'orillas' de la comunidad de profesionales de la protección de datos* nos ha llevado a trabajar con facultades de Derecho en proyectos de investigación, emplear métodos más propios de la investigación humanística (como el análisis cuantitativo de contenido QCA o el análisis de modelos mentales), y en definitiva, adquirir las competencias necesarias para poder realizar un análisis en profundidad del RGPD y de otras normas relacionadas, de modo que pudiéramos abordar con confianza nuestra misión y ofrecer resultados que realmente respalden la implementación del RGPD en el proceso de desarrollo de software y sistemas. De hecho, uno de nuestros resultados es un documento de orientación de políticas (*policy brief*) para la ingeniería de privacidad y protección de datos que se le ha hecho llegar a la Comisión Europea. Por último, existen también herramientas software que utilizan habitualmente los ingenieros y donde se han integrado los resultados de nuestra investigación, pero, en este caso, se trata de herramientas creadas por terceros que han aplicado nuestro enfoque —y que, además, sirven como validación de nuestros resultados—.

El presente documento está **organizado en 10 capítulos** cuya relación se muestra en la Figura 1. Los primeros capítulos (1 a 3), aun cuando incluyen varias contribuciones originales del autor, sirven fundamentalmente para motivar la necesidad de esta investigación y las deficiencias existentes respecto al estado del arte. El capítulo 4 sirve como núcleo en el que engarza el resto del documento, al presentar el enfoque de ingeniería basada en modelos en el que incardinan las demás contribuciones. Los capítulos 5 a 9 presentan las contribuciones respectivas a cada una de las cinco disciplinas de la ingeniería del software y sistemas en cuyo ejercicio pretendemos introducir la privacidad y protección de datos: tres de ellas dirigidas al producto (gestión de riesgos, ingeniería de requisitos, y diseño) y las otras dos dirigidas al proceso de desarrollo (aseguramiento de sistemas e ingeniería de métodos). Finalmente, el último capítulo evalúa los resultados obtenidos y los proyecta hacia el futuro.

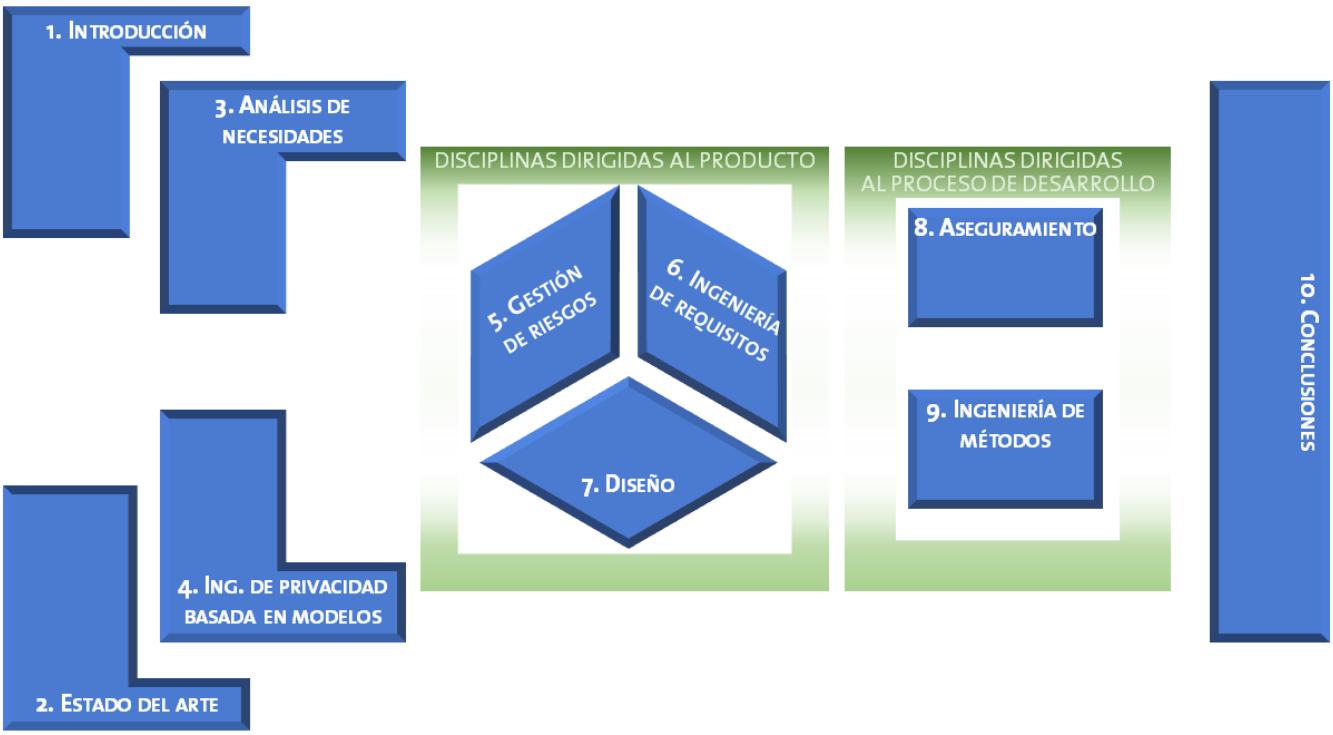


Figura 1. Estructura del documento.

A continuación, se explica brevemente el contenido de cada capítulo:

1. **Introducción.** Se introduce la ingeniería de privacidad y protección de datos como el trasfondo general y la motivación que enmarca nuestro trabajo de investigación, describe los objetivos de investigación asociados, y explica el enfoque conceptual de nuestro trabajo, relacionándolo con las cinco disciplinas de ingeniería de software y sistemas donde pretendemos introducir la privacidad y la protección de datos, y enumera las herramientas metodológicas que utilizaremos para llevar a cabo nuestra investigación. Además, se proporciona un resumen del resto del documento.
2. **Estado general del arte y de la práctica.** Relaciona los enfoques que se han venido utilizando hasta ahora para afrontar los desafíos que estamos abordando, y por qué no han tenido éxito en resolverlos adecuadamente. En particular, establecemos el marco epistémico-conceptual que encuadra nuestra investigación, analizando diferentes concepciones de privacidad y términos relacionados, e introduciendo nuestra definición operativa y nuestro modelo de adversario. Analizamos en este mismo capítulo cómo se abordan la privacidad y la protección de datos desde los diferentes paradigmas dominantes.
3. **Evaluación de necesidades y análisis de deficiencias:** Se presenta en detalle un análisis de las funciones que pueden proporcionar las disciplinas de ingeniería de software y sistemas para respaldar la privacidad y la protección de datos, y que no están cubiertas adecuadamente por otros enfoques. Obtenemos estas necesidades de dos fuentes: un conjunto de entrevistas con ingenieros en ejercicio y de un análisis sistemático y exhaustivo del texto legal del RGPD.
4. **Ingeniería basada en modelos para la privacidad y protección de datos.** Explica el enfoque de la ingeniería de software y sistemas basada en modelos (MBSSE) y muestra cómo la ingeniería de privacidad y protección de datos puede encajar como una perspectiva en esa disciplina más amplia, describiendo un lenguaje de modelado de aspectos específico de dominio (DSAL) para la privacidad y la protección de datos, abstracto (independiente de la tecnología de modelado) y con múltiples puntos de vista.
5. **Minimización de datos y métricas de riesgos de privacidad y protección de datos.** Construye un enfoque multidimensional de los diferentes significados que se han atribuido al concepto de *'minimización'* y otros relacionados, y deriva de ellos una propuesta de métricas para medir los riesgos de privacidad para los interesados desde una organización que lleva a cabo actividades de tratamiento de datos personales.
6. **Ingeniería de requisitos de privacidad y protección de datos.** Introduce un enfoque dual (basado en riesgos y orientado a metas) para gestionar los requisitos de privacidad y protección de datos, un modelo de requisitos derivados de estándares técnicos de privacidad y su correspondencia con los requisitos legales.
7. **Lenguaje de patrones de diseño para la privacidad y protección de datos.** Proporciona un modelo del léxico, la morfología y la sintaxis de un lenguaje de patrones de diseño (soluciones reutilizables y ampliamente probadas) para la privacidad.
8. **Modelos normativos reutilizables para el aseguramiento de la privacidad y protección de datos.** Proporciona varios modelos de elementos de privacidad y protección de datos para fines de aseguramiento o garantía: un modelo orientado al proceso de los aspectos de evaluación de impacto de protección de datos (EIPD o DPIA) del RGPD, un modelo de patrones de argumentación de los mismos, un modelo orientado al proceso de un estándar para EIPD, un modelo orientado al proceso de directrices para EIPD en una industria específica, y correspondencias entre ellos.
9. **Ingeniería de métodos de privacidad y protección de datos.** Modela las actividades de las diferentes disciplinas involucradas en la ingeniería de privacidad y protección de datos, y propone una arquitectura de una herramienta para implementar estos métodos a través de un conjunto de herramientas basadas principalmente en herramientas de modelado convencionales.
10. **Valoración, trabajo futuro y conclusiones.** Presenta el resumen de un documento de orientación de políticas enviado a la CE y una agenda de investigación, evalúa los

resultados e impactos del trabajo de investigación (con un enfoque especial en la adopción de resultados y transferencia tecnológica), esboza el trabajo futuro que puede derivar de los resultados, y proporciona algunos comentarios finales.

Como se ha indicado, este documento ofrece una versión resumida en castellano de memoria la tesis doctoral indicada. Para mayor detalle, se remite a los capítulos correspondientes de la memoria original en lengua inglesa, a donde se hará referencia explícita desde distintas secciones de este mismo documento. Todo el documento, incluidas las tablas, está escrito en castellano, pero las figuras extraídas de la memoria de la tesis se han mantenido en la lengua inglesa original por la dificultad de rehacerlas.

1.2 Motivación y necesidad del trabajo de investigación

Tal como pronosticó Marc Andreessen (fundador de Netscape, e inversor en Twitter y otras grandes empresas) [17] en 2011 en el Wall Street Journal, *el software se ha comido el mundo*: escuchamos música en Spotify o iTunes; buscamos trabajo en LinkedIn, pareja en Tinder y alojamiento turístico en AirBnB; el periodismo ha entregado su modelo de negocio a algoritmos de recomendación inextricables; varias industrias han sido desintermediadas por la '*uberización'* y la '*economía de plataformas*'; y las comunicaciones personales y comerciales ya no se basan en llamadas telefónicas, ni siquiera en mensajes de correo electrónico, sino en servicios proporcionados por aplicaciones como WhatsApp, WeChat, Slack o Zoom. Incluso la cadena de valor de la propia industria de los teléfonos móviles se ha invertido: el éxito de una línea de *smartphones* depende más del software disponible en el correspondiente mercado de aplicaciones que del propio hardware del teléfono.

Bajo este imperio del software han **florecido negocios de éxito mundial**: como muestra, las cinco *mayores corporaciones por capitalización de mercado* en EE. UU. (a fecha de diciembre de 2021) son empresas cuyo negocio está centrado en el software, conocidas colectivamente como *Gigantes Tecnológicos o GAFAM* (por el acrónimo de sus nombres: Google-Alphabet, Apple, Facebook-Meta, Amazon y Microsoft). El florecimiento de estos negocios aprovecha la combinación de una conectividad ubicua, la capacidad de cómputo disponible tanto en los centros de datos distribuidos como en los dispositivos que los usuarios finales llevan en su propio bolsillo, unos procesos y herramientas de desarrollo optimizados, y unas tasas de interés históricamente bajas que han favorecido la inversión en acciones de alto potencial crecimiento y los fondos de cobertura; y, para colmo, se han visto impulsados por un salto nunca visto al mundo en línea causado por el shock de la pandemia de COVID-19.

Pero, por mucho que el software domine nuestro mundo cada vez más, su valor **depende de los datos que procesa**: el software está hecho de estructuras y algoritmos que no sirven de nada hasta que se rellenan y se alimentan con datos. Por eso se dice que *los datos son el nuevo petróleo* [18], la materia prima que alimentará a las nuevas industrias de la Cuarta Revolución Industrial; un '*material'* que, como el petróleo, por supuesto, requiere ser refinado en información, conocimiento y, finalmente, sabiduría; pero, sin embargo, se trata de un '*material'* que no está hecho de materia, sino de bits que pueden duplicarse, renovarse y generarse a voluntad a bajo coste.

Pero *«nunca llueve a gusto de todos»*, o, dicho de una manera más formal, la innovación (en el sentido schumpeteriano) que lleva el progreso a una industria siempre implica efectos destructivos en otros lugares. Siguiendo la analogía de la industria petrolera, igual que esta ha sido fuente de contaminación y derrames de petróleo, también la industria de este 'nuevo petróleo' de los datos de este puede generar externalidades negativas en forma de *'contaminación' o 'vertidos' de datos* [19].

En particular, los usuarios de los servicios basados en datos y los ciudadanos en general se ven a menudo **afectados negativamente por el tratamiento de sus datos personales**. La misma inmaterialidad que hace que los datos sean particularmente aptos para procesarlos a coste cero abre infinitas posibilidades de explotarlos para cualquier propósito sin escrúpulos. De hecho, muchos de los servicios prestados por los susodichos gigantes tecnológicos y por otras muchas empresas más pequeñas se basan

en explotar datos personales: los servicios no se proporcionan ante un pago monetario, sino a cambio de datos personales de los usuarios, que, una vez debidamente acumulados y perfilados, permiten que las empresas de Internet vendan servicios basados en datos (principalmente servicios de publicidad) a sus auténticos clientes —lo se recoge el dicho «*si no eres el cliente, entonces eres el producto*» [20], [21]—. Esto expone la dualidad que veníamos comentando: si bien los usuarios de Internet obtienen servicios personalizados y convenientes a un coste reducido, pierden el control de sus datos personales, cuyo valor se refleja en la cuenta de resultados del proveedor de servicios. Las consecuencias son daños tangibles a los interesados (a los que se refieren los datos, los '*data subjects*') en forma de discriminación, pérdida de reputación, etc., contra los que carecen de ningún poder de negociación.

En su día, el CEO de Sun Microsystems declaró que había que superar la privacidad porque, al fin y al cabo, ya no teníamos ninguna, y el de Facebook que la privacidad ya no era una norma social [22], [23]. Esto se ha venido a resumir en el dicho de que «*la privacidad está muerta*»… pero eso está lejos de la realidad. Día tras día, los principales medios de comunicación informan sobre escándalos de datos personales que abarcan desde fugas de datos e infracciones que rompen los controles de seguridad, a usos inapropiados de datos personales que incluyen la venta no autorizada a intermediarios de datos y terceros que los reprocesan para propósitos dudosos. Esta lluvia continua de noticias sobre violaciones masivas de la privacidad muestra que los problemas de privacidad existen, pero también demuestra la **relevancia social que se le otorga** (como muestran, por ejemplo, las encuestas del Eurobarómetro de la UE), que va de la mano de una indignación pública contra la industria tecnológica que crece con cada nuevo escándalo. Otra prueba anecdótica de que la privacidad está de moda son las fotografías mostradas en la Figura 2, realizadas por el autor a fines de 2019 en el aeropuerto de Bruselas, donde Google (uno de los susodichos Gigantes Tecnológicos) había colocado grandes anuncios luminosos… para dar a conocer las características de privacidad de sus servicios: hasta qué punto los proveedores de servicios habrán entendido el valor de la privacidad para el cliente, que lo abordan en sus propias acciones de marketing.

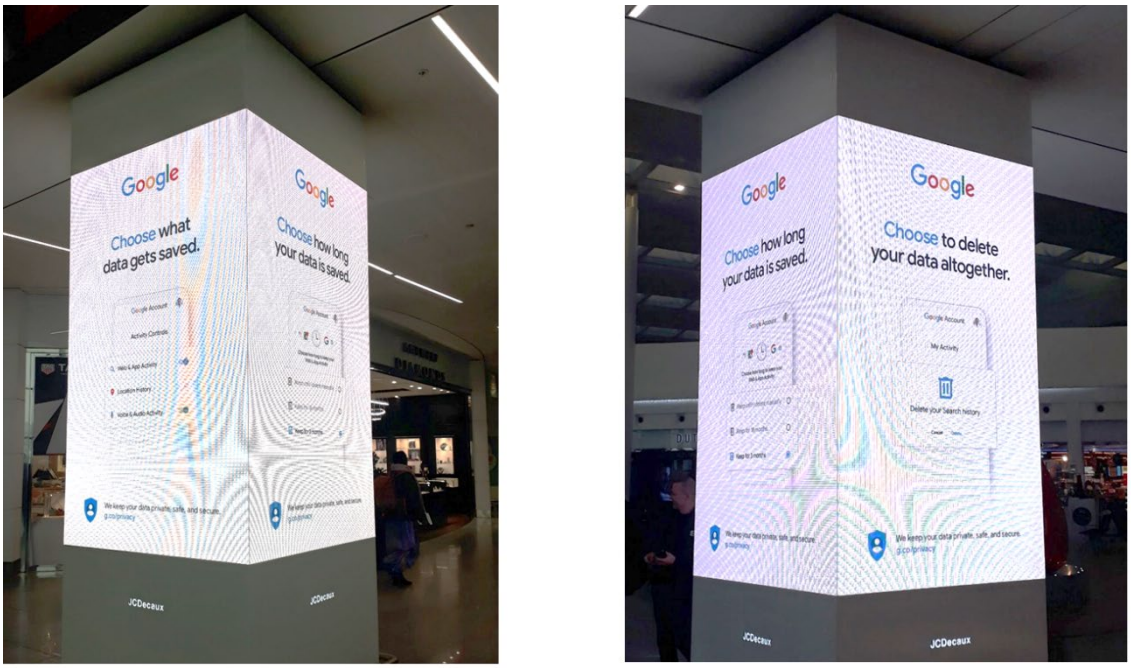


Figura 2. Anuncios luminosos en el aeropuerto de Bruselas, publicitando las características de gestión de datos personales en los servicios de Google

Este aumento en la sensibilidad pública a los problemas de privacidad ha ido de la mano de la promulgación de marcos legales de privacidad más amplios y estrictos en todo el mundo, de los cuales el Reglamento General de Protección de Datos de la UE (RGPD o GDPR) es el ejemplo primordial [24]. El RGPD despertó un interés mundial en cuestiones de privacidad en los años entre su promulgación (2016) y su aplicabilidad (2018), como

muestran las tendencias en las búsquedas web de términos relacionados con la privacidad (Figura 3, datos de Google Trends) e incluso tras de la moda del momento, el interés se ha mantenido más alto que antes.

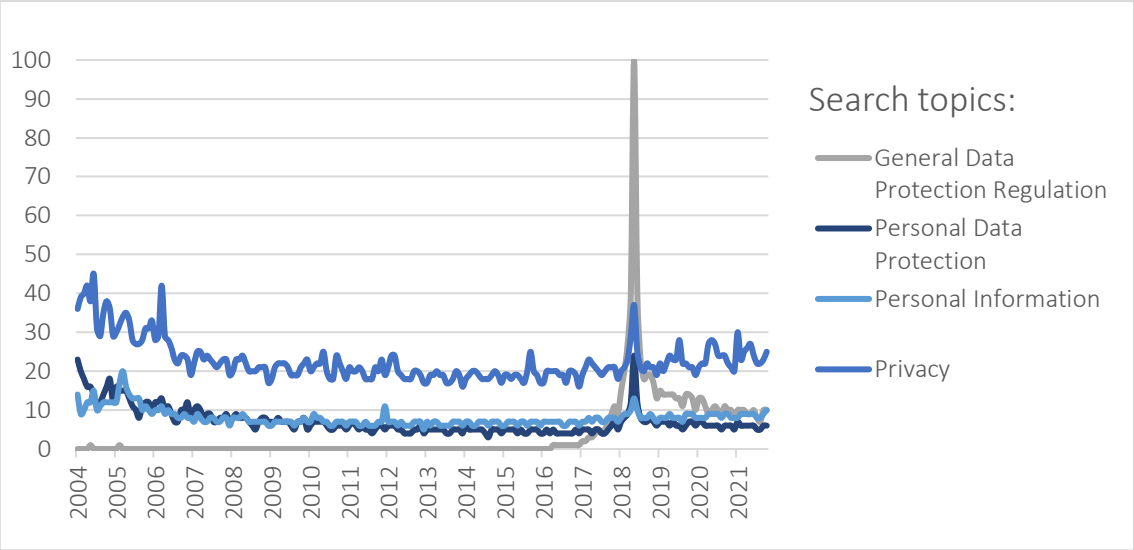


Figura 3. Relevancia a lo largo del tiempo de las consultas realizadas en el buscador de Google en algunos temas relacionados con la privacidad.

Sin embargo, por mucho que los marcos legales puedan influir en la protección de la privacidad y la protección de datos, esta también depende en realidad de **aspectos sociotécnicos** resultantes de las complejas interacciones entre los numerosos actores implicados: proveedores de servicios de Internet dirigidos a particulares (B2C), de servicios entre empresas (B2B), y de plataformas e infraestructuras de comunicaciones y almacenamiento; desarrolladores de todo tipo de componentes de software y hardware necesarios para prestar esos servicios, y gestores de los respectivos proyectos de desarrollo; los propios usuarios interesados a través de sus actitudes hacia la divulgación que guían su uso de la tecnología; legisladores que promulguen el marco jurídico y autoridades de supervisión que lo hagan cumplir, etc.

Si bien nuestra investigación tiene en cuenta estas relaciones complejas entre diferentes actores, se dirige en particular y principalmente a alguien cuyo papel en la protección de la privacidad había sido desdeñado tradicionalmente: el **ingeniero involucrado en el desarrollo** de los productos, servicios y sistemas que acabarán procesando los datos personales una vez que se pongan en funcionamiento. Como se dice a menudo, «*el código [del software] es ley*» [25], en el sentido de que el soporte tecnológico regula nuestras acciones (al ofrecer funcionalidades que respaldan, imponen o excluyen acciones específicas) tanto como lo hace el marco legal (al permitirlas, hacerlas cumplir o prohibirlas); de ahí el aumento de los marcos técnicos, organizativos y legales que enfatizan la necesidad de incorporar medidas específicas dentro del software para proteger la privacidad en la práctica. Por lo tanto, parece sensato que los ingenieros (quienes, en última instancia, están a cargo de desarrollar los componentes de software en los que se basan los servicios) participen en las actividades de protección de la privacidad para *cumplir con su responsabilidad*.

Puede parecer evidente que la privacidad y la protección de datos requieren la cooperación de los ingenieros que implementan los productos, sistemas y servicios que tratarán los datos personales. O, dicho de otra manera: la introducción de características de privacidad y protección de datos afectará al trabajo de los ingenieros, como lo han señalado repetidamente numerosos medios profesionales (Figura 4). De hecho, podemos preguntarnos, si los ingenieros han aplicado sistemáticamente buenas prácticas durante mucho tiempo para tratar con otras categorías de requisitos (por ejemplo, a través de la ingeniería de software seguro), *¿por qué no se puede aplicar el mismo enfoque para tratar la privacidad y la protección de datos?*

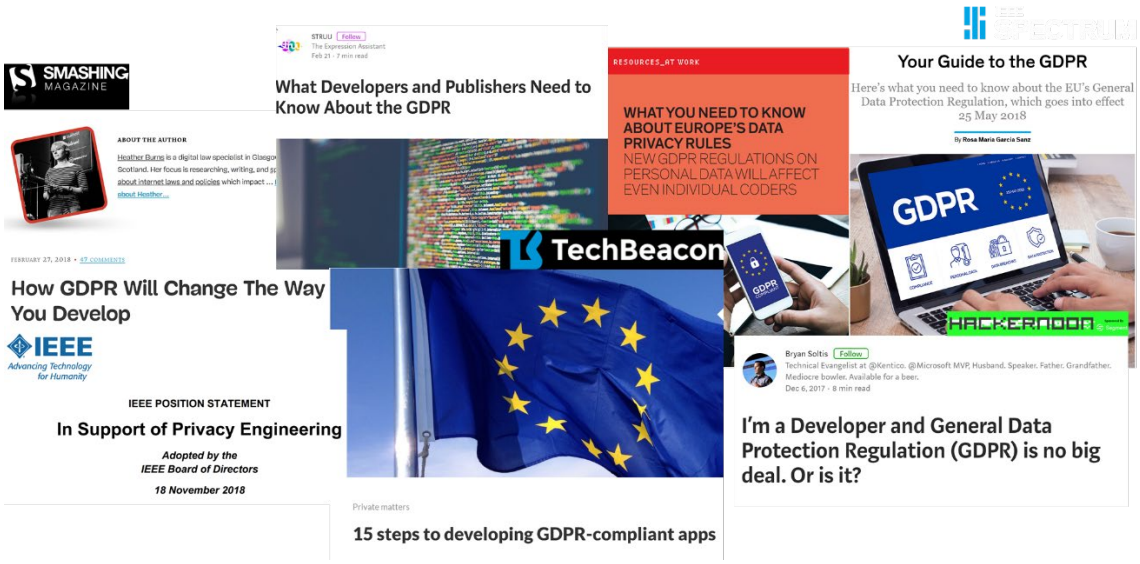


Figura 4. Collage de diferentes noticias de medios profesionales orientados a la ingeniería que destacan el papel del RGPD (y de la privacidad en general) en las actividades de ingeniería de software y sistemas.

El RGPD establece una serie de principios vinculantes de protección de datos, derechos individuales y obligaciones legales para garantizar la protección de los datos personales de los ciudadanos de la UE. Sin embargo, **el enfoque legal no es suficiente** si no viene acompañado de medidas técnicas para proteger la privacidad y los datos personales en la práctica. Para abordar adecuadamente la privacidad y la protección de datos personales, estas deben considerarse de manera proactiva durante el diseño y desarrollo de productos, servicios y sistemas. Esta noción está capturada por los *principios de Privacidad y Protección de Datos por Diseño (PbD / DPbD),* que promueven que la privacidad y la protección de datos se consideren desde el inicio de un proyecto y a lo largo de todas las actividades involucradas en el ciclo de vida de desarrollo del sistema (SDLC), en lugar de abordarlas como una cuestión añadida en el último momento del proceso de desarrollo. Sin embargo, para que estos principios resulten viables, es necesario **involucrar efectivamente a los ingenieros**, ya que son responsables en última instancia de concebir, elaborar, construir y mantener los sistemas, servicios y productos de software y hardware que deben cumplir con el RGPD. Aún así, a pesar del interés despertado por los principios PbD / DPbD en el ámbito regulatorio, aún no han logrado una adopción generalizada y activa en la práctica de la ingeniería. Esto responde a un **desajuste entre la mentalidad legal y la tecnológica**, que, desgraciadamente, ha llevado a descuidar o directamente pasar por alto la privacidad y la protección de datos en la mayoría de los trabajos relevantes sobre ingeniería de datos. Como consecuencia, desde la mentalidad de los ingenieros, la privacidad y la protección de datos generalmente se consideran solo desde la perspectiva de garantizar la seguridad de los datos (si acaso); y tienden a confiar en el cumplimiento de las políticas de privacidad en lugar de en los diseños técnicos que eligen en función de otros requisitos y restricciones (distintos de la privacidad y la protección de datos). Sin embargo, *las innovaciones regulatorias de privacidad y protección de datos sí que impactan en la tecnología y el proceso de*

ingeniería. A modo de ejemplo, el derecho al olvido o el derecho a la portabilidad de los datos, además de permitir su ejercicio frente a los responsables del tratamiento, implican que estos deben implementar las funcionalidades necesarias en sus productos para poder dar respuesta a las solicitudes de ejercicio por parte de sus usuarios. Y esto tiene un impacto real a lo largo del ciclo de desarrollo del producto, ya que implica introducir requisitos operativos para hacer valer esos derechos, modelar las categorías de datos personales afectados, determinar de las funciones y el comportamiento del sistema a petición de los usuarios, e implementar y validar esos comportamientos. Otras innovaciones regulatorias, como la responsabilidad proactiva (o rendición de cuentas, *accountability*) o la evaluación de impacto de la protección de datos, afectan directamente al proceso de desarrollo.

En el ámbito puramente tecnológico, las soluciones se han investigado y elaborado durante mucho tiempo para crear **tecnologías de mejora de la privacidad (PET)** que fomentan la protección de datos y responden a las preocupaciones de privacidad. Sin embargo, *las tecnologías de mejora de la privacidad siguen siendo desconocidas para la mayoría de los ingenieros*, debido al desacoplamiento entre los PET y la práctica de la ingeniería y el desarrollo sistemáticos, lo que hace que los ingenieros desconozcan o no conozcan la aplicabilidad adecuada de tales soluciones.

Por tanto, los ingenieros no suelen estar en el foco de los métodos y herramientas de protección de la privacidad (que, en general, se dirigen al personal legal, de administración y de operaciones de TI), e incluso los principios de PbD / DPbD carecen de guías prácticas y operativas que los ingenieros puedan seguir fácilmente. Pero las disciplinas de la ingeniería de software y sistemas tienen una larga tradición exitosa de proporcionar orientación sistemática para abordar problemas complejos en el desarrollo de software y sistemas. Por ello, nos planteamos este trabajo de investigación para proporcionar nuestras modestas contribuciones hacia una mayor consideración de la privacidad y la protección de datos dentro del ciclo de vida del desarrollo de software y sistemas, ya que postulamos que la sabiduría acumulada durante los años por las comunidades de práctica de software e ingeniería de sistemas en sus respectivos cuerpos de conocimiento puede aprovecharse y expandirse para responder a las demandas de servicios más respetuosos con la privacidad.

La investigación académica ha demostrado consistentemente [26]–[29] que los desarrolladores e ingenieros (que generalmente no son expertos en privacidad en absoluto), encuentran que la privacidad y la protección de datos son ajenas a su trabajo y, lo que es más importante, *rara vez usan herramientas de gestión de la privacidad, ya que encuentran que están más orientadas al ámbito legal* que a las actividades de ingeniería. La misma investigación ha encontrado que serán más similares a tomar decisiones que protejan la privacidad y la protección de datos cuando *el proceso esté integrado en su flujo de trabajo y herramientas de desarrollo* habituales.

Por su parte, las empresas expresan una *demanda insatisfecha por parte de sus clientes de contar con herramientas y métodos* que faciliten la introducción de la privacidad y la protección de datos —y el cumplimiento del RGPD en particular— sin interrumpir sus flujos de trabajo de desarrollo ni introducir mucha carga sobre ellos.

En síntesis, la regulación pide a los ingenieros que se adhieran a los principios de PbD y apliquen soluciones de protección de datos a lo largo de sus proyectos, y los desarrollos tecnológicos proporcionan tales soluciones en forma de PET. Sin embargo, para lograr eso, los ingenieros exigen **métodos y herramientas más cercanas a su mentalidad** y les permitan introducir sistemáticamente tales soluciones dentro del flujo de trabajo y el entorno al que están acostumbrados. Para dar respuesta a los retos que acabamos de introducir, argumentamos que *los ingenieros deben estar dotados de herramientas metodológicas y tecnológicas* que permitan la aplicación sistemática de los principios de protección de datos para lograr el cumplimiento del marco regulatorio, manteniendo al mismo tiempo una economía de recursos, conviviendo con otros requisitos y logrando los objetivos generales del negocio.

Esta visión se realiza en el campo de la **Ingeniería de Privacidad y Protección de Datos**, una definición seminal de la cual fue proporcionada por Gürses y Del Álamo como «*un campo naciente de investigación y práctica que persigue enfoques sistemáticos para el inicio y la aplicación de soluciones orientadas a la privacidad a través de sistemas y procesos de desarrollo de software*» [30]. Las piedras angulares del campo de la Ingeniería de Privacidad y Protección de Datos son precisamente los *métodos* empleados por los ingenieros para capturar y abordar los problemas de privacidad durante las diferentes etapas del desarrollo de los sistemas sociotécnicos, y las *herramientas* de software que apoyan a los ingenieros durante el proceso para realizar las tareas o actividades de ingeniería prescritas por dichos métodos. Este campo está integrado dentro de las disciplinas más amplias de la *ingeniería de software y sistemas*, que tienen comunidades maduras de práctica que han generado cuerpos de conocimiento bien establecidos y ampliamente aceptados aplicados por los ingenieros en su trabajo diario. Por lo tanto, la adopción y aceptación de métodos y herramientas de ingeniería de privacidad y protección de datos requiere su *integración* dentro del gran legado de los métodos y herramientas de ingeniería de software y sistemas.

Nuestro trabajo surge como respuesta a tal **demanda de Ingeniería de Privacidad y Protección de Datos**, perfectamente capturada por estas recomendaciones emitidas por ENISA: «*La comunidad de investigación necesita investigar más a fondo en ingeniería de privacidad, especialmente con un enfoque multidisciplinario. Este proceso debe ser apoyado por agencias de financiación de la investigación*» [31, p. 51], «*Los proveedores de herramientas de desarrollo de software y la comunidad de investigación deben ofrecer herramientas que permitan la implementación intuitiva de las propiedades de privacidad.*» [31, p. 52]

Postulamos que *los ingenieros no necesitan ser expertos en privacidad, sin embargo, enfrentarán problemas de privacidad en su trabajo de todos modos*. En consecuencia, nuestro trabajo tiene como objetivo facilitarles el tratamiento de esos problemas proporcionándoles *herramientas y métodos que no los obliguen a salir de su zona de confort*. El enfoque para hacerlo consiste en agregar funciones de privacidad y protección de datos a software de propósito general y herramientas de ingeniería de sistemas que ya están en uso por los ingenieros; e integrar las actividades de ingeniería relacionadas con la privacidad y la protección de datos dentro de los métodos que ya siguen. Los ingenieros tienen una sólida formación en elicitación de requisitos, análisis de software, diseño arquitectónico y detallado, modelado de datos, validación de software, garantía de calidad, gestión de proyectos, etc. y están acostumbrados a aplicar herramientas de software que les ayudan a realizar de manera sistemática y económica tales tareas. Nuestro trabajo tiene como objetivo aprovechar toda esa inteligencia en aras de la privacidad y la protección de datos.

1.3 Objetivos

En el contexto que acabamos de presentar, abogamos por *incluir las funciones de privacidad y protección de datos en las herramientas que los ingenieros usan habitualmente* (una necesidad que ha sido recogida por ENISA), en lugar de obligarles a aprender a usar otras herramientas desconocidas para ellos. Asimismo, defendemos la integración de las actividades de privacidad y protección de datos en las diferentes etapas del *Ciclo de Vida del Desarrollo de Sistemas* (SDLC – System Development Lifecycle), y en los métodos y flujos de trabajo que siguen los ingenieros, en lugar de considerarlas en actividades separadas e inconexas. Además, para que se puedan aplicar de manera cotidiana, apoyamos la creación de un *cuerpo de conocimiento* fácilmente disponible, descrito en términos compatibles con la mentalidad de los ingenieros, y que destile el conocimiento acumulado por la comunidad de práctica e investigación en privacidad. En consecuencia, consideramos que los conocimientos técnicos en materia de privacidad y protección de datos deben incorporarse a la práctica general de la ingeniería de software y sistemas, proporcionando a los ingenieros métodos y herramientas que estén más cerca de su experiencia.

Por lo tanto, en este trabajo, asumimos la misión de llevar el conocimiento establecido de privacidad y protección de datos a la práctica general de la ingeniería de software y sistemas, proporcionando a los ingenieros métodos y herramientas que operativicen los principios y la regulación de privacidad y protección de datos, y que se integren con los que ya utilizan habitualmente en las diferentes actividades que tienen lugar a lo largo de las etapas del ciclo de vida de desarrollo de sistemas (SDLC), llevando a cabo el paradigma de la privacidad por diseño; para que, en última instancia, puedan crear sistemas que cumplan con el RGPD, se adhieran a los principios de protección de datos y cuiden los derechos de los interesados. (Figura 5)

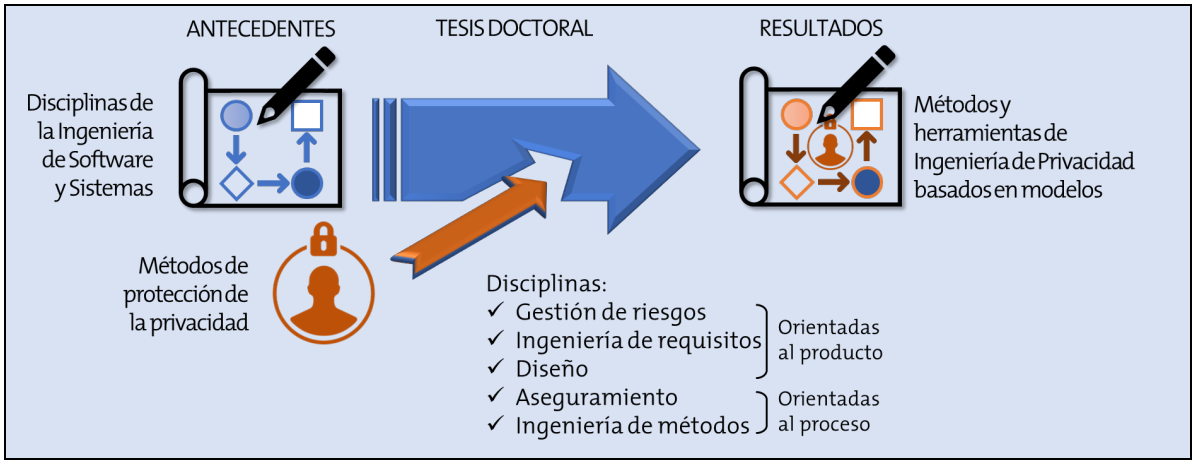


Figura 5. Representación gráfica de nuestra misión.

Esta misión global se instrumenta a través de los **siguientes seis objetivos específicos**:

1. *Identificar las deficiencias respecto a las necesidades de los ingenieros con respecto a las cuestiones de privacidad y protección de datos* (según lo establecido por las regulaciones legales u otras fuentes) *que pueden abordarse desde las diferentes disciplinas de* la *ingeniería de software y sistemas* (a saber, gestión de riesgos, ingeniería de requisitos, diseño de sistemas basado en modelos, aseguramiento de sistemas, e ingeniería de métodos)
2. *Integrar la privacidad y protección de datos desde el diseño diseño y las actividades de protección de datos dentro de los métodos de ingeniería de sistemas y software convencionales existentes,* alineándose con las actividades principales de los diferentes enfoques del ciclo de vida de desarrollo de sistemas (SDLC). En particular, tenemos la intención de aprovechar las posibilidades que proporciona la Ingeniería de Sistemas y Software Basada en Modelos (MBSSE) para introducir aspectos transversales con un impacto mínimo para el desarrollo del sistema, enriqueciendo los modelos de sistemas con atributos que relacionan y vinculan el modelo con preocupaciones de privacidad y protección de datos.
3. *Operativizar las cuestiones de privacidad y protección de datos en elementos orientados a la ingeniería (por ejemplo, requisitos, propiedades, métricas).* Los conceptos de privacidad y protección de datos a menudo parecen demasiado abstractos para los ingenieros y desconectados de su profesión, quienes necesitan que esos conceptos se pongan en práctica, es decir, se conviertan en características que puedan observarse, entenderse y medirse más fácilmente, y que puedan mapearse y rastrearse hasta esos conceptos abstractos para evaluar su cumplimiento. Los ingenieros deben tratar con conceptos con los que están acostumbrados a trabajar en sus respectivas disciplinas (por ejemplo, requisitos que pueden introducirse como elementos de backlog, índices de riesgo, estructuras de datos y otros modelos de diseño, casos de aseguramiento), y que de hecho estén alineados con los que aparecen en las herramientas software que los propios ingenieros emplean en su trabajo.
4. *Empoderar a los ingenieros generalistas para aprovechar los conocimientos existentes sobre protección de datos, incluso si no son expertos en el campo.* Dado que los ingenieros generalistas no son expertos en privacidad o protección

de datos, no podrán aplicar adecuadamente los métodos de privacidad y protección de datos en su trabajo diario a menos que tengan un cuerpo de conocimiento fácilmente disponible a mano que destile la sabiduría acumulada por la comunidad de práctica e investigación de privacidad y protección de datos, y que esté representada en términos compatibles con su mentalidad.

5. *Demostrar la validez de nuestro enfoque y promover la práctica más amplia y la adopción de la ingeniería de privacidad y protección de datos en general y de nuestras contribuciones en particular.* Con este fin, además de difundir nuestros resultados a través de artículos de revistas académicas y documentos de conferencias, interactuaremos también con las principales comunidades existentes de ingeniería de software y sistemas.
6. *Dirigir recomendaciones a las instituciones para la evolución de políticas y regulaciones legales que puedan aumentar la adopción de la privacidad y la protección de datos a lo largo del ciclo de vida de desarrollo de productos y servicios.* Contrariamente a los proveedores de servicios, los desarrolladores de software y componentes del sistema que fabrican estos sistemas no son el objetivo directo del marco regulatorio actual, ni esto se adapta a sus necesidades. Consideramos que la regulación debe abordar toda la cadena de suministro de desarrollo de servicios, y especialmente dirigirse a las organizaciones que desarrollan sistemas.

En particular, abordaremos la introducción de la privacidad y la protección de datos en varias disciplinas **de ingeniería** de software y en las que donde se puede lograr el mayor impacto, debido a la relevancia que tienen en el cumplimiento del marco regulatorio (cf. Tabla 5 infra), la madurez de los métodos de ingeniería de privacidad y protección de datos en esas áreas, y la existencia de métodos y herramientas ampliamente utilizados para la ingeniería de software. Tales disciplinas son, a saber:

1. **Gestión de riesgos, que apoya la ejecución de** *evaluaciones de impacto de privacidad y protección* de datos desde la perspectiva de la ingeniería, alineadas con los requisitos del art. 35 del RGPD y la guía WP29, para identificar, evaluar, evaluar y mitigar los riesgos para los interesados que puedan surgir de las actividades de tratamiento que tratan con sus datos personales. [32]
2. **Ingeniería de requisitos,** que respalda la *operativización de los objetivos de alto nivel establecidos por el RGPD* (por ejemplo, principios de privacidad, derechos de los interesados, obligaciones de responsables y encargados) en requisitos de diseño específicos (controles de privacidad), la elección adecuada de soluciones (por ejemplo, PET y patrones de privacidad) vinculados a esos controles, y la especificación sistemática general, gestión, análisis, trazabilidad, validación y verificación de la privacidad y Requisitos de protección de datos.
3. **Diseño de sistemas basado en modelos,** que ayuda a los ingenieros a analizar y diseñar los sistemas *en desarrollo desde la perspectiva* de la privacidad y la protección de datos, enriqueciendo los modelos de sistemas con la introducción de atributos relevantes de protección de datos y privacidad (incluidos, por ejemplo, modelos de información personal y de actividades de tratamiento de datos); analizando la alineación de dichos modelos con los principios**,** regulaciones y estrategias de privacidad; y transformando los modelos para que se adhieran a Principios de protección de datos.
4. **Aseguramiento,** que apoya la demostración del cumplimiento del RGPD y la *observancia del principio de responsabilidad* a través de la captura sistemática de evidencias, su asociación con requisitos y artefactos, trazabilidad al RGPD y argumentación de cumplimiento derivada de esas evidencias.
5. **Ingeniería de** *métodos*, que apoya la *organización del trabajo realizado en las otras disciplinas* y *su integración entre sí y con* marcos de desarrollo de procesos de *propósito general*, a través de una definición estructurada de los procesos que

establecen, los artefactos que consumen y producen, y los roles involucrados en ellos, fomentando la reutilización y la interoperabilidad de métodos definidos independientemente del proyecto.

Los objetivos de investigación planteados **contribuyen a cumplir distintos Objetivos de Desarrollo Sostenible (ODS)**, en particular, se encuentran alineados con las Metas 8.2., 8.3, 9.c, 16.3 y 16.6.

1.4 Métodos aplicados

Debido a la amplitud del alcance de nuestras contribuciones, se han utilizado **diferentes métodos** para llevar a cabo la investigación asociada a cada una de ellas.

- El estado del arte y de la práctica se han construido a través de un *análisis de los paradigmas dominantes* en la ingeniería de privacidad y protección de datos. Empleamos varios métodos específicos para analizar el trabajo relacionado en algunas de las disciplinas que tratamos: *revisión* general (análisis terciario) para analizar las revisiones existentes sobre métricas de privacidad y definiciones de entropía, *visión general crítica* (revisión integradora) para analizar el estado del arte y MBSSE en privacidad. Para examinar el estado del arte en patrones de privacidad (Sec. 7.1), hemos recurrido a la realización de un *estudio de mapeo sistemático* de acuerdo con las directrices propuestas por Petersen et al. [33], [34].
- Las contribuciones generales se han desarrollado siguiendo las pautas metodológicas de *Design Science* ('*ciencia del diseño*'). Se trata de un paradigma de investigación constructivo que tiene como objetivo desarrollar conocimiento prescriptivo, es decir, se centra en mejorar algún aspecto de los problemas existentes en contextos específicos, a través del diseño iterativo de artefactos portadores de conocimiento que resuelven los problemas encontrados, en lugar de explicar los fenómenos existentes. Igual que en nuestra investigación, Design Science suele dirigirse a profesionales de un campo determinado, y sus artefactos pueden incluir modelos de procesos y metodologías. Planteamos seguir las directrices que debe seguir el proceso de investigación de Design Science según Hevner [35]: diseño como artefacto, relevancia del problema, evaluación del diseño, contribuciones de investigación, rigor en la investigación, diseño como proceso de búsqueda, y presentación de la investigación.
- Por otra parte, las contribuciones específicas presentadas en cada capítulo emplean sus propios métodos, detallados en las respectivas secciones.
    - Para definir operativamente el objeto de nuestra investigación, se utiliza el marco de los *conceptos esencialmente cuestionados* y un *modelo de adversario* (v. Sec. 2.2).
    - Para determinar las necesidades de los ingenieros, se lleva a cabo un análisis de brechas mediante la aplicación del *método de los Modelos Mentales* (v. sec. 3.1), en combinación con un análisis textual del RGPD.
    - Para definir un lenguaje específico de aspectos de dominio (DSAL) para la privacidad y protección de datos, se sigue el enfoque de *ingeniería de software y sistemas basada en modelos MBSSE* (v. Sec. 4.1)
    - Se aplica el *análisis de contenido cualitativo (QCA)* (presentado en la sección 5.1) para determinar las dimensiones que se consideran bajo el concepto de '*minimización de datos*'. Este mismo método se utiliza más adelante para el modelado de varios estándares y partes del RGPD (Secs. 6.3 y 8.2).
    - Para la derivación de requisitos de privacidad y protección de datos, se propone seguir un método de *operativización jerárquica* a partir de requisitos estándar (v. Sec. 6.1)
    - Se han seguido los métodos propuestos por Meszaros & Doble, y por Noble (v. Sec. 7.1) para la definición de *lenguajes de patrones y sus relaciones*, en nuestro caso aplicado al lenguaje de patrones de privacidad.
    - Para la definición del metamodelo de métodos de ingeniería de privacidad, se ha seguido el enfoque de *construcción de métodos propuesto por la ingeniería*

situacional de métodos según SEMDM (v. Sec.9.1), y el modelado funcional de procesos *SIPOC propuesto por la metodología Six Sigma* (v. Sec.9.1).

Además, como nuestras contribuciones abordan disciplinas metodológicas, aprovecharemos los métodos de cada *disciplina* para crear nuestras contribuciones, a saber, de: métodos de gestión de riesgo (por ejemplo, CORAS), aseguramiento de sistemas (v. Sec. 8.2), etc.

2 Estado general del arte y de la práctica

Antes de acometer propiamente nuestros objetivos de investigación, ofreceremos una visión global del entorno en el que se desarrollarán, desde dos perspectivas:

1) El *contexto epistemológico*, analizando el concepto de privacidad que enmarca y da soporte a nuestro trabajo.
2) El *contexto paradigmático* que domina las soluciones actuales para la privacidad, y cuyas lagunas vienen a inspirar nuestras contribuciones.

Comenzaremos sumergiéndonos en la búsqueda de una definición operativa del concepto de 'privacidad', a partir de algunos esfuerzos previos existentes para especificarlo. Así, repasaremos distintas *definiciones seminales* del concepto de privacidad: desde el derecho a permanecer tranquilo frente a intromisiones (Warren y Brandeis, 1890), pasando por la autodeterminación (Westin, 1967), y el control selectivo de la información personal (Altman, 1975), hasta llegar a la integridad contextual (Nissenbaum, 2004). Se analizan asimismo diferentes *descomposiciones del concepto de privacidad* más cercanas a la ingeniería, a través de: el objeto de la privacidad, los principios de privacidad, sus características operativas, funciones de gestión organizativa, daños y riesgos de los que protege, y las propiedades y metas de la ingeniería de privacidad y protección d datos. Para nuestro trabajo, resultan especialmente relevantes los principios de la norma ISO/IEC 29100, los principios del mencionado Reglamento General de Protección de Datos de la Unión Europea, junto con los derechos y obligaciones que establece, y las amenazas definidas en la metodología LINDDUN. La Figura 6 resume la historia de algunas de estas conceptualizaciones de la privacidad.

La coexistencia de tantas conceptualizaciones distintas de la idea de 'privacidad' nos plantea un desafío, ya que corremos el riesgo de condenar nuestro esfuerzo por apoyar a la ingeniería a la privacidad, debido a la propia indefinición de su objeto. Responderemos a ese desafío desde el prisma de los *conceptos esencialmente controvertidos*, un enfoque utilizado en las ciencias sociales y las humanidades para definir conceptos cualitativos y abstractos, para cuya definición es difícil alcanzar un consenso. Así, presentaremos las respuestas a una serie de cuestiones que nos ayuden a definir el concepto de privacidad dentro del ámbito de nuestra investigación. Asimismo, definimos nuestro *modelo de adversario* que enmarca los riesgos para los derechos de los interesados cuya protección queremos lograr, que resulta de la interacción entre tres roles con distintas capacidades e intereses: un ingeniero '*honesto pero despreocupado*', el custodio de los datos personales (que engloba a encargados y responsables del tratamiento), y unos intermediarios de datos sin escrúpulos que tratan de subvertir la explotación de los datos personales.

Para concluir este capítulo, acometemos una crítica reflexiva sobre los *paradigmas dominantes* en la práctica de la privacidad y la protección de datos y su relación con la ingeniería. Encontramos cómo estos paradigmas, centrados en la regulación o en la gestión de los servicios en operación, suelen olvidar el rol del ingeniero responsable de concebir y desarrollar el software que da soporte a esos servicios. Incluso aquellos enfoques más orientados a la tecnología, como la privacidad y protección de datos desde el diseño, o las tecnologías de mejora de la privacidad (PETs), pasan por alto la necesidad de ofrecerles a los ingenieros soluciones que sean operativas y sistemáticas, para que puedan aplicarlas en proyectos concretos sin necesidad de conocimiento experto.

Por cuestiones de extensión, se remite al lector a la memoria de la tesis doctoral del autor [1, Secs 2.2.3, 2.2.4, 2.2.5]para obtener información adicional sobre aspectos como las regulaciones de privacidad en dominios específicos (p.ej. la red eléctrica inteligente o *smart grid*, la tecnología financiera o *fintech*, y el vehículo conectado), los proyectos de investigación e innovación financiados por los programas de la UE, o el panorama de la normalización técnica y estandarización, incluyendo distintos estándares de ISO e IEC

relacionados con la privacidad, ya sea enfocados a disciplinas del ciclo de vida del desarrollo (SDLC), a tecnologías específicas (tarjetas inteligentes o *eCards*, Internet de las Cosas o *IoT*, cadenas de bloques o *blockchain*, sistemas de intercambio electrónico de datos o *EDI*…), a industrias concretas (transporte, educación, financiero, ciudades inteligentes o '*smart cities*', salud…), así como otros estándares de facto de organismos industriales (como OASIS, W3C o IEEE).

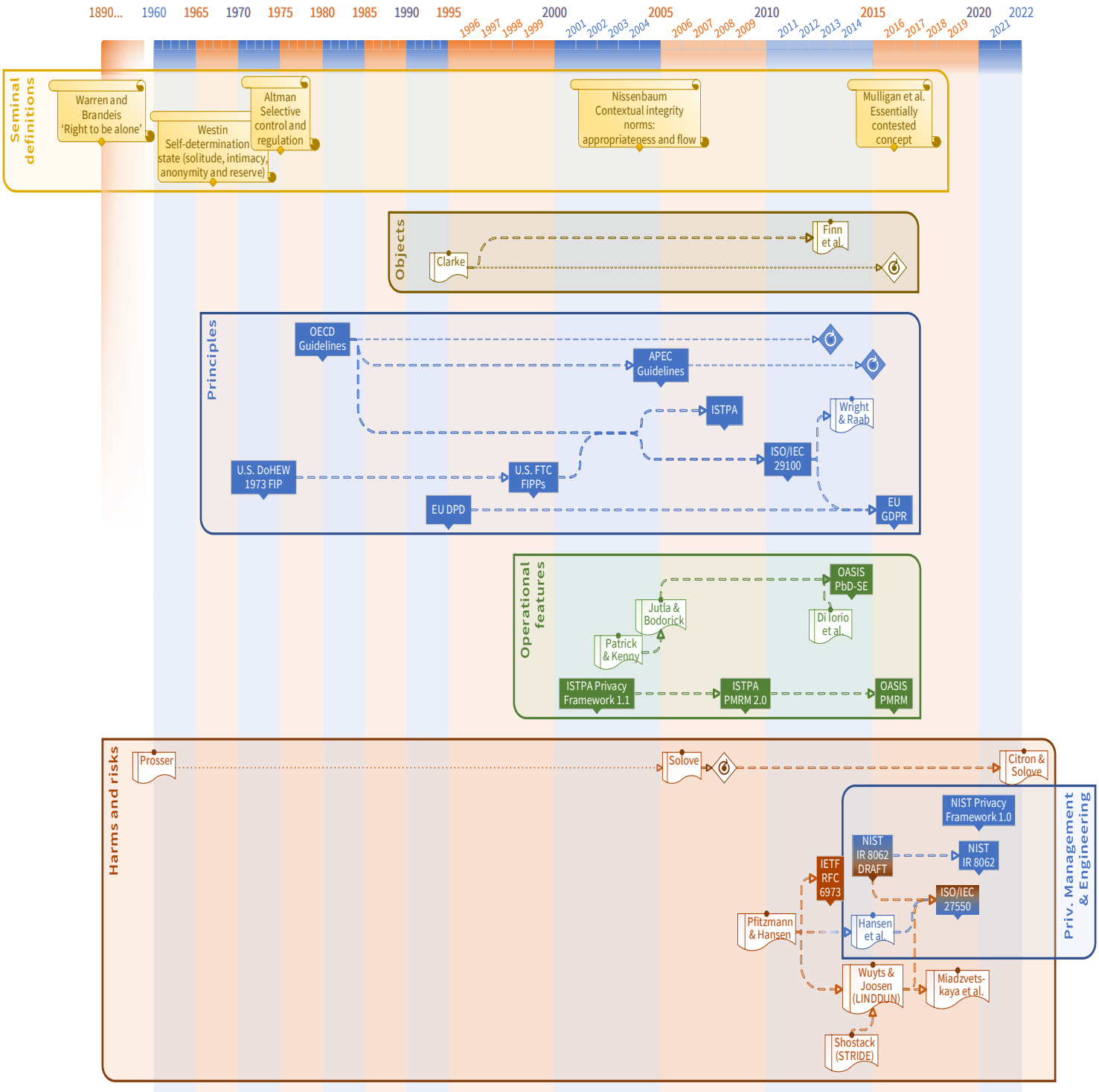


Figura 6. Línea temporal de las distintas conceptualizaciones de la privacidad.

2.1 Resumen de conceptos

El concepto moderno de privacidad resulta reciente en la historia de la humanidad —no obstante, es más antiguo de lo que podría parecer en un principio— y su evolución ha ido paralela a la introducción de nuevas tecnologías. Así, a finales del s. XIX, el desarrollo del obturador mecánico y la película fotográfica facilitaron la aparición de cámaras 'instantáneas' con las que se podían tomar fotografías subrepticiamente sin conocimiento del sujeto fotografiado. Esto impulsó la primera conceptualización de la privacidad, llevada a cabo por **Warren y Brandeis** en 1890 [36]. Para ellos, el derecho a la privacidad se equiparaba con una especie de *derecho a la intimidad*[1] que, por ejemplo, protegía a los ciudadanos de los chismes y las intromisiones de periodistas y fotoperiodistas. Este derecho se distingue tanto del derecho a la protección de la reputación y el buen nombre (un daño a la privacidad no requiere que la información divulgada sea dañina o falsa), como del derecho de propiedad intelectual (la información personal queda protegida independientemente de que exista una obra intelectual o de quién sea su autor).

De manera análoga, siete décadas después, los avances en las tecnologías de la información y la comunicación motivaron a **Westin** a repensar el concepto de privacidad. Introduce la privacidad como «*la reivindicación por individuos, grupos o instituciones de determinar por sí mismos cuándo, cómo y en qué medida se comunica información sobre ellos a otros. Visto en términos de la relación del individuo con la participación social, la privacidad es el retiro voluntario y temporal de una persona de la sociedad en general a través de medios físicos o psicológicos, ya sea en un estado de soledad* [libre de observación] o de *intimidad de un grupo pequeño* [aislamiento necesario para una asociación cercana] *o, cuando se encuentra entre grupos más grandes, en una condición de anonimato* [desconocido y libre de identificación] *o reserva* [limitación de la divulgación a otros].» [37, p. 7]. Según él, la privacidad desempeña diferentes funciones en las sociedades modernas: *autonomía personal*, *liberación emocional*, *autoevaluación* y *comunicación limitada y protegida* [37, pp. 32–39]. Estos cuatro estados o condiciones de privacidad con los que Westin complementó su definición (*soledad, intimidad, anonimato y reserva*) aparecerán de una u otra manera en las diferentes conceptualizaciones o privacidad que presentamos en el resto de este capítulo.

En la misma época, **Altman** comenzó a teorizar sobre la privacidad desde el punto de vista de la psicología social, pero su trabajo, y especialmente la extensión por Petroni a través de la teoría de la *Gestión de la Privacidad de la Comunicación* (CPM - Communication Privacy Management), se llegó a convertir en fundamental para comprender y administrar la privacidad en los servicios modernos de redes sociales en línea. Para Altman, la privacidad puede definirse como «*el control selectivo del acceso al yo*» [38, p. 24]. Existen distintos mecanismos de conducta para ejercer este control que regulan los límites interpersonales para lograr un nivel óptimo de privacidad [39], dentro de un proceso social, bidireccional y dialéctico. Una consecuencia relevante de esta contribución es que la privacidad no solo puede regularse a nivel individual, sino también a nivel grupal (podemos hablar de la privacidad de una familia, un equipo, etc.)

Mucho más recientemente, ante la explosión de los servicios basados en Internet y la electrónica de consumo, **Nissenbaum** consideró que era necesaria una nueva conceptualización de la privacidad, motivada en este caso por la aparición de mecanismos electrónicos de vigilancia pública. Mientras que los problemas de privacidad anteriores involucraban a «*agentes gubernamentales que buscaban ampliar su acceso a los* [datos de] *ciudadanos; u* [organizaciones que] *recopilan o divulgan información sensible, confidencial o personal; o la intromisión en espacios o esferas normalmente considerados privados o personales*»; la vigilancia pública no caía en ninguna de estas tres situaciones, pero planteaba un problema de privacidad. Como alternativa, Nissenbaum propone [40] que las violaciones de privacidad suceden cuando se rompen las normas sociales dependientes del contexto que regulan los flujos de datos

[1] "*Right to be alone*" en el original que se puede traducir por derecho a "estar solo", pero también por "estar tranquilo".

(«*integridad contextual*»). Según ella, todas las transacciones en cualquier esfera de la vida tienen lugar en un contexto (por ejemplo, entre amigos, comprando en una tienda, en clase, visitando al médico, etc.) donde existen normas sociales que regulan los flujos de información. Estas normas pueden abordar la *conveniencia* de divulgar cierta información personal o el *flujo o distribución* de información a otras partes, de distinta manera en función del contexto. Por lo tanto, la privacidad como integridad contextual ya no es una cuestión dicotómica, sino que depende de un «*conjunto más rico y completo de parámetros relevantes*» [41, Ch. 7]: sujeto de datos, remitente, destinatario, tipo de información y tipo de transmisión. Lo que es apropiado en un contexto no lo será en otro.

Todas esas definiciones han demostrado su utilidad en distintos campos; sin embargo, si nuestro objetivo es introducir la privacidad y la protección de datos en los procesos de ingeniería, necesitamos una definición más operativa que podamos aplicar directamente en las actividades correspondientes, a través de, por ejemplo, requisitos específicos, amenazas o modelos de diseño. Por eso, analizamos a continuación los significados atribuidos al concepto de la privacidad desde perspectivas más cercanas al **dominio de la ingeniería**:

- Según el *objeto de la privacidad*, Clarke definió seis tipos de privacidad en distintas obras [42], [43]: privacidad de la persona o corporal, del comportamiento (incluyendo los medios de comunicación de masas), de las comunicaciones personales (incluyendo su intercepción), de los datos personales, de la experiencia personal (incluyendo lo que hoy llamaríamos 'metadatos'), y de la información. Más tarde, Later, Finn, Wright, y Friedewald [44] definieron la privacidad de la acción como extensión de la del comportamiento, la de la imagen como extensión de la de los datos, y añadieron categorías de privacidad del pensamiento y los sentimientos, localización y espacio, y privacidad de asociación (incluyendo la privacidad grupal).
- Desde un enfoque más operativo, la Alianza Internacional de Privacidad, Seguridad y Confianza (ISTPA) [45], [46], y, más tarde, OASIS [47, Sec. 4], propusieron un marco de referencia con diez *servicios necesarios para la gestión de la privacidad*: negociación o acuerdo, control o uso, validación, certificación, auditoría y seguridad, ejecución, interacción, uso, agente, y acceso.
- Podemos hallar otras definiciones relevantes de la privacidad orientadas al mundo de la ingeniería en las *funciones de gestión de la privacidad* del marco del NIST [48] : Identificar, Gobernar, Controlar y Comunicar,
- Desde el punto de vista de los *factores humanos*, Patrick y Kenny [49] se centraron en las características que debe mostrar la interfaz persona-ordenador (IPO o HCI) de un sistema que respete la privacidad: consciencia, comprensión, control, y consentimiento. Más tarde, Jutla y Bodorik [50] expandieron esta categorización para abordar el control del usuario, resultando en estos siete factores: comprensión, consciencia, elección, consentimiento, contexto, confinamiento (de los datos), y consistencia.
- Di Ioro *et al.* definieron en el proyecto BIRO tres *propiedades* que debe implementar cualquier sistema que cumpla con los principios de privacidad [51]: minimizar la identificabilidad, la enlazabilidad, y la observabilidad. El Grupo de Trabajo sobre Documentación de la Privacidad desde el Diseño para Ingenieros de Software de OASIS (PbD-SE), codirigido por Jutla y Cavoukian, recogió esas propiedades y añadió otras cuatro [52], [53]: trazabilidad, auditabilidad, rendición proactiva y responsable de cuentas (*accountability*), y disposición para la verificación.
- Desde un punto de vista negativo, hay varios enfoques que definen la privacidad a partir de un conjunto posibles violaciones de ella, daños derivados de estas, o riesgos de que se produzcan.
    - Prosser en 1960 hablaba [54] de *daños y perjuicios* por intrusión, revelación, publicidad y apropiación.
    - Siguiendo este camino, Solove [55], [56, Ch. 5] define una taxonomía de *actividades que pueden dañar la privacidad*, y que abarcan la recogida de información (daños por vigilancia, por interrogación), el procesamiento de la información (daños por agregación, identificación, inseguridad, uso

secundario, y exclusión), la diseminación de información (por ruptura de la confidencialidad, revelación, exposición, accesibilidad aumentada, chantaje, apropiación, distorsión, y diseminación), o la invasión en asuntos privados (por intrusión, o interferencia en las decisiones).
    - En otra clasificación del mismo autor [57], los *daños a la privacidad* se categorizan como físicos, reputacionales, económicos, psicológicos, a la autonomía, o de discriminación.
    - Finalmente, desde un enfoque más cercano a la ingeniería, los trabajos de Pfitzmann y Hanssen [58], y más tarde Wuyts y Joosen [59], definen un conjunto de 7 *categorías de amenazas* (que responden a las siglas LINDDUN en inglés): enlazabilidad, identficabilidad, no repudio, detectabilidad, revelación de información, desconocimiento, y no cumplimiento.

Para especificar el contenido que abarca el concepto de 'privacidad', se ha recurrido a menudo a establecer listas de **principios de privacidad** a los que atenerse para cumplir con ella. Aunque las fuentes que los especifican no proporcionan una definición explícita acerca de lo que es un 'principio', podemos resumirlo como aquellos valores o intereses fundamentales en los que se puede descomponer el concepto de la privacidad. A lo largo de los años, diversos órganos (gobiernos, parlamentos, comités asesores, organismos y asociaciones internacionales, organismos de estandarización, etc.) han definido multitud de familias de principios de privacidad, las más importantes de los cuales se listan más adelante en la Tabla 1 (pág. 23).

De todos ellos, los más relevantes para nuestro trabajo son los de la norma ISO/IEC 29100 y el propio RGPD. La norma ISO/IEC 29100 [60] define un conjunto de conceptos, actores y principios relacionados con la protección de datos personales (o '*Información personal identificable*' o "*PII*"). Los **principios de privacidad de ISO/IEC 29100** abarcan muchos de los principios existentes en familias anteriores, reorganizados y ampliados. La lista completa de principios es de esta norma es:

1. Consentimiento y elección (en la práctica también incluye elementos del principio de "notificación" de otras familias).
2. Legitimidad y especificación del propósito.
3. Limitación de la recogida.
4. Minimización de datos.
5. Limitación de uso, retención y divulgación;
6. Precisión y calidad.
7. Apertura, transparencia y aviso.
8. Participación individual y acceso (incluidos los derechos de acceso y rectificación);
9. responsabilidad (mucho más detallada que en anteriores familias, incluidas las responsabilidades con terceros, la gestión de violaciones de datos, los recursos y compensaciones, etc.).
10. Seguridad de la información.
11. Cumplimiento de la privacidad (referido a la gobernanza del cumplimiento o '*compliance*', incluida la auditoría y supervisión internas y externas, la evaluación del riesgo de privacidad y la cooperación con las autoridades de supervisión).

Cabe destacar la relevancia que estos principios otorgan a las cuestiones de minimización de datos (abordadas en los principios de limitación de la recogida, minimización de datos, y limitación de uso, retención y divulgación).

Cada uno de los principios de la ISO/IEC 29100 se descompone en un conjunto de requisitos que pueden «*guiar el diseño, desarrollo e implementación de políticas de privacidad y controles de privacidad*», y ser «*utilizados como línea de base en la monitorización y medición del rendimiento, la evaluación comparativa y las cuestiones de auditoría de los programas de gestión de privacidad dentro de una organización*». Los principios ISO/IEC 29100 se consideran aplicables en cualquier contexto y tienen un propósito global; sin embargo, al contrario de, por ejemplo, los principios de la OCDE, ISO/IEC 29100 considera que sus requisitos son «*complementarios a los requisitos legales para la protección de la PII»* y afirma que «*no pretenden ser un modelo global de*

políticas ni un marco legislativo». De hecho, gran parte de su contenido puede considerarse como meta-requisitos que dejan espacio explícito para que la «*ley aplicable*» detalle su contenido. Aun así, es innegable la influencia de los principios ISO/IEC 29100 en marcos legislativos posteriores, como el propio RGPD.

Los principios de la **legislación de la UE en materia de protección de datos** se introdujeron en la Directiva de Protección de Datos de la UE (DPD) [61] y se actualizaron en el Reglamento General de Protección de Datos (RGPD o GDPR). El RGPD incluye una serie de principios de privacidad en su art. 5 (*Principios relativos al tratamiento de datos personales*), a saber: licitud, lealtad y transparencia (desarrollado en el resto del cap. II y en el cap. III, sec. 2 y 3 del RGPD, y abordando también el consentimiento), limitación de la finalidad, minimización de datos, exactitud, limitación del plazo de conservación, integridad y confidencialidad, y responsabilidad proactiva ('*accountability*'). Aunque, a primera vista, podría parecer desconcertante que algunos principios de otras familias anteriores falten en el RGPD, muchos de ellos se abordan realmente en otros capítulos de esta norma, p. ej. los derechos del interesado se recogen en su capítulo III, las obligaciones del responsable y del encargado del tratamiento en el cap. IV y los flujos transfronterizos de datos (o transferencias de datos personales) en el cap. V (arts. 44-50).

Como continuación de este marco conceptual, merece la pena reflexionar sobre la relación de la privacidad con **otros conceptos cercanos**. Por una parte, encontramos que una idea muy próxima es el de la *'protección de datos [personales]'*. En ocasiones, esta se entiende como el conjunto de mecanismos técnicos y organizativos que garantizan la *confidencialidad* de la información personal; en otras, como la *autodeterminación* sobre la información personal propia (en el ámbito legal); o incluso, como mero *cumplimiento* legal (frente a la protección más proactiva que ofrecería una 'privacidad' así denominada y enfocada a la minimización de datos). En nuestro caso, empleamos a lo largo de nuestro trabajo la locución «privacidad y protección de datos», haciendo honor a la realidad de la variación terminológica en su uso en el ámbito legal y de la ingeniería.[2]

Por otra parte, existen claras diferencias (que delimitan el ámbito de nuestra investigación) entre la *la (ciber-)seguridad y la privacidad (y la protección de datos)*. Así, al contrario que aquella, la privacidad y la protección de datos:

- Se ciñen a *datos de tipo personal* (es decir, no abarcan aquellos datos que se refieren a organizaciones, maquinaria, etc., ni tampoco datos anónimos).
- Son *Derechos Humanos*, recogidos como tales en textos constitucionales, mandatados por las leyes que los desarrollan, y considerados por los estándares de Responsabilidad Social Corporativa.
- Consideran a la *organización poseedora de los datos como un adversario potencial* de quien hay que protegerse (a pesar de que es, a la vez, el administrador legítimo de los datos y el custodio a quien se le confían, como veremos en la siguiente sección).
- Se siguen aplicando a aquellos *datos que no son confidenciales*, o que han dejado de ser secretos; ya que el control por el interesado sigue siendo relevante.
- Se desarrollan en *principios* que no suelen tenerse en cuenta en el dominio de la ciberseguridad: consentimiento libre y control por el interesado, minimización de datos (y anonimato), responsabilidad proactiva, transparencia algorítmica, integridad contextual, etc.; así como en los derechos de los interesados de acceso, rectificación, portabilidad, etc.

[2] Somos conscientes de que el derecho a la privacidad y el derecho a la protección de datos personales no son sinónimos, parten de distintas tradiciones jurídicas, difieren en su ámbito objetivo (p. ej. la 'privacidad del domicilio' tiene poco que ver con los datos personales), y se recogen por separado en los textos legales (cf. art. 7 y art. 8 de la Carta de los Derechos Fundamentales de la UE, RGPD y Directiva 'e-Privacy', etc.) No obstante, desde la profesión de la ingeniería, se hace un uso indistinto de dichos términos. Para honrar a ambas disciplinas, en nuestro trabajo hablamos en conjunto de 'privacidad y protección de datos'.

Tabla 1. Historia de las familias de principios de privacidad.

Fecha	Fuente
1973	*Principles for fair information practice* (primera versión), del informe '*Records, Computers and the Rights of Citizens*'
1977	Objetivos del informe '*Personal Privacy in an Information Society*' (*US Privacy Protection Study Commission*)
1980, 2013	Directrices de la OCDE sobre protección de la Privacidad - *OECD Privacy Guidelines*
1981, 2018	Convención 108 del Consejo de Europa - *Council of Europe Convention 108*
1988	*Australian Privacy Act Information Privacy Principles*
1990	Principios rectores sobre la reglamentación de los ficheros computadorizados (sic) de datos personales de la ONU - *UN Guidelines for the Regulation of computerized personal data files*
1995	Directiva de Protección de Datos de la UE // *EU Data Protection Directive*
1996, 2012, 2021	*Hong Kong Personal Data (Privacy) Ordinance data protection principles*
1997	Repertorio de recomendaciones prácticas sobre la protección de datos de los trabajadores de la Organización Internacional del Trabajo (OIT) - *International Labour Organization (ILO) Code of Practice on the Protection of Workers' Personal Data*
1998	*Fair Information Practice Principles (FIPPs), 1998 US FTC Privacy Online report.*
1999	*Best Principles for Health Privacy (Health Privacy Project, Georgetown University's Institute for Health Care Research and Policy).*
2000	Principios internacionales de privacidad de puerto seguro entre la UE y EE. UU. - *International (EU-US) Safe Harbor Privacy Principles*
2000, 2003	*Canadian Standards Association Model Code for the Protection of Personal Information* (incorporados en la *Personal Information Protection and Electronic Documents Act* de Canadá *'PIPEDA'*)
2003, 2009	*American Institute of Certified Public Accountants, Inc. (AICPA) & Canadian Institute of Chartered Accountants (CICA) Privacy Framework / Generally Accepted Privacy Principles (GAPP)*
2003, 2020	Obligaciones de la *Japan Personal Information Protection Act*
2005, 2015	*APEC (Asia-Pacific Economic Cooperation) Privacy Framework principles.*
2008	*International Security, Trust and Privacy Alliance (ISTPA) 'Analysis of Privacy Principles: Making Privacy Operational'*
2011	Tecnología de la información. Técnicas de seguridad. Marco de privacidad ISO/IEC 29100 - *ISO/IEC 29100 - Information technology — Security techniques — Privacy framework*
2012	*GSMA Privacy Principles*
2016	Reglamento General de Protección de Datos de la UE – *EU General Data Protection Regulation*
2016	Principios del Escudo de la privacidad UE-EE. UU. - *EU-U.S. Privacy Shield Framework Principles Issued by the U.S. Department of Commerce*

2.2 Definición operativa y modelo de adversario para la privacidad y la protección de datos

Tal variedad de enfoques y familias de principios de privacidad revela que estamos tratando con un concepto difuso y que escapa a una definición cerrada (expresado cómicamente, se podría recurrir al conocido chiste: «*estos son mis principios [de privacidad] y, si no le gustan, aquí tengo otros*»). Incluso los autores más reconocidos en el área se han quejado de la falta de una definición homogénea: «*Es un concepto en desorden. Nadie puede articular lo que significa*» (Solove) [56, p. 1], «*Pocos valores tan fundamentales para la sociedad como la privacidad se han dejado tan indefinidos en la teoría social o han sido objeto de una escritura tan vaga y confusa por parte de los científicos sociales*» (Westin) [37, p. 5], «*La privacidad es un concepto elusivo porque es un concepto elástico*» (Margulis) [62, Sec. 2.5], «*Los intentos de dar cuenta del significado de "privacidad" han sido casi tan variados como numerosos. Tal vez esto no debería ser una sorpresa. Tampoco hay relatos universalmente aceptados del significado de "libertad", "justicia" o "igualdad".* [...] *Los teóricos han concluido que el concepto de privacidad es un concepto vago, elusivo, infectado de ambigüedades perniciosas e incluso indefinible*» (Allen) [63, p. 5].

Encontramos la respuesta a este desafío en la definición de la privacidad como un **concepto esencialmente controvertido (o cuestionado)**, proporcionada por Mulligan, Koopman y Doty [64]: en las ciencias sociales, se denomina así a aquellos conceptos cualitativos y abstractos, con una serie de características, y «*cuyo uso adecuado implica inevitablemente interminables disputas sobre sus usos adecuados por parte de sus usuarios*» (Gallie) [65] en la esencia del propio concepto, tras la búsqueda continua de un ideal. En particular, Mulligan *et al.* demuestran cómo la privacidad cumple con todos los criterios que caracterizan los conceptos esencialmente controvertidos (son objeto de valoración, internamente complejos, cuentan con descripciones diversas, son abiertos, las distintas definiciones se reconocen recíprocamente, existen ejemplos no discutidos, y se definen a partir de una competición en progreso continuo), y analizan en qué dimensiones pueden aparecer los desacuerdos, y cuyo valor concreto, por tanto, se debe especificar en cada caso para aclarar sobre qué definición de privacidad se trabaja. Gracias a ello, aunque no alcancemos una definición única e incuestionable del concepto de privacidad, esto no deberá suponer ningún problema insuperable para continuar con nuestra investigación (por mucho que a simple vista pudiera parecer extraño).

En resumen, en nuestra investigación debemos identificar en qué dimensiones pueden variar las distintas definiciones de privacidad, dónde están los puntos de acuerdo y dónde los desacuerdos, y explicar nuestra la posición ante cada uno. Por ello derivamos nuestra *definición operativa de privacidad,* construida a través del prisma de los conceptos esencialmente cuestionados para dar respuesta a las preguntas planteadas por cada una de las dimensiones identificadas por Mulligan *et al.* (según se resume en la Tabla 2)

Tabla 2. Definición operativa de la privacidad construida a partir de las dimensiones de la privacidad como concepto esencialmente controvertido.

Dimensión de la privacidad		***Pregunta***	**Respuesta desde nuestra definición operativa**
Teoría	*Metas*	*¿Privacidad para qué?, ¿qué proporciona?*	Autodeterminación y equilibrio de poder
	Justificación	*¿Por qué debería ser privado?*	Cumplimiento con los Derechos Humanos y con el marco normativo
	Concepto en contraste	*¿Qué no es privado?*	Flujo sin restricción de datos personales
	Ejemplar	*¿Cuál puede ser un ejemplo?*	Confirmación de mayoría de edad mediante credenciales anónimas
Protección	*Objeto*	*¿Privacidad de qué / sobre qué?*	Datos personales (tal como los define el RGPD). Secundariamente, otros objetivos.
	Sujeto	*¿La privacidad de quién está en riesgo?*	Interesados (sujetos de datos personales) Interesados secundarios
Daño	*Acción*	*¿Qué infringe la privacidad?*	Violaciones de los principios de privacidad y obligaciones de responsables y encargados, en especial: - Privacidad desde el diseño - Responsabilidad proactiva - Minimización de datos - Limitación de la comunicación - Limitación de la finalidad
	Infractor	*¿Quién infringe la privacidad?*	+ Ingeniero honesto pero despreocupado + Custodio de los datos personales bajo escrutinio + Intermediario de datos sin escrúpulos) (v. “modelo de adversario” más abajo)
	Por parte de quién	*¿Contra quién protege la privacidad?*	Cualquier destinatario (autorizado o no) de los datos personales
Provisión	*Mecanismo*	*¿Cómo se proporciona la privacidad?*	Disciplinas de la ingeniería de privacidad y protección de datos (gestión de riesgos, ingeniería de requisitos, aseguramiento de sistemas)
	Proveedor	*¿Quién [se supone que] debe proporcionarla?*	Responsables y encargados del tratamiento, más la cadena de suministro (incluyendo a los ingenieros de desarrollo)
Ámbito	*Límites sociales*	*¿Dónde se encuentra la privacidad?*	UE y en todo el mundo
	Escala temporal	*¿Durante cuánto tiempo se requiere?*	Depende del alcance de los riesgos para los interesados
	Alcance cuantitativo	*¿Hasta dónde se extiende?*	

En relación con esta definición operativa presentamos asimismo nuestro modelo de adversario. La privacidad es una categoría de requisitos que determina una serie de comportamientos que el sistema debe evitar, por lo que se suele abordar a partir de la definición del objeto que se intenta proteger (rasgo que comparte con la ciberseguridad, la resiliencia, etc.) Por ello, es frecuente recurrir a un modelo de las causas de esos comportamientos a evitar, en lo que se conoce como un 'modelo de adversario'. En nuestro caso, seguimos la conceptualización de modelos de adversario de Do et al. [66] para definir un modelo de adversario (resumido en la Figura 7 y la Tabla 3 y explicado a continuación) a partir de tres agentes, para quienes asumimos una serie de privilegios de acceso, metas y capacidades.

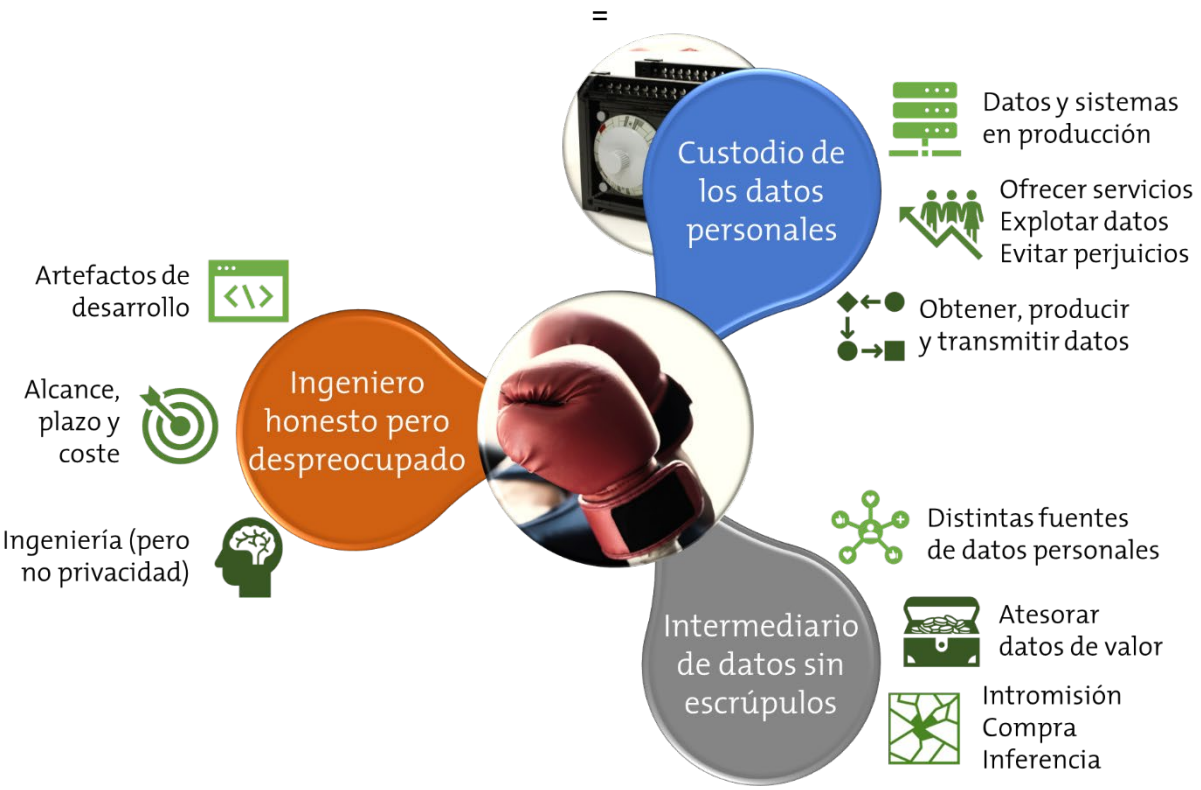


Figura 7. Modelo de adversario para la ingeniería de privacidad y protección de datos.

Tabla 3. Modelo de adversario para la ingeniería de privacidad y protección de datos.

Adversario	Acceso a	Metas	Capacidades
Ingeniero honesto pero despreocupado	Artefactos de desarrollo, pero no al sistema en producción.	Producir un sistema bajo la triple restricción de gestión de proyectos: alcance (prestaciones), plazo, y coste.	Conocimientos, habilidades y experiencia en ingeniería de software y sistemas. No tanta en ingeniería de privacidad y protección de datos.
Custodio (bajo supervisión) de los datos personales	Sistemas en producción que tratan (y almacenan) datos personales.	Generar beneficio económico: - ofreciendo sus productos y servicios, - explotando datos personales, y - evitando consecuencias negativas. Respetar principios de responsabilidad social corporativa.	Producción de datos personales (ya sean facilitados voluntariamente por el interesado, observados, derivados, inferidos, u obtenidos de fuentes secundarias). Transmisión y comunicación de datos personales.
Intermediario de datos sin escrúpulos	Otras fuentes de datos personales (en distinto grado) del propio interesado o de otros.	Atesorar el máximo de datos personales, del máximo de fuentes, sobre el máximo de individuos, y del máximo valor económico.	Acceso parcial (en distinto grado) a los datos personales albergados por el custodio mediante: - compra del acceso a los datos, - intromisión en sus sistemas o comunicaciones, o - potencia de cómputo para generar inferencias estadísticas a partir de los datos personales.

El *'ingeniero honesto pero despreocupado'*[3] se ocupa de determinadas actividades del ciclo de vida de desarrollo del sistema, para crear un producto que satisfaga, en la mayor medida posible, la triple restricción de la gestión de proyectos: alcance (prestaciones), plazo, y costes. Aunque no tiene acceso a los datos personales del sistema en ejecución, sus acciones sobre el sistema en desarrollo son cruciales para determinar los atributos de privacidad y la protección de datos que exhibirá el sistema: podrá especificar los requisitos que debe cumplir el sistema, las amenazas de las que debe protegerse, el diseño y la estructura de las soluciones en las que consiste, e incluso el proceso de desarrollo para crearlo.

Sin embargo, sus objetivos no incluyen la privacidad o la protección de datos, aunque tampoco las atacan ni se oponen activamente a ellas: simplemente, son reacios a gastar recursos o sacrificar otras prestaciones en aras de la privacidad y la protección de datos. No pretenden sabotearlas, pero carecen de incentivos económicos, experiencia o

[3] El nombre de este agente hace un juego de palabras que remite al atacante «honesto pero curioso» que suele mencionarse en los análisis de adversarios en ciberseguridad para referirse a aquellos agentes maliciosos que intentan acceden a información confidencial en un intercambio de mensajes, sin violar las especificaciones de su comportamiento en el protocolo de intercambio.

conocimientos para protegerla activamente (de ahí el calificativo de «honesto pero despreocupado»).

El *'custodio de los datos personales'* procesa o trata datos personales de los interesados (en el sentido más amplio: recopila, almacena, analiza, etc. datos que se pueden relacionar con un individuo) y accede sin restricciones a los sistemas que opera y que contienen o transmiten estos datos (a salvo de las posibles medidas técnicas aplicadas para proteger su seguridad). Desde la perspectiva del RGPD, este custodio de datos puede ser un responsable o un encargado del tratamiento. Como organización que opera con éxito, su objetivo es maximizar su ganancia económica, apoyándose en la explotación de datos personales para mantener su operativa de negocio, realizar perfiles de usuario, o venderlos al mejor postor. Asumimos que, ya sea por razones económicas o de responsabilidad corporativa, el custodio pretende comportarse éticamente y no va a atacar proactivamente la privacidad (en caso contrario, desbordaría nuestro modelo de adversario).

El *'intermediario de datos sin escrúpulos'* trata de acumular todos los datos personales que le sea posible conseguir y venderlos en un mercado de datos, motivado por el valor económico que puede extraer de los datos personales a los que accede. Intentará obtener tantos datos, sobre tantas personas, y de tantas fuentes como sea capaz (por medios lícitos o ilícitos). Asumimos, en todo caso, que tiene acceso a múltiples fuentes que le permitirán inferir nuevos datos a través de la aplicación de métodos estadísticos.

2.3 Paradigmas dominantes actualmente para abordar la privacidad y protección de datos

En la sección anterior se ha enmarcado el concepto de privacidad, como elemento central de nuestra investigación, y hemos delimitado el uso que daremos a este concepto, así como el modelo de adversario frente al que buscamos protección. Ante ello, cabe preguntarse si la manera en la que se aborda actualmente la cuestión de la privacidad protege de este modelo de adversario. En particular, nos preguntamos *cómo se considera al ingeniero 'honesto pero despreocupado'* en los enfoques que prevalecen para abordar la privacidad y la protección de dados.

Para ello, merece la pena que nos detengamos en los **agentes que intervienen en relación con la privacidad** según la norma ISO/IEC 29100 [60, Sec. 4] (cuyos principios hemos presentado en la sección anterior), y que coinciden básicamente con los que define el RGPD [67] Art. 4:

- El *interesado* (*data subject*) es la «*persona física identificada o identificable* [...] *cuya identidad pueda determinarse, directa o indirectamente*» (es decir, por referencia a «*un identificador* [...] *o a uno o más elementos* [...] *de dicha persona física*»), a quien se refieren los datos personales. Así, será un dato personal cualquier dato que se refiera a un interesado que pueda ser identificado directa o indirectamente (lo que los opone a los datos anónimos, pero incluye a los datos seudónimos).[4]
- El *responsable del tratamiento* (*data controller*) es una «*persona física o jurídica* [...] *que, solo o junto con otros, determina los fines y los medios del tratamiento* [...]». Es quien decide a qué tratamientos se someten los datos personales, para qué (propósito) y cómo (medios) y, por lo tanto, es responsable de cumplir con la normativa aplicable de privacidad y protección de datos. No obstante, puede encargar a otros (por ejemplo, a un proveedor de servicios) que le proporcionen esos medios (v. a continuación).
- El *encargado del tratamiento* (*data processor*) es la «*persona física o jurídica* [...] *que trata los datos personales por cuenta del responsable del tratamiento*». Aunque el

[4] En otras fuentes se habla de '*información personal*' o '*PII*' ('*información personal identificable*') en lugar de 'datos personales' (y, en lugar de interesados o '*data subjects*', se refieren a '*PII principals*'). Aunque existen sutiles diferencias entre ambos conceptos, podemos tomarlos como sinónimos a los efectos de nuestra investigación.

responsable del tratamiento es quien toma las decisiones sobre los tratamientos de datos personales, el encargado tiene el compromiso y la responsabilidad de llevar a cabo esas decisiones a través de los medios que facilita.

- Un *tercero* (*third party*) es «*una persona física o jurídica* [...] *distinta del interesado, el responsable del tratamiento, el encargado del tratamiento y las personas que, bajo la autoridad directa del responsable o del encargado del tratamiento, estén autorizadas a tratar datos personales*». Cabe señalar que este tercero puede convertirse en un responsable en sus propios términos una vez que haya recibido los datos personales [60, Para. 4.2.4].[5]
- Un *destinatario* es «*la persona física o jurídica, autoridad pública, servicio u otro organismo al que se comunican datos personales, se trate o no de un tercero*». Este término general abarca a cualquier persona que reciba datos personales del responsable del tratamiento, es decir, puede tratarse de un tercero, pero también de un encargado del tratamiento, un corresponsable (*joint controller*), o un empleado del responsable del tratamiento.

Encontramos, pues, que los responsables y encargados del tratamiento son corresponsables de las actividades de tratamiento de datos, pero el alcance de su responsabilidad es diferente, ya que este último actúa en nombre de los primeros, en lugar de satisfacer sus propios objetivos. Esto se aborda mediante marcos legales y reglamentarios de privacidad y protección de datos, que generalmente incluyen obligaciones específicas de responsables y encargados. Pero… *¿no falta alguien en esa lista?* Hay un rol que no aparece y cuyo papel es crítico a la hora de determinar el comportamiento de los tratamientos de datos personales y, por tanto, su respeto a la privacidad y la protección de datos: *quienes proveen* los componentes, aplicaciones, servicios y sistemas desarrollados para dar soporte al tratamiento y, en particular, los ingenieros encargados de crear del sistema. Esto se intenta reflejar en la Figura 8, donde presentamos los distintos flujos de datos (lícitos o no) que pueden darse (según [60, Sec. 4.3]) entre los distintos agentes presentados más arriba, y cómo por debajo de todos ellos está el rol 'oculto' de los **ingenieros que desarrollan un sistema**, y que han sido tradicionalmente ignorados por los marcos normativos tradicionales de privacidad y protección de datos.

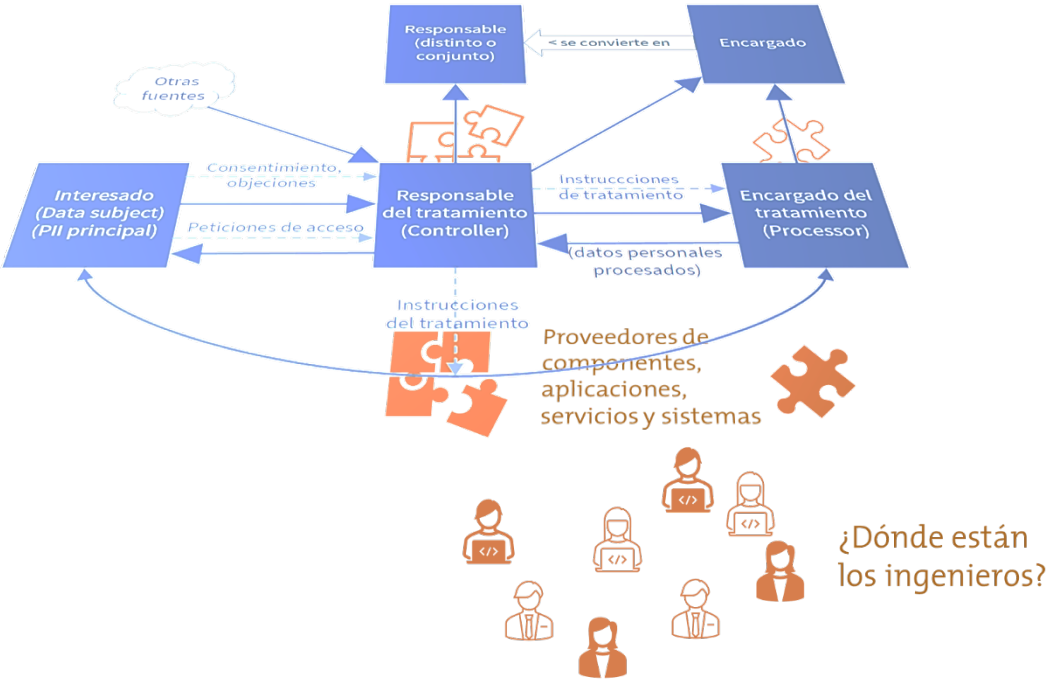


Figura 8. El ingeniero tras los roles en relación con la privacidad y la protección de datos.

[5] Téngase en cuenta también que esto representa el concepto de "*tercero*" bajo el RGPD, que puede diferir de conceptos con nombres similares en otros marcos legales[157].

De hecho, los **paradigmas dominantes** en el campo de la privacidad y la protección de datos *no son suficientes* para abordar el problema desde el ámbito de la ingeniería (como hemos esbozado en la motivación presentada en la sección 1.2), lo que pretende representar la Figura 9, desarrollada a continuación:

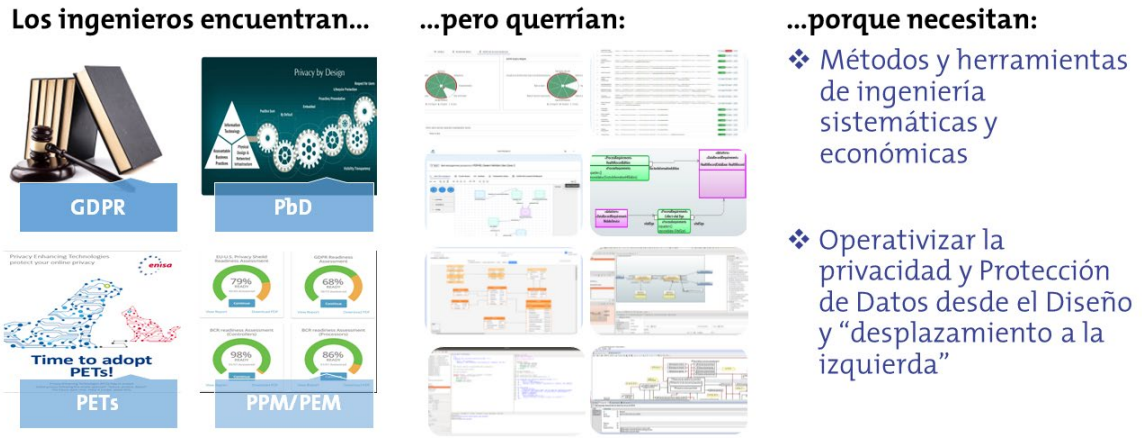


Figura 9. La brecha de la privacidad en la ingeniería.

1. La **privacidad y la protección de datos desde la regulación** (p.ej. RGPD) emplean conceptos ajenos al ámbito de la ingeniería, con los que los ingenieros no están preparados para lidiar, y que les resultan ambiguos desde la perspectiva de la implementación práctica. A menudo, esto degrada el papel de la privacidad y la protección de datos en el desarrollo de sistemas a una mera cuestión de documentar las políticas de privacidad. En ocasiones, los ingenieros simplemente reciben tales políticas como un marco estricto del que no pueden desviarse, sin haber recibido información sobre cómo implementarlas; en otras, se intenta obtener el consentimiento del interesado, de cualquier manera, para aquellos tratamientos que los ingenieros han decidido sin ningún análisis sobre sus consecuencias. Aparte de los problemas con el consentimiento en sí (por ejemplo, que no sea realmente espontáneo o informado), este enfoque desdeña por completo otros principios como el de minimización de datos [68].
2. La **privacidad y protección de datos desde el diseño (PbD / DPbD)** abogan por introducir las cuestiones de privacidad desde el inicio de un proyecto, en lugar de abordarlas como una ocurrencia tardía. Sin embargo, en la práctica, esto no ha resultado más que un eslogan con pocas herramientas prácticas que apoyen su aplicación para hacerlo efectivo en la ingeniería de software [69]. Por mucho que el PbD/DPbD se afirme en textos legales y declaraciones institucionales, se postula como un desiderátum construido a partir de principios abstractos y aspiracionales, sin un correlato técnico operativo que deje claro "*el quién, qué, cuándo, por qué y cómo*" [31], [70], [71] de ese diseño.
3. Las **tecnologías de mejora de la privacidad (PETs)**, en sus distintas variantes como '*PETs duras*' u orientadas a la minimización [72], [73][74][75], '*PETs blandas*' u orientadas a la transparencia [72][73], [76]–[79][80][31][75], y *PETs de interacción social*, abordan la privacidad desde distintos enfoques, respectivamente denominados como "*privacidad por política*" (o '*blanda*', que acepta que el interesado ha perdido el control de sus datos y debe confiar en responsables y encargados del tratamiento como custodios de sus datos), "*privacidad por arquitectura*" [81] (o '*dura*', que considera al custodio de datos o al proveedor de servicios como un adversario en sí mismo) y "*privacidad por interacción"* [81]. En realidad, estas tecnologías ofrecen un enfoque 'artesanal' para abordar la privacidad y protección de datos, ya que no proporcionan los métodos sistemáticos y económicos que serían esperables de un proceso ingenieril.
4. Las **herramientas de gestión de privacidad y protección de datos** abarcan otras áreas corporativas como la Tecnología de la Información o la Dirección Operativa, pero no abordan el proceso de desarrollo. Estas herramientas pueden agruparse en dos grandes familias [82], que, aunque están relacionadas superficialmente con nuestros

objetivos (pretenden mejorar la protección de la privacidad en los servicios que ofrecen las organizaciones), no abordan los problemas que enfrentan los ingenieros durante el ciclo de vida de desarrollo de sistemas. Por un lado, las herramientas de *Gestión de Programas de Privacidad* (PPM) se dirigen al departamento legal, incluyendo herramientas como gestores de evaluación, herramientas de mapeo de datos, administradores de consentimiento, gestores de documentos sobre privacidad, sistemas de seguimiento de consultas sobre incidentes, peticiones de acceso a datos, etc. Estas herramientas están completamente desconectadas de la perspectiva de la ingeniería, ya que no pueden usar como entrada ningún producto de ingeniería de software y sistemas, ni vinculan los consentimientos al aseguramiento del proceso de desarrollo, ni asignan requisitos legales a actividades o procesos de ingeniería, ni abordan el ciclo de vida de desarrollo del sistema, etc. Por otro lado, las herramientas de *gestión de privacidad empresarial (EPM)* se dirigen a los departamentos de tecnologías de la información, y ofrecen funciones que van desde el monitoreo de la actividad, al descubrimiento de datos, disociación o anonimización, etc. A pesar de la apariencia, estas herramientas abordan el funcionamiento de los sistemas en operación, más que su desarrollo, y en su mayoría desempeñan un papel de apoyo cuyos beneficiarios finales son, de nuevo, los departamentos jurídicos, ya que sus resultados no pueden ser consumidos por el ciclo de vida de la ingeniería.

Frente a ello, abogamos por un cambio de paradigma que corresponde a lo que se ha dado en llamar **Ingeniería de la Privacidad**, un concepto adoptado por muchos autores, y que ofrece una visión *complementaria a las de esos cuatro paradigmas* dominantes que acabamos de introducir y, en particular, proporciona un punto en el que podemos basar nuestro trabajo para dar una *respuesta contra sus inconvenientes* [7], [30], [31], [81], [83]–[91]

1. Frente al enfoque puramente regulatorio, nuestro trabajo basado en la ingeniería de privacidad considera cómo se puede integrar la implementación de estas normativas con las principales herramientas y métodos que los ingenieros emplean y aplican a diario, para ayudarles a cumplirlas.
2. Con respecto a la privacidad y protección de datos desde el diseño (PbD / DPbD), no pretendemos quedarnos en unos principios generales a los que aspirar, sino que nuestro objetivo es permitir que los ingenieros introduzcan efectivamente la privacidad a lo largo del ciclo de vida de la ingeniería de software, a través de herramientas y métodos que integran la práctica de la privacidad en diferentes disciplinas de la ingeniería de software.
3. En relación con las PET, nuestro trabajo tiene como objetivo proporcionar métodos que vayan más allá de las soluciones artesanales, y ayudar a que los ingenieros sistematicen la implementación de los controles de privacidad más adecuados sin exigirles ser expertos en privacidad y protección de datos, ni que tengan que realizar una investigación profunda sobre el estado del arte.
4. Y, en contraste con los enfoques de gestión, nuestro trabajo se dirigirá a los ingenieros, con el objetivo de que puedan utilizar herramientas y métodos adaptados a sus necesidades, mientras se integran con los que utilizan en su trabajo diario.

En definitiva, se trata de ofrecer métodos y soluciones *alineados con el punto de vista de los ingenieros, para integrarlos en su trabajo cotidiano* y en los métodos y herramientas que están acostumbrados a emplear, *minimizando su necesidad de conocimiento experto* en el ámbito de la privacidad y la protección de datos. Como se dijo anteriormente, consideramos este enfoque de Ingeniería de la Privacidad como complementario a los paradigmas previamente existentes, ya que entendemos que la privacidad es un tema complejo que debe abordarse desde múltiples perspectivas (en lo que se ha dado en llamar '*privacidad por red*'), sin caer en un solucionismo tecnológico que no dejaría de ser otro reduccionismo [92]

3 Evaluación de necesidades y análisis de deficiencias

Este capítulo presenta los resultados de un doble análisis para enmarcar y justificar la necesidad de las contribuciones presentadas en el resto de nuestro trabajo de investigación.

Anteriormente, hemos mencionado que nuestras contribuciones a la ingeniería de software y sistemas para la privacidad y la protección de datos abordarían cinco disciplinas (gestión de riesgos, ingeniería de requisitos, diseño de sistemas basado en modelos, aseguramiento de sistemas, e ingeniería de métodos). Esta elección es el resultado de un exhaustivo **proceso de elicitación de necesidades**, a través del cual hemos analizado las necesidades que conlleva la operativización de los contenidos del RGPD en términos de ingeniería de software y sistemas. En concreto, para asegurarnos de la *alineación de nuestros resultados con el ámbito regulatorio y las necesidades de los ingenieros:*

(1) Hemos involucrado a esos ingenieros a través de un conjunto de entrevistas dirigidas, donde hemos aplicado la técnica de investigación conocida como *'Modelos Mentales'*.
(2) Hemos analizado sistemáticamente las implicaciones de los *contenidos del RGPD en las actividades de ingeniería*, las lagunas que no quedan cubiertas por los enfoques actuales y su correspondencia con las actividades y procesos de las principales disciplinas de software e ingeniería de sistemas.

3.1 Método para la captura de necesidades y análisis de deficiencias: modelos mentales orientados a tareas

Cuando las personas llevamos a cabo una actividad, lo hacemos según nuestro propio modelo mental (consciente o inconsciente, aunque sea impreciso, limitado, o incluso erróneo) [93]–[95]: nuestra representación del mundo que nos rodea, nuestra interpretación de los procesos que adivinamos a partir de lo que podemos ver, y nuestras creencias sobre las consecuencias de nuestras acciones. Todo ello depende de nuestra experiencia previa y, aunque nuestros modelos mentales pueden sufrir adaptaciones, la inercia suele prevalecer sobre la flexibilidad.

Por ello, cuando se diseña un proceso de trabajo, es fundamental que esté alineado con los modelos mentales de quienes lo van a llevar a cabo. Para ello, existen técnicas que tratan de asegurar este alineamiento. En particular, **Mental Models [96]** es una *técnica de investigación de usuario generativa, cualitativa y centrada en el espacio del problema*, cuyo objetivo es aclarar cuáles son las *tareas* que pretenden realizar las personas, sus *objetivos*, las *acciones* que realizan para alcanzarlos y las *preocupaciones* que enfrentan al realizar alguna actividad. Por lo tanto, este método resulta perfectamente adecuado para extraer las necesidades del usuario —propósito para el que lo estamos aplicando en nuestro trabajo—, correspondiendo dicho usuario, en nuestro caso, al ingeniero que desarrolla sistemas, y sus tareas, a aquellas para satisfacer los requisitos de privacidad y la protección de datos en los sistemas que desarrollan.

El método gira en torno a una serie de *entrevistas a los usuarios del sistema, producto, o método bajo desarrollo, semiestructuradas y no dirigidas*. Para ello, es necesario definir y preparar la entrevista, ejecutarla, y analizar sus resultados.

El primer paso consiste llevar a cabo una '*segmentación basada en tareas'* de los perfiles de los entrevistados potenciales, para definir los distintos segmentos del público objetivo que deben ser muestreados para las entrevistas. Los criterios de segmentación no son demográficos o psicográficos (al contrario que en el marketing), sino que la segmentación debe centrarse en distinguir los comportamientos que lleva a cabo cada grupo.

A continuación, se deben preparar las entrevistas de los Modelos Mentales como **entrevistas semiestructuradas**: están orientadas hacia un objetivo específico (la captura del modelo mental del usuario, incluyendo sus tareas, etc.) y requieren de una preparación; sin embargo, tienen una naturaleza abierta que las diferencia de las encuestas o cuestionarios cerrados. El alcance de la entrevista debe establecerse a partir de los objetivos del proyecto general en cuestión, y se puede crear *una lista abierta de indicaciones* a partir de los comportamientos de lluvia de ideas durante la fase de segmentación. Los entrevistadores deben asegurarse de que la entrevista resulte en la obtención del modelo mental del usuario, pero no deben llevarlos a temas específicos (o, de ninguna manera, sugerir posibles respuestas). La Tabla 4 presenta ejemplos de las preguntas utilizadas como guía en nuestras entrevistas a ingenieros de software y sistemas.

Tabla 4. Indicaciones para las entrevistas semidirigidas a los ingenieros de software y sistemas acerca de sus tareas en relación con la privacidad y protección de datos.

Sobre el tratamiento de datos personales:

1. *¿Cuál es la relación de los productos que desarrollas con los datos personales / información personal?*
2. *¿Los productos que desarrollas usan datos personales?* [Decimos '*usan*' para evitar cualquier mención al '*tratamiento*' de datos personales.] [Alternativamente:] *¿Sabe si sus sistemas tratan con datos personales?*
3. *¿Qué datos personales usa? ¿Obtiene algún dato personal aparte de ellos* [por ejemplo, el perfil de usuario]*? ¿Algún otro dato personal que recopile?* [Alternativamente:] *¿Sabe qué se consideran estrictamente datos personales?*
4. [Para cada categoría de datos personales:] *¿Con qué fines se emplean esos datos personales? ¿Qué hace con los datos recopilados? ¿De dónde se recopila esta información* [por ejemplo, un mercado de datos]*?*

Acerca de la protección de datos personales:

5. *Al desarrollar, ¿tiene en cuenta la privacidad y la protección de datos personales?*
6. *¿Cómo aborda la privacidad y la protección de datos (si lo hace)?*
7. [Centrarse en las tareas a realizar, en lugar de en las actividades para realizarlas:] *¿Qué hace para proteger los datos personales? ¿Quién lleva a cabo eso? ¿Tiene soporte de herramientas para esto? ¿Cree que es relevante?* [Alternativamente:] *¿Por qué? ¿O quizás solo a veces? ¿O 'no para mi sistema'?*

Acerca de la privacidad y la protección de datos en el ciclo de vida del desarrollo (SDLC):

- *¿Cuándo introduce la protección de datos en su flujo de trabajo de ingeniería?* [por ejemplo, en qué etapa/actividades:] *¿requisitos?*, *¿validación?*, *¿operación?*

Acerca del cumplimiento del RGPD:

8. *¿Qué implicaciones prácticas conoce del RGPD?*
9. *¿Qué dificultades encuentra para proteger adecuadamente los datos personales?*
10. *¿Se siente lo suficientemente seguro como para saber si está abordándolo adecuadamente?*
11. *¿Cómo sabe que los datos personales que recopila cumple con el RGPD?*
12. *¿Cómo sabe que no tiene ningún otro dato personal?*
13. *¿Sabes cuánto costaría cumplir?*
14. *¿Sabes cómo lograría cumplir* [por ejemplo, recursos a seguir, métodos o procesos]*?*

Sobre las medidas técnicas de protección de datos y seguridad durante el proceso de desarrollo:

15. *¿Cómo garantiza la integridad de los datos personales que almacena?*
16. *¿Tiene algún control sobre quién puede acceder a sus interfaces de programación ('APIs')?*
17. *¿Cómo acceden a ello? ¿Cómo inician sesión?*
18. *¿Tiene alguna medida técnica para garantizar que los desarrolladores no almacenen contraseñas?*
19. *¿Gestiona la información personal que podría ser almacenada por los usuarios?*
20. *¿Ejecuta algún servicio de telemetría?*
21. *¿Envías directamente su código a un repositorio público de código abierto?*

…

El objetivo de la entrevista es obtener una **lista de tareas** que los entrevistados buscan resolver. En el lenguaje de los modelos mentales, el concepto de 'tarea' es más amplio que lo que podría esperarse: son *«acciones, pensamientos, sentimientos, filosofías y motivaciones»* que pueden abarcar *tareas propiamente dichas* (acción o paso llevado para cumplir algún trabajo u objetivo), *tareas implícitas* (afirmación de que alguien deba realizar una tarea), *tareas de terceros* (que realiza alguien distinto del interesado), *filosofías* (lógica que determina que una tarea deba hacerse), y *motivaciones* (emocionales que impulsan a alguien a realizar una tarea). Las tareas recogidas deben tener un nivel de granularidad apropiado para el alcance del proyecto en cuestión. Asimismo, las tareas no deben confundirse con otros conceptos, relacionados pero diferentes, como preferencias, deseos, o expectativas, medios de producción, exposiciones de hechos, explicaciones del funcionamiento interno, circunstancias, y quejas.

Las **técnicas para la entrevista** se pueden resumir en un conjunto de siete reglas:

1) Acciones en lugar de sentimientos (comportamientos y filosofías en lugar de preferencias de productos).
2) Preguntas abiertas.
3) Utilizar los términos del usuario (aplicando una '*escucha activa*').
4) Seguir la conversación (a partir de las respuestas del entrevistado).
5) Objetivos más que herramientas (no *qué* utiliza, sino *para qué*.
6) Experiencia inmediata (no de un pasado histórico, o de otros individuos).
7) Granularidad apropiada.

Tras las entrevistas, se lleva a cabo un proceso de análisis para **construir el modelo mental y representarlo en un diagrama** (llamado igualmente '*modelo mental*'). La Figura 10 muestra el aspecto de un diagrama de Modelo Mental de ejemplo, explicado a continuación.

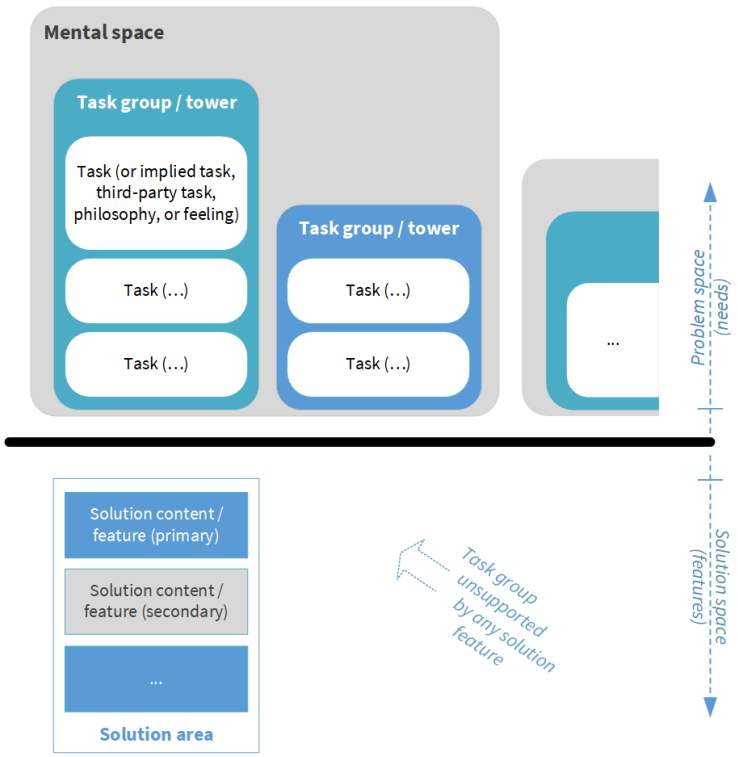


Figura 10. Notación de los diagramas de Modelos Mentales.

En un diagrama de modelo mental, las tareas se agrupan por afinidad en '*espacios mentales*' y se contraponen a las características del producto (sistema, método, etc.) que aborda esas tareas. Para ello, se peinan las transcripciones de las entrevistas buscando *tareas atómicas*. Luego, el investigador busca patrones que permitan agruparlas en

definiciones de *tareas comunes*, estas a su vez en *torres* que representan grupos conceptuales y, finalmente, en *espacios mentales* que representan objetivos y tiempos similares. Entonces la propia organización puede sugerir nuevas tareas en función de su experiencia.

Una vez construido el modelo, se aborda la *alineación y análisis de brechas*: se dibuja una lista o mapa de contenido con todas las características actuales (o planificadas); y cada elemento se alinea debajo de la(s) torre(s) que soporta. Ahí, se detectan desalineaciones, tales como huecos en torres (tareas) que corresponden a una funcionalidad dada, huecos en la funcionalidad de una torre, y escasez de funcionalidades primarias para una tarea. El diagrama resultante muestra las tareas que están adecuadamente cubiertas por las funciones actuales o planificadas, y aquellas en las que se necesitarían nuevos desarrollos.

3.2 Análisis de deficiencias en el desarrollo de sistemas de los métodos y herramientas de gestión de privacidad y protección de datos

Una vez realizada una serie de entrevistas siguiendo el mencionado método de los Modelos Mentales, el principal resultado hallado fue precisamente la **ausencia de concienciación y conocimientos de privacidad y protección de datos**, por lo que no se manifestaron muchas tareas que los ingenieros estuvieran llevando a cabo, lo que nos llevó a recurrir al modelado manual de las tareas que se pudieran inferir del contenido del propio RGPD. El resultado de todo ello es un análisis de deficiencias de los actuales métodos y herramientas de gestión de la privacidad en el desarrollo de sistemas, en relación con los contenidos del RGPD, y una correspondencia entre esas deficiencias y las características de las disciplinas de ingeniería desde las que proponemos abordar la privacidad y la protección de datos. El Modelo Mental se estructura así en **cuatro espacios mentales**:

1. *Ingeniería de Requisitos y Análisis*, tanto desde una perspectiva orientada a metas que aborda la captura y el análisis de requisitos, como desde otra basada en la evaluación del riesgo que se vincula más con la disciplina de Gestión de Riesgos.
2. *Diseño y Arquitectura de Software y Sistemas*, que aborda el proceso para trasladarse desde los requisitos a las soluciones que las satisfagan, incluyendo la toma de decisiones respecto al tratamiento de datos personales, la distribución del tratamiento a lo largo de la cadena de valor, y la toma de medidas de protección.
3. *Gobernanza de Proyecto y Aseguramiento de Sistemas*, como disciplina transversal que va en paralelo al ciclo de vida del desarrollo (SDLC) y a los usuarios de sus resultados, y que aborda cuestiones como la demostración del cumplimiento responsable, la distribución de la responsabilidad a lo largo de la cadena de valor, el registro de operaciones de tratamiento, la notificación de violaciones de datos personales, y la certificación.
4. *Operación del Sistema*, que, a pesar de estar más allá del alcance estricto de nuestro trabajo, nos iluminará sobre las características que deben abordar las otras disciplinas, y donde se abordan las operaciones de tratamiento de datos personales, el respeto a los principios de privacidad durante la ejecución de dichas operaciones, y las medidas de respaldo al ejercicio de los derechos de los interesados.

La organización de estos cuatro espacios se ilustra en la página siguiente en la Figura 11, apaisada (tan solo se muestra la estructura general, los detalles de la figura son visibles en el fichero PNG adjunto a esta memoria).No entraremos aquí en detalle en el contenido de cada uno de esos espacios mentales, para lo que remitimos al lector a la memoria de la tesis doctoral donde se recoge y desarrolla este trabajo de investigación [1, Secs 3.3, 3.4, 3.5].

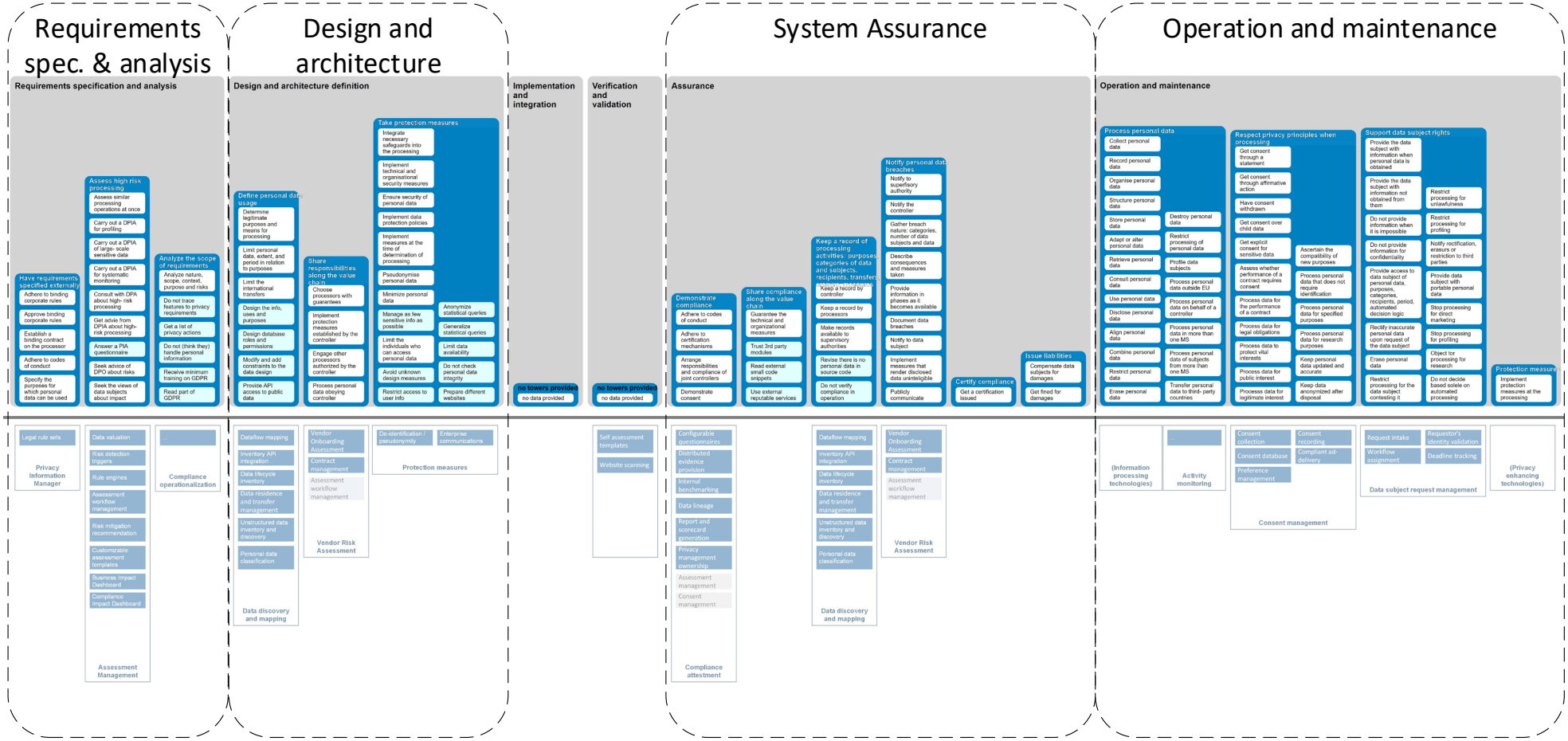


Figura 11. Modelo Mental del Soporte al RGPD.

A partir del análisis de este modelo, podemos extraer dos tipos de resultados que mostramos a continuación.

En primer lugar, si analizamos el papel de las disciplinas correspondientes en relación con su *soporte a la ingeniería de privacidad y protección de datos* y al **cumplimiento con el RGPD** en particular, podemos ver cómo muchas de las actividades típicamente asociadas a las distintas disciplinas de la ingeniería pueden contribuir al cumplimiento de la regulación de protección de datos. La Tabla 5 intenta sintetizar los resultados de nuestro análisis acerca del papel de cada una de las disciplinas de ingeniería que abordamos en nuestro trabajo en relación con el RGPD (de nuevo, remitimos al lector al correspondiente capítulo de la tesis doctoral del autor para una explicación más detallada [1, Secs 3.3, 3.4, 3.5]).

Tabla 5. Correspondencia con el RGPD de las disciplinas de la ingeniería de software.

Correspondencia con el RGPD de las disciplinas de ingeniería
Gestión de riesgos
• Evaluaciones de impacto de protección de datos (EIPD) (art. 35, art. 36 + recomendaciones del GT29/WP29 y las autoridades de supervisión nacionales), incluyendo la determinación de la necesidad de una EIPD, los contenidos mínimos de la evaluación, y las actividades de consulta previa, y las acciones en caso de tratamientos de alto riesgo. • Seguridad del tratamiento (art. 32.2) • Riesgos por violaciones de datos personales (art. 33.1, art. 34.1) • Riesgos en la cadena de suministro: encargados del tratamiento (Art. 28), encargados secundarios (Art. 28.4), responsables conjuntos (Art. 26), destinatarios en terceros países (Art. 44-50) y en general (Art. 29), terceras partes. • Riesgos financieros para la organización: derecho a indemnización y responsabilidad (Art. 82), multas administrativas (Art. 83), sanciones (Art. 84), derecho a tutela judicial (Art. 79). • Otros riesgos derivados de discriminación por el perfilado de los interesados (considerando 71, Art. 22).
Ingeniería de requisitos
• Principios (Capítulo 2) • Derechos de los interesados (Capítulo 3) • Obligaciones y responsabilidad del responsable y encargado del tratamiento (Capítulo 4, esp. art. 24 – 34) • Medidas técnicas y organizativas (art. 24, art. 32.1) • Transferencias a terceros países (Capítulo 5) • Determinación del ámbito de los requisitos: naturaleza y categorías de los datos, incluyendo en su caso aquellos especialmente sensibles (art. 9); licitud del tratamiento (art. 6), ámbito y contexto del tratamiento, especialmente la elaboración de perfiles (art. 22); finalidad, incluyendo finalidades secundarias de investigación (art. 89); tamaño de la organización (art. 30); riesgos de los tratamientos (art. 24.1, art. 35.1, art. 5.1.b) …
Diseño
• Identificación datos personales y actividades de tratamiento (art. 1–4) • Prerrequisito para respaldar desde la ingeniería a los derechos del interesado (Cap. 3) y los principios del tratamiento, especialmente el principio de licitud. • Respaldo de las obligaciones del responsable y el encargado (Cap. 4). • Medidas de protección de datos (art. 24, art. 25), medidas técnicas y organizativas de seguridad (art. 32.1) incluyendo control de acceso (art. 32.4, art 25.2), medidas aplicables por los encargados (art. 28.1, art. 28.3) y encargados secundarios (art. 28.4), salvaguardas en transferencias internacionales (art. 40, art. 46, art. 49) y para finalidades de investigación (art. 89). • Soporte a derechos de los interesados (Cap. 2): necesario modelar los datos personales en el diseño para poder determinar el ámbito de los derechos.

Aseguramiento
• Principios de responsabilidad proactiva o rendición de cuentas (art. 5.2), y transparencia (art. 5.1.a, art. 12). Demostración a las autoridades del cumplimiento (art. 24.1, art. 31, art. 38), y por los encargados a los responsables (art. 28.3.h.) • Distribución de responsabilidad entre distintas partes: responsables conjuntos (art. 26), encargados (art. 28.4), encargados secundarios (art. 28.5), transferencias internacionales (art. 44, art. 47). Asistencia en el cumplimiento de los encargados, certificación códigos de conducta, políticas… • Registro de evidencias de: consentimiento del interesado (art. 7), medidas de protección (art. 24, art. 28.3.c, art. 32, art. 35.7.d), resultados de EIPDs (art. 25), registros de tratamientos de datos (art. 30), notificaciones de violaciones (art. 33, art. 34)… • Regulación suplementaria: pautas del GT29/WP29 y el CEPD/EDPB (art.70), excepciones (art. 49, Ch. 10), códigos de conducta (art. 40, 41), certificación (art. 42), decisiones de adecuación de transferencias internacionales (art. 45), cláusulas contractuales estándar (art. 46) y normas corporativas vinculantes (art. 47).
Ingeniería de métodos
• Fuentes normativas complementarias (incluyendo auto-, co-, y cuasi-regulaciones): pautas y opiniones del GT29/WP29 y el CEPD/EDPB, códigos de conducta, normas corporativas vinculantes, cláusulas contractuales normalizadas, requisitos de las EIPDs, provisiones y excepciones de Estados miembros y actos delegados, enmiendas futuras, etc.

En segundo lugar, si analizamos el respaldo actual a estas necesidades de los ingenieros por parte de las **funciones de herramientas existentes de gestión de privacidad**, de acuerdo con la clasificación de la IAPP [97] (International Association of Privacy Professionals), encontramos que algunos tipos de herramientas abordan algunas de esas tareas: herramientas de gestión de EIPDs ('*PIA tools*', incluyendo cuestionarios, reglas, etc.), de gestión de riesgos de proveedores ('*Vendor Risk Management*'), de operativización del cumplimiento ('*compliance*'), tecnologías de mejora de la privacidad ('*PETs*', incluyendo gestión de identidad, seudonimato, comunicaciones cifradas, etc.). Sin embargo, deberíamos hablar más bien de la falta de respaldo por esas herramientas: su problema general en relación con nuestro propósito es que *no representan los activos en términos de artefactos de ingeniería* (p. ej. modelos de flujo de datos, componentes arquitectónicos, etc.) Como explicamos al analizar los paradigmas dominantes en el capítulo 2, bien se dirigen a equipos legales, o bien a equipos operación de sistemas, pero no a los ingenieros que los desarrollan. Además, frecuentemente se centran en cuantificar el impacto de un incumplimiento para la compañía, más que el impacto del perjuicio potencial para el interesado.

Finalmente, como **conclusiones de ámbito más general** de este análisis de deficiencias, podemos observar que:

- La privacidad y la protección de datos se deben tener en cuenta a lo largo de todo el ciclo de desarrollo software.
- No obstante, esto *desafía el modelo mental asentado* de los ingenieros de software.
- Las herramientas actuales de gestión de la privacidad no están integradas con las necesidades de los ingenieros de software, ni con sus herramientas habituales.
- Cuando los ingenieros de software declaran abordar aspectos de privacidad y protección de datos, se centran en realidad en aspectos de ciberseguridad de la información.

4 Ingeniería basada en modelos para la privacidad y la protección de datos

El desarrollo de un sistema parte normalmente de una visión general y abstracta, para ir avanzando hacia un diseño cada vez más refinado. Este proceso suele incluir una actividad de modelado del sistema, que abstrae una serie de elementos (y relaciones entre ellos) que se consideran relevantes para caracterizar el sistema. De este modo, cuando estemos abordando el estudio del sistema, podemos reemplazarlo por ese modelo simplificado que, de alguna manera, representa al sistema (y viceversa). Los modelos se pueden aprovechar para integrar fácilmente tanto elementos específicos del sistema como aspectos transversales, por lo que se ajustan bien a nuestro propósito de integrar en la ingeniería de sistemas y software convencional las cuestiones de privacidad y protección de datos, y esa es la razón por la que confiamos en ellos para respaldar nuestra investigación.

Antes de pasar a los modelos específicos utilizados para cada una de las subdisciplinas de ingeniería que abordamos (gestión de riesgos, ingeniería de requisitos, diseño de sistemas, aseguramiento de sistemas, e ingeniería de métdos), este capítulo muestra cómo la ingeniería basada en modelos puede contribuir a introducir la privacidad y la protección de datos en el ciclo de vida de desarrollo, y ofrece un *lenguaje específico de aspectos de dominio* para introducir el modelado de los aspectos de privacidad y protección de datos en lenguajes de modelado ya existentes.

4.1 Ingeniería de software y sistemas basada en modelos (MBSSE)

Cuando estamos tratando con un sistema, un **modelo** de ese sistema es una *abstracción que puede reemplazar el sistema a efectos del estudio* que se esté llevando a cabo [98] La amplitud de este concepto permite aplicar el término '*modelo*' a muchas realidades (v. Figura 12): desde las miniaturas que reproducen un modelo a escala de una máquina industrial hasta las ecuaciones que describen el modelo matemático de un fenómeno físico. En el caso de los sistemas software, los modelos se encuentran a caballo entre esos dos tipos: no dejan de ser '*máquinas*' (programas) ejecutables, pero se expresan en un lenguaje formal de modelado, y se pueden representan de distintas maneras, incluyendo diagramas, cuya forma depende del tipo de lenguaje de modelado que se use.

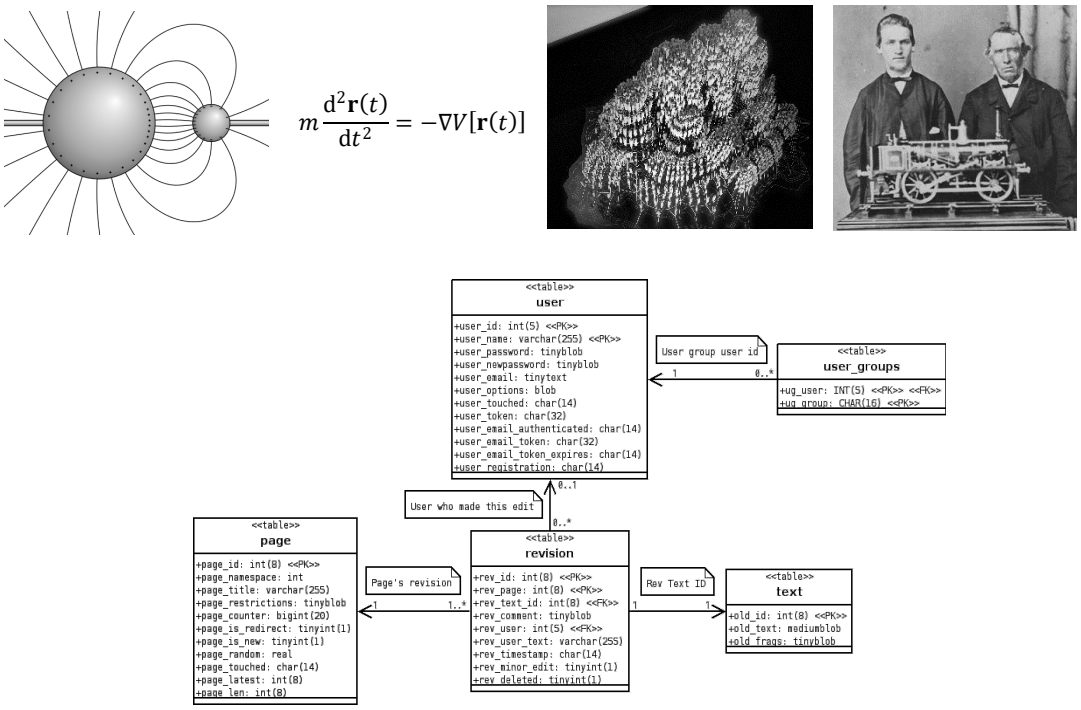


Figura 12. Varios modelos utilizados en ingeniería. (de izquierda a derecha y de arriba abajo): representaciones gráfica y formulada del campo eléctrico, modelo polifunicular de la basílica de la Sagrada Familia de Barcelona, modelo en miniatura de una locomotora, y diagrama de clases UML que representa el modelo de datos de un sistema software.

Tal como indica el título de nuestro trabajo de investigación, nuestras propuestas se adhieren al enfoque de la **ingeniería de software y sistemas basada en modelos (MBSSE)**, que proporciona una forma de *estructurar sistemáticamente y de formalizar el conocimiento en torno a un sistema y su contexto* mediante un conjunto de modelos. Varias *iniciativas* respaldan este enfoque para el desarrollo de sistemas (con sutiles matices entre sí), como la Arquitectura Dirigida por Modelos (MDA) del OMG [99], la Ingeniería de Sistemas Basada en Modelos (MBSE) de INCOSE [100] o el estándar de Software e Ingeniería de Sistemas Basado en Modelos (MBSSE) ISO/IEC/IEEE (en elaboración) [101]. Sin embargo, el uso de modelos en el desarrollo de software y sistemas trasciende a todos estos foros, ya que los modelos del software están presentes en todas partes (aunque a veces disfrazados bajo otro nombre): desde el esquema de una base de datos escrita en SQL, hasta las docenas de herramientas que introducen tipos estructurados en datos JSON o en APIs RESTful, hasta el gemelo digital de un objeto o de un proceso físico que evoluciona en tiempo real a lo largo del ciclo de vida del objeto en sí, a plataformas de baja cantidad de código (*low-code*) que aprovechan el desarrollo visual declarativo para abstraer el código subyacente que generaron, y lenguajes específicos de dominio (DSL) creados para abordar problemas en dominios o industrias específicas. Es en este sentido amplio que el uso del modelado se aborda desde nuestra investigación.

A continuación se presentan los principales conceptos que intervienen en los métodos de ingeniería basados en MBSSE, cuyas relaciones se muestran en la Figura 13.

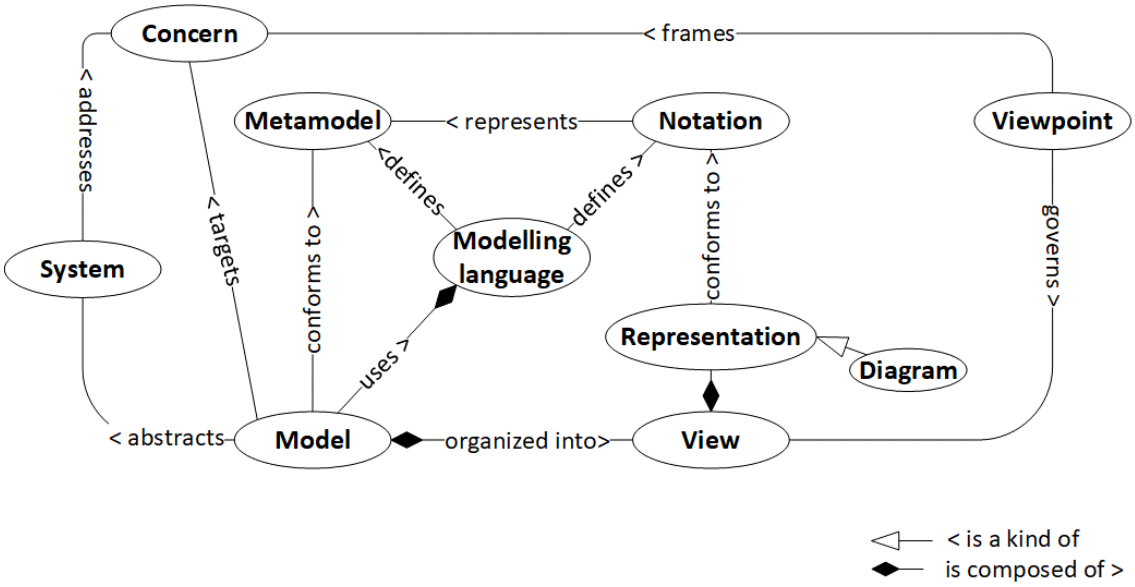


Figura 13. Mapa conceptual de la ingeniería de software y sistemas basada en modelos (MBSSE)

En MBSSE, los modelos deben expresarse mediante algún tipo de **lenguaje de modelado**, *estándar, bien definido, y compartido* por todas las personas involucradas en el proceso y las herramientas software de desarrollo que utilicen. Típicamente, estos lenguajes proporcionan *notaciones gráficas* que permiten representar el modelo (o parte), en forma de *diagramas* (aunque no se debe confundir el modelo con su representación visual). Cuando un modelo se crea siguiendo un lenguaje de modelado, se dice que dicho modelo es conforme con el *metamodelo* que especifica el lenguaje de modelado. Este metamodelo especifica los elementos del lenguaje de modelado, es decir, es el modelo del lenguaje. Por ejemplo, si estamos creando un sistema de teleducación, nuestro modelo incluirá elementos como '*profesor*', '*estudiante*' o '*asignatura*'. Estos elementos y su contenido, definidos en un lenguaje de modelado como UML, se representarán como '*clases*', '*interfaces*', '*atributos*' o '*métodos*': pues bien, estos últimos son los términos del lenguaje UML constituyen el *metamodelo de UML*.

Por otra parte, como abstracción del sistema, el modelo no pretende reproducirlo en su totalidad, sino ofrecer solo una *vista* que incluye aquellos elementos necesarios para comprenderlo desde un **punto de vista** específico y responder a un interés específico. Por ejemplo, un modelo de un edificio puede ofrecer vistas en alzado, planta, y vista lateral, e incluir solo elementos relevantes para el punto de vista de fontanería (p.ej. desagües) o electricidad (p.ej. interruptores). En el caso de un sistema software, encontraremos, p.ej. una vista *funcional* que describe los procesos de negocio, casos de uso, roles y usuarios; una vista de *arquitectura de despliegue* que detalla el sistema y sus componentes, las interfaces internas y externas, o la topología de red; una vista de *estructura de datos* que identifica los tipos y las categorías de información que se procesan, etc.

Con esto, tenemos lo necesario para responder al lugar que *la privacidad y la protección de datos* tienen en el modelado: se trata de uno de los *intereses que se pueden modelar desde su propio punto de vista*, y, por tanto, requiere dotar al lenguaje de modelado de los conceptos propios de este punto de vista, para que los modelos puedan reflejarlo. Más en concreto, dado que la privacidad y la protección de datos son aspectos transversales (pueden verse afectadas por la estructura el sistema, su dinámica, su despliegue, etc.) hablamos de la definición de un lenguaje específico de aspectos de dominio (DSAL) [102], que añade extensiones (un '*perfil*') a los lenguajes de modelado para que aquellos elementos de los modelos puedan enriquecerse (etiquetarse o '*estereotiparse*') con conceptos relevantes para la privacidad y la protección de datos.

4.2 Lenguaje específico de modelado de aspectos de para la privacidad y la protección de datos

Como se indicado, siguiendo el paradigma de MBSSE, hemos desarrollado un **lenguaje específico de aspectos de dominio (DSAL) para la privacidad y la protección de datos**, cuyos elementos presentamos en esta sección. Se trata de un lenguaje que define una serie de conceptos propios de la privacidad y protección de datos como ámbito de aplicación transversal, pensados para extender lenguajes de modelado ya existentes, de manera que se puede enriquecer el modelo de un sistema bajo estudio con información sobre aquellas características relevantes para la privacidad. Por ejemplo, el modelo de un sistema puede incluir una serie de '*clases*' o de '*tablas*' que representan distintos tipos de datos que trata el sistema: el DSAL de privacidad permite etiquetar algunas de esas clases con el '*estereotipo*' *«Personal Data»* (*dato personal*), indicando, además, si se trata de datos sensibles, qué condición hace lícito su tratamiento, etc.

Podemos agrupar los estereotipos (elementos de nuestro lenguaje) de nuestro DSAL de privacidad y protección de datos en tres grandes categorías, en función del **punto de vista de los elementos a los que se aplican**:

- Desde un *punto de vista estructural*, debemos considerar los propios *datos personales*, los *interesados* (*«Data subject»*) a los que se refieren, y las *políticas* que determinan su uso. Los datos personales van acompañados de información de *trazabilidad de procedencia* (*provenance*) que pueden explicar cuál es su origen y la base de licitud de su tratamiento (p.ej. un registro de consentimiento, un análisis de impacto de interés legítimo, etc.).
- Desde un *punto de vista comportamental*, encontramos los *procesos de las actividades de tratamiento* (*«Data processing operation»*) de datos personales, las *entidades externas* (*«External entity»*) a un sistema que producen o consumen los datos tratados, y los *flujos de datos* (*«Data flow»*) entre procesos y entre estos y las entidades.
- Desde el *punto de vista de la arquitectura de despliegue*, tenemos dos elementos básicos: los *nodos* (*«Processing node»*) en los que se ejecutan las distintas actividades de tratamiento (p.ej. servidores de aplicaciones, sistemas operativos, máquinas virtuales, hardware), y los *enlaces* (*«Communications link»*) a través de los que transitan los flujos de datos; más los *dominios* (*«Realm»*) que agrupan distintos nodos y enlaces controlados por un mismo agente, directamente (por ejemplo, bajo un mismo responsable del tratamiento) o indirectamente (por ejemplo, dentro de una misma jurisdicción).

En las páginas siguientes, la Tabla 6 lista los distintos **estereotipos** que podemos utilizar para modelar los aspectos de privacidad y protección de datos de los puntos de vista mencionados. Cada uno de los estereotipos incluidos responde a una o varias preguntas, que se presentan en la misma tabla, y pueden incluir, a su vez, atributos adicionales relativos al estereotipo (siguiendo el ejemplo del *«dato personal»*, este puede ir acompañado de información que trace el origen del que proviene). La última columna representa a qué tipo de elementos se puede aplicar cada uno de los estereotipos: así, el estereotipo «*dato personal*» se aplica a los elementos que modelen estructuras de datos (llámense '*clases*', '*tablas*', etc., en función del lenguaje de modelado del software con el que se trabaje).

Por otro lado, más allá del modelado del sistema, hay otros conceptos que también son relevantes para la privacidad y la protección de datos, y que se pueden modelar en el **resto de las disciplinas** de la ingeniería de software y sistemas (gestión de riesgos, ingeniería de requisitos, aseguramiento de sistemas, e ingeniería de métodos), como recoge la Tabla 7.

Tabla 6. Estereotipos de privacidad y protección de datos que se pueden adjuntar a los elementos de un modelo de sistema.

Estereotipo	Preguntas que aborda	Contenidos	Elementos a los que aplica
«Personal Data» (dato personal)	¿Qué datos son personales?	Categoría, incluyendo: - tipo de datos	a) Clasificador: clase, interfaz, tipo de datos, definición de estructura, tabla de entidad… b) Característica estructural: propiedad, atributo, extremo de asociación, parte, columna de tabla… d) Parámetro de operación, argumento de función… c) Datos de instancia: especificación de instancia, ranura (*slot*), especificación de valor, objeto, registro de datos, tupla, valor de celda… d) Estereotipo e) Elemento de información (en flujos de información)
	¿Son datos sensibles?	- sensibilidad - rango de posibles valores - distribución de frecuencias	
	¿De dónde provienen? ¿Por qué se pueden recoger y tartar?	Trazabilidad de procedencia ('*provenance*') de los datos, incluyendo: - origen de los datos - base legitimadora y evidencias que la respalden.	
	¿Cómo sé que son personales?	Procedencia de la categoría	
«Data subject» (interesado)	¿De quién trata este dato?	- Categoría de interesado - Dato personal	- Actor, entidad externa, usuario, parte involucrada, rol - Otros clasificadores - Instancias y estereotipos sobre cualquiera de ellos - Elementos no incluidos en el modelo
«Usage policy» (política de uso)	¿Quién está autorizado a acceder? ¿Cómo sabe que está autorizado?	Especificación de control de acceso (v. definiciones de minimización en el Capítulo 5) - Finalidad - Visibilidad - Granularidad - Retención Procedencia de las reglas de acceso	Elemento etiquetado como «personal data»

Estereotipo	Preguntas que aborda	Contenidos	Elementos a los que aplica
«Data Processing Operation» (proceso de actividad de tratamiento de datos)	¿Qué procesos tratan con cada dato? ¿A qué tratamientos se somete cada dato? ¿Para qué finalidades se usa?	- Datos personales bajo tratamiento. - Finalidad y contexto del tratamiento.	Comportamientos y características comportamentales: actividades y acciones (procesos), así como operaciones, interacciones, casos de uso…
«Data Flow» (flujo de datos)	¿Cómo viajan los datos entre los procesos? ¿Qué datos van de un proceso a otro?	- Datos personales - Actividades de tratamiento	Flujo de datos (y de control), flujo de información, arcos de actividad (arco de flujo de objetos), mensaje.
«External entity» (entidad externa)	¿Quién envía información al Sistema y la recibe de este?		Clasificadores: componentes, actors… Instancias de clasificador Entidades externas
«Data Storage» (almacén de datos)	¿Dónde y durante cuánto tiempo se almacenan los datos?	- Retención - Datos personales	a) Comportamientos b) Componentes c) Instancias de componente
«Right supporting» (respaldo a derechos)		- Data subject right - Personal data addressed	Comportamientos y características comportamentales
«Processing node» (nodo de tratamiento)	¿Dónde se lleva a cabo el tratamiento (y se recogen los datos, etc.)?	- Localización - Tratamiento de datos personales	Nodo y sus instancias
«Communications link» (enlace de Comunicaciones)		- Protocolo - Flujo de datos	a) Asociación, camino de comunicación b) Enlace
«Realm» (dominio)	¿Quién trata datos personales? ¿Bajo qué jurisdicción?	- Responsabilidad del rol - Jurisdicción	a) Conjuntos de elementos b) Componente, subsistema, nodo c) Instancias de cualquiera de ellos
Common attributes (atributos comunes)	¿Cómo se protegen los datos?	Medidas y características de protección	—

Tabla 7. Otros puntos de vista del modelado relevantes para la privacidad y la protección de datos.

Punto de vista del modelado	Contenidos	Preguntas que aborda
Modelo de requisitos	- Plantillas / marcos de requisitos de privacidad y protección de datos - Fuentes de los requisitos - Trazas a requisitos funcionales y elementos del sistema	¿Qué características debe implementar el sistema? ¿Por qué?
Modelo de riesgos	Activos Amenazas Factores de riesgo	¿Qué necesitamos proteger? ¿De qué necesitamos protegernos? ¿Cómo de probable y dañino podría ser?
	Controles (mecanismos de protección)	¿Cómo podemos (nosotros y los demás) protegernos de esos riesgos?
Modelo de aseguramiento	Marco regulatorio de referencia Evidencias Argumentaciones	¿Con qué debo cumplir? ¿Cómo puedo garantizar y demostrar que cumplo?
Modelo de métodos	Procesos de desarrollo, personas involucradas, y productos (artefactos) Recursos de desarrollo: marco normativo de privacidad y protección de datos, bases de conocimiento, etc.	¿Cómo organizo el desarrollo de un sistema que cumpla con la privacidad y protección de datos? ¿De qué recursos dispongo?

5 Minimización de datos y métricas de riesgos de privacidad y protección de datos

En este capítulo, presentamos nuestras contribuciones en relación con la disciplina de gestión de riesgos de privacidad y protección de datos. Las Evaluaciones de Impacto de Protección de Datos (EIPD, o DPIA en inglés) determinan que el responsable del tratamiento debe tomar medidas en caso de que los riesgos para los interesados sean altos. Ahora bien, cabe preguntarse "*¿cuándo es alto ese riesgo?".* Siguiendo el lema de que "*no se puede gestionar lo que no se puede medir*", podemos decir que un *objetivo clave en la gestión de riesgos debe ser su medición*. En la privacidad y la protección de datos, estos riesgos se deben referir a los **riesgos para el interesado** (no para la organización), por lo que consideramos que un *buen indicador del riesgo es la cantidad de información de los interesados* que maneja la organización responsable de su tratamiento, ya que, cuantos menos datos haya, menor será el riesgo de que se usen maliciosamente.

En particular, recurrimos al **principio de 'minimización de datos'**, que se considera una piedra angular en la privacidad y la protección de datos, y que puede definirse informalmente como la utilización de tan pocos datos personales como sea estrictamente necesario para una finalidad determinada. Aunque el propio nombre de '*minimización'* nos induce a pensar en una magnitud que se puede medir (o, al menos, se puede comparar), en la práctica, bajo este concepto se entremezclan otros como la protección de datos por defecto, la limitación en la recogida de datos, la limitación del plazo de almacenamiento, la limitación de uso, retención y comunicación, la limitación del propósito, la especificación del propósito, la protección de datos del usuario o incluso el seudonimato, el anonimato, la desvinculabilidad, la no observabilidad, la indetectabilidad, etc. Esto nos ha llevado a revisitar la literatura para crear una teoría que abarque los diferentes significados que se encuentran bajo los diferentes principios y enfoques relacionados con la minimización de datos.

5.1 Método de análisis cuantitativo de contenidos

Para determinar qué se entiende por 'minimización' en las distintas fuentes, hemos empleado un *método de investigación* denominado *análisis cualitativo de contenidos (QCA),* mediante el que se extraen los significados que están contenidos dentro de una o más fuentes textuales y se los asigna a un conjunto limitado de categorías. Aunque este método proviene de las ciencias sociales y humanidades, se ha utilizado también en ocasiones en la ingeniería de requisitos de software.

El QCA toma una serie de textos como entrada y recorre su contenido marcando sistemáticamente ('codificando') el significado de cada parte del texto para la identificación de temas o patrones. Los principales pasos de este método son [103]–[108], [109, Ch. 11], [110]:

1. Preparar los datos, es decir, crear el conjunto de textos.
2. Definir la unidad de análisis, es decir, los ejemplares que clasificar mediante códigos.
3. Desarrollar categorías y un esquema de codificación.
4. Probar el esquema de codificación con parte del texto.
5. Codificar el texto.
6. Extraer conclusiones de los datos codificados.
7. Informar los métodos y conclusiones.

El proceso se lleva a cabo de manera iterativa, y se apoya en el uso de herramientas tecnológicas y metodológicas para marcar y desarrollar las categorías. Dentro de las

distintas variantes de QCA que existe, nosotros aplicamos en particular el método comparativo constante.

Hemos aplicado el método QCA para la definición de las *dimensiones en las que se puede descomponer el concepto de 'minimización de datos'* (otros relacionados con él), a partir de las siguientes *fuentes*:

1. Las estrategias y tácticas de privacidad de Hoepman [111]–[113] incluidas en la norma técnica ISO/IEC 27550 [114].
2. '*A terminology for talking about privacy by data minimization*' (de Pfitzmann y Hansen) [58], y su evolución en la recomendación estándar de Internet (IETF) RFC 6973 [115] [116].
3. Los principios de privacidad de la sección 5 de la norma técnica ISO/IEC 29100 [60, Sec. 5].
4. Aquellos contenidos de los Criterios Comunes [117] (como evolución de la norma ISO 15408) que estén relacionados con la privacidad [118].
5. El propio RGPD [67].
6. La definición seminal de Cavoukian de los principios de Privacidad por Diseño [70] y su operativización [119], [120]
7. '*Engineering Privacy by Design Reloaded*' (de Gürses y Troncoso) [88].

5.2 Definición multidimensional de la minimización en la privacidad y la protección de datos

Como resultado de aplicar el método, hemos sintetizado *25 preguntas* que identifican las dimensiones a lo largo de las cuales se pueden '*minimizar*' los datos personales. La Figura 14 presenta esas 25 dimensiones, seguida de una breve explicación de cada una de ellas (el detalle de los resultados en relación con cada una de las fuentes y la explicación detallada del significado de cada dimensión se recoge en la memoria de la tesis doctoral del autor en la que se enmarca este trabajo de investigación [1, Sec. 5.1.2])

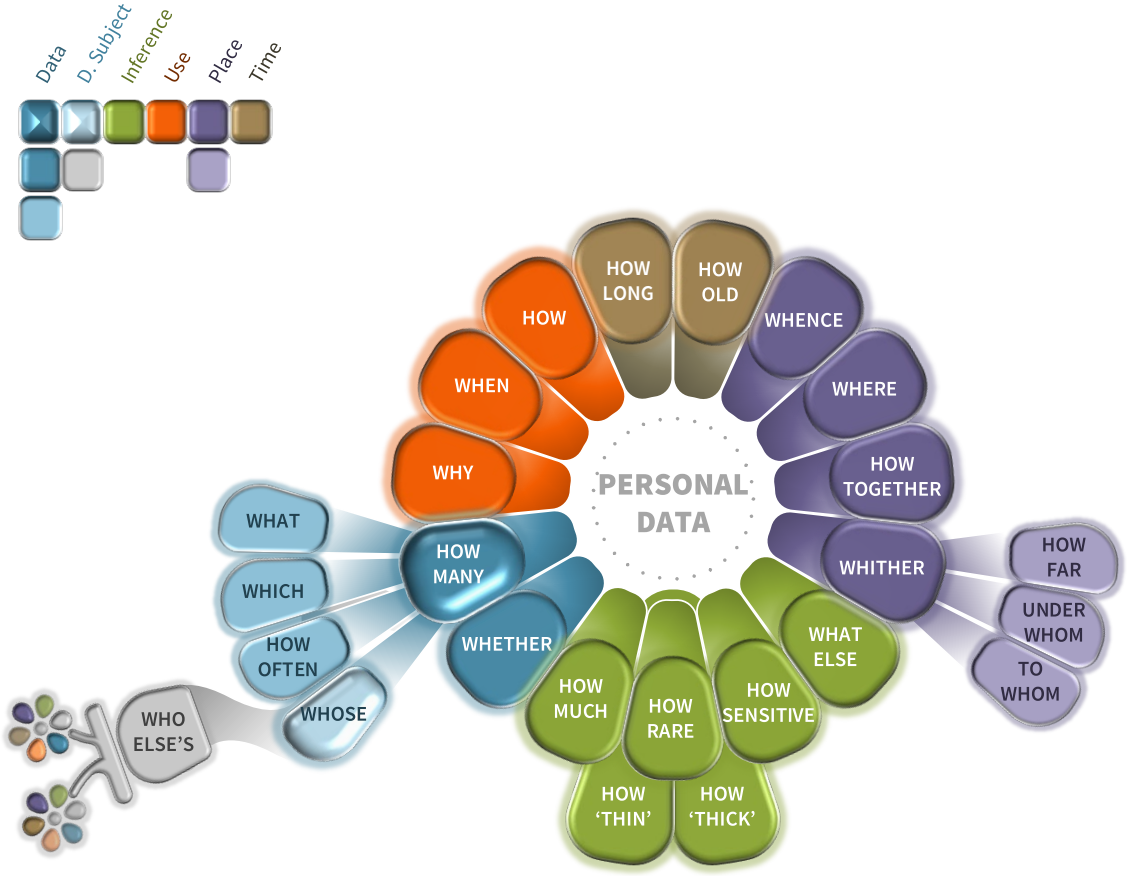


Figura 14. Dimensiones de la minimización de datos personales.

- *¿Cuántos datos?* Cantidad de datos personales recopilados, almacenados, utilizados o procesados.
    - *¿Qué datos?* Número de atributos, características o tipos de datos personales de datos personales (recopilados, almacenados, utilizados o procesados).
    - *¿Cuáles?* Número de registros de datos o muestras creadas de un tipo de atributo dado.
    - *¿Con qué frecuencia?* Frecuencia de captura (o recogida) de datos de un atributo dado.
    - *¿Sobre quién?* Número mínimo de personas sobre las que se recopilan, utilizan o procesan datos.
- *¿Por qué?* Finalidades para las que se utilizan los datos.
- *¿Cuándo?* Eventos de recogida de datos o transacciones que la requieran.
- *¿Cómo?* Actividades de tratamiento a las que se someten los datos.
- *¿Cuánto tiempo?* Cantidad de tiempo que se retienen los datos (permanecen en el sistema).
- *¿Qué antigüedad?* Obsolescencia de los datos personales tratados.
- *¿De dónde?* Fuentes de las que se obtienen los datos.
- *¿En dónde?* Entidades donde se almacena y procesa un dato.
- *¿Cómo de juntos?* Cantidad de datos procesados dentro de una entidad determinada.
- *¿Adónde?* Destinos potenciales o reales de los datos personales.
    - *¿A quién?* Destinatarios (entidades, empleados, etc.) a las que fluyen o se transfieren los datos.
    - *¿Bajo quién?* Jurisdicciones aplicables a los datos en tránsito o en reposo.
    - *¿Hasta dónde?* Distancia física desde donde se puede acceder a los datos.
- *¿Qué más?* Datos potencialmente derivados o inferidos.
- *¿Qué sensibilidad?* Sensibilidad de los datos tratados.
- *¿Cómo de raro?* Entropía de información de los datos personales tratados.
- *¿Cuánto?* Daño que puede ser causado (y compensado) por un tratamiento ilícito de los datos.
- *¿Cómo de 'fino'?* Precisión (granularidad) de los datos.
- *¿Cómo de 'grueso'?* Valor de los datos en torno a áreas semánticamente cercanas.
- *¿Es posible?* Posibilidad mínima de recopilación de datos.
- *¿Quién más?* Otros interesados cuyos datos personales se divulgan secundariamente como consecuencia de hacerlo con los del interesado principal.

5.3 Métricas de ámbito organizativo, orientadas a riesgos, y centradas en los sujetos interesados, para la privacidad y la protección de datos

Siguiendo la idea de usar la minimización de datos como sustituto aproximado para medir los riesgos de privacidad y protección de datos, en esta sección, ofrecemos un enfoque práctico para calcular métricas de riesgos que aplican varias de las dimensiones presentadas. En particular, recurrimos al concepto de la **entropía de la información** para capturar la cantidad de 'nueva información' que proporciona cada nuevo dato tratado. Así, proponemos las siguientes *métricas para distintos riesgos* asociados a la información personal (la presentación detallada y el razonamiento empleado para su obtención están recogidas en memoria de la tesis doctoral del autor [1, Sec. 5.2.1]).

Dada una fuente X de datos personales que puede tomar valores de un conjunto $\mathcal{X} = \{x_i\}$ con distinta probabilidad $P(x_i) = p_i$, y un conjunto (*dataset*) $\hat{X}$ de muestras de ese tipo datos X, con una frecuencia relativa para cada valor $f(x_i) = f_i$, definimos:

(1) El *riesgo de información de la variable* X para el dataset $\hat{X}$ como la entropía cruzada desde X a $\hat{X}$, cuyo valor es $\boldsymbol{H_{\hat{X}}}(X) = \sum_x -f_i \cdot \log_2 p_i$

Si ahora contamos con un conjunto de fuentes $\{X_{j=1,\dots,n}\}$ de distintos tipos de datos personales, que pueden tomar respectivamente valores $x_{i_{(j)}} \in \mathcal{X}_j$ con probabilidad conjunta $P\left(x_{i_{(1)}}, \dots, x_{i_{(n)}}\right)$, definimos:

(2) El *riesgo de información del dataset* asociado $\mathfrak{X}$, que contiene r_j registros generados con frecuencia v_j para m_j individuos como $H(\mathfrak{X}) \approx \sum_{X_j \in \mathfrak{X}} m_j^2 \cdot H_{\widehat{X_j}}(X_j)^2 \cdot v_j \cdot r_j$, donde los términos m_j y $H_{\widehat{X_j}}(X_j)$ aparecen repetidos porque reflejan tanto un mayor impacto como un mayor incentivo a ser atacados.

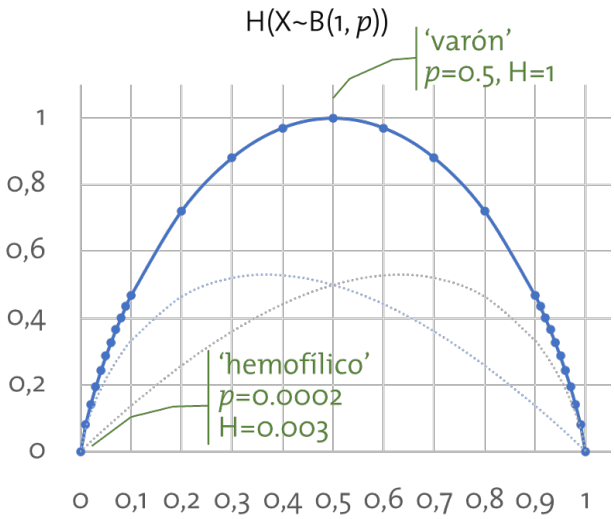
(3) El *riesgo de identificabilidad* de las fuentes asociadas a las variables $X_{1,\dots,n}$ como la entropía conjunta de todas ellas $\boldsymbol{H^I}(\{X_{1,\dots,n}\}) \triangleq H(X_1, \dots, X_n) = \sum_x -P(x_{1,i}, \dots, x_{n,i}) \cdot \log_2 P(x_{1,i}, \dots, x_{n,i})$

(4) El *riesgo de enlazabilidad* de las fuentes asociadas a las variables $X_{1,\dots,n}$ como la entropía conjunta que tendrían si fueran fuentes de datos independientes $H(X_1 \otimes \dots \otimes X_n)$ (donde $\otimes$ representa la creación de una nueva variable cuya distribución conjunta es el producto de las marginales), que equivale a la suma de las entropías individuales $H(X_1) + \cdots + H(X_n)$ y que vale $\boldsymbol{H^L}(\{X_{1,\dots,n}\}) \triangleq H(X_1 \otimes \dots \otimes X_n) = \sum_j \sum_x -P(x_{j,i}) \cdot \log_2 P(x_{j,i})$

(5) El *riesgo de revelación* de las variables $X_{1,\dots,n}$ almacenadas en distintas fuentes X^S como la media ponderada entre entropía $H(X^S)$ y perplejidad $2^{H(X^S)}$ con el valor $\boldsymbol{H^S} = w \sum_S H(X^S) + (1-w) \log_2 \sum_S 2^{H(X^S)}$

Cabe recordar que nuestro objetivo aquí es disponer métricas de riesgos orientadas a las Evaluaciones de Impacto que debe realizar una organización (pueden existir otras métricas de privacidad y protección de datos desde otras perspectivas).

A continuación, proponemos la aplicación del método PMRM de particionado de riesgos de objetivos múltiples [121], [122] para abordar los **riesgos asociados a datos sensibles**, mitigando un problema en el que caen habitualmente los enfoques puramente basados en la entropía, cuando se trata con estos datos sensibles. En concreto, sucede que cuando encontramos un tipo de dato personal sensible, este conlleva asociados riesgos extremos (es decir, tienen asociado un gran daño, pero son de baja probabilidad), especialmente cuando solo algunos valores poco frecuentes del atributo son altamente sensibles (pensemos, por ejemplo, en un marcador genético de una enfermedad rara, o en la adscripción a una ideología minoritaria y perseguida por un régimen dictatorial). Sin embargo, al aplicar la fórmula tradicional riesgo = probabilidad × impacto, los altos impactos asociados a estos eventos se diluyen en su baja probabilidad, es decir, hay tan pocos individuos que pueden verse afectados por el riesgo, que parece que no valiera la pena protegerles. La Figura 15 ilustra el problema, donde parece que, contra lo que cabría intuir, el riesgo asociado a conocer que una persona tiene sexo biológico de 'varón' es menor que el asociado a conocer que 'padece hemofilia'



$$H(X) = \sum_x -p_i \cdot \log_2 p_i$$

Figura 15. Ilustración del comportamiento contraintuitivo de la valoración del riesgo de privacidad basada en la entropía de la información.

Este problema al que nos enfrentamos con los atributos sensibles es bien conocido en la gestión de riesgos bajo el nombre de '*falacia del valor esperado*': al incluir en el valor del riesgo la esperanza matemática de riesgos normales y riesgos extremos, acabamos dando más peso a una situación frecuente con bajo impacto que a un evento raro con alto impacto, y se fomentan aquellas medidas que controlan los riesgos comunes frente a las que reducen los riesgos extremos. El problema se debe a que estamos ponderando conjuntamente dos cantidades que en realidad son inconmensurables (expresan magnitudes que no se pueden sumar): los riesgos normales y los riesgos catastróficos, que en realidad obedecen a realidades distintas.

La solución que proponemos es utilizar el *Método de particionado de riesgos de objetivos múltiples (PMRM)*, que trata a los riesgos moderados y a los riesgos extremos como dos objetivos independientes que se deben optimizar por separado. De esta manera, se define una nueva métrica de riesgo que pondera solamente aquellos riesgos que se encuentren en el extremo superior de la función de distribución acumulada, es decir, que cuenten con una probabilidad de rebasamiento[6] suficiente (Figura 16).

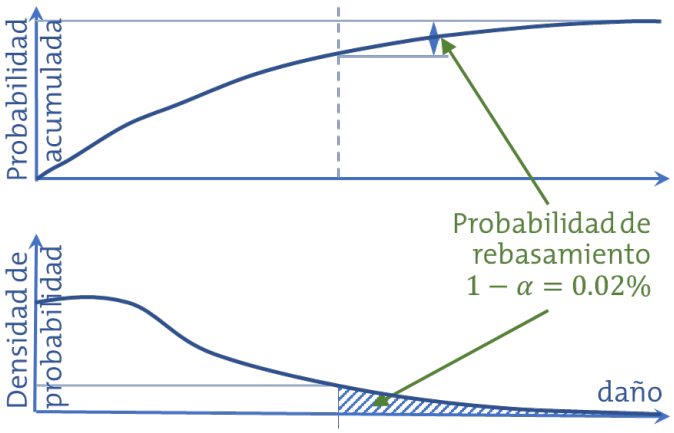


Figura 16. Análisis diferenciado de los riesgos extremos en el método PMRM

[6] El nombre 'probabilidad de rebasamiento' se refiere a que el daño producido por un riesgo podrá rebasar un umbral de daños dado con dicha probabilidad, pero a su vez desvela que el origen de este método se encuentra en la ingeniería hidráulica, donde el rebasamiento se produce físicamente sobre una presa o dique.

En nuestro caso, además, la función de daño (impacto) a la que se aplica PMRM, corresponde a un valor de la '*sorpresividad'* (cantidad de información asociada a cada dato) ajustada en un peso *w* función de la sensibilidad del valor específico del dato personal: $S(x) = -w\ x \cdot \log_2 p_X(x)$

El resultado es que tenemos una nueva medida de los riesgos extremos, cuyo valor viene dado por

$$H^*(X) = \sum_{\log_2 p_i < \frac{-F_S^{-1}(\alpha)}{w_i}} -\frac{p_i}{1-\alpha} \cdot w_i \log_2 p_i$$

Finalmente, proponemos la **combinación de mediciones de riesgo en un método de decisión** a través del *Proceso de Análisis Jerárquico (AHP)* [123], [124], para integrar varias métricas de riesgo independientes en el proceso de elección entre distintos proveedores de servicios, lo que requiere:

i) La definición de un conjunto de criterios que representan distintos objetivos que se persiguen (en nuestro caso, la minimización de los distintos tipos de riesgo de información, enlazabilidad, identificabilidad, revelación, etc., así como los riesgos extremos recién presentados).
ii) Opcionalmente, la descomposición sucesiva de cada uno de esos criterios en subcriterios, para los que se aplica este mismo método iterativamente.
iii) Una serie de juicios de valor en cuanto a la importancia relativa entre cada par de criterios bajo consideración. Los juicios de valor pueden ser cuantitativos o cualitativos (p.ej. "es algo más importante que"). En nuestro caso, podemos plantearlo como las preferencias en cuanto a la privacidad de distintos consumidores de servicios (p.ej. responsables del tratamiento que contratan a encargados del tratamiento). Distintos consumidores de servicios pueden tener distintas preferencias y, por tanto, asignar distintos pesos a cada criterio.
iv) La definición de una serie de alternativas disponibles para abordar el problema. En nuestro caso, se trata de los distintos proveedores de servicios (encargados del tratamiento).
v) Para cada criterio, una serie de juicios de valor en cuanto a la bondad relativa de cada una de las alternativas disponibles (p.ej. para el criterio 1, la alternativa A se comporta "mucho peor que" la alternativa B).

A partir de ahí, mediante una serie de operaciones con matrices, el AHP emite:

(1) Un vector con la importancia global de cada criterio, evaluada numéricamente.
(2) Para cada criterio, un vector con el comportamiento de cada alternativa.
(3) Un vector con el comportamiento global de cada alternativa.

Como resultado, permite emitir un ranking de las alternativas preferibles. En concreto, la Figura 17 muestra el resultado de aplicar el AHP a la elección entre proveedores de servicios en la nube (PSNs) en función de los parámetros de privacidad que emiten en un acuerdo de nivel de servicio de privacidad (PLA), por un lado, en relación con criterios individuales y, por el otro, en relación con preferencias agregadas de cada consumidor de servicios (CSNs) [2], [125]:

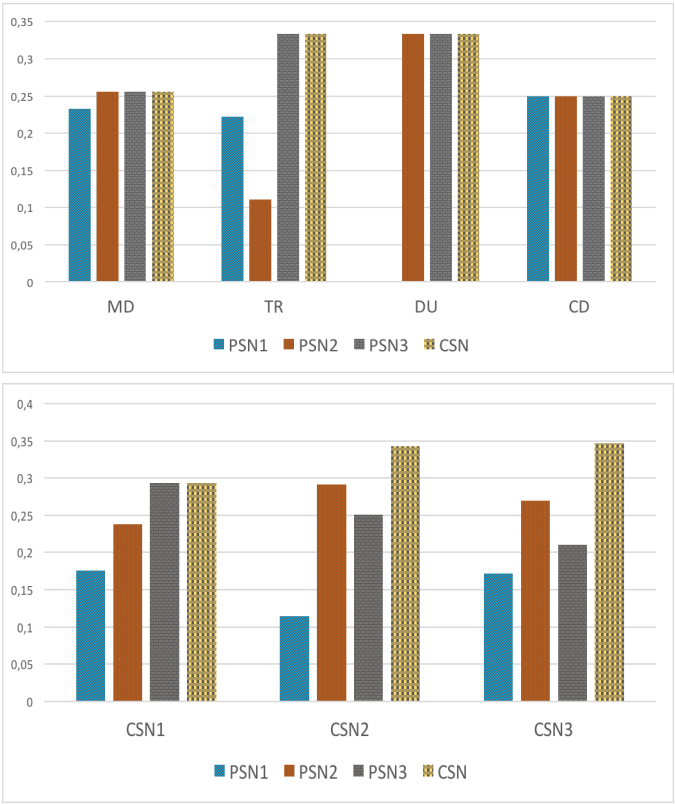


Figura 17. Resultado de la aplicación del AHP para comparar las ofertas de servicio de diferentes proveedores de servicios en la nube respecto a varios criterios de privacidad, y adecuación a los requisitos de los diferentes consumidores de servicios en la nube.

6 Ingeniería de requisitos de privacidad y protección de datos

Como se ha comentado anteriormente, la privacidad y la protección de datos son problemas sociotécnicos que deben responder a múltiples intereses de distintos interesados (legales, de los usuarios, comerciales). Para construir un sistema que responda de manera óptima a esos intereses, es necesario disponer de una especificación detallada de los requisitos que debe cumplir; sin embargo, las necesidades de privacidad y protección de datos se expresan a un nivel muy alejado de la tecnología que entienden los ingenieros y además, como requisitos extrafuncionales, se formulan de manera inherentemente independiente del proyecto y del sistema (p.ej. "*cifrar los datos en tránsito*" significa poco si no especificamos cuáles son esos datos y por dónde transitan), lo que dificulta aún más su capacidad de gestión. Este capítulo intenta ofrecer nuestra respuesta a estas cuestiones, mostrando cómo se puede integrar los requisitos de privacidad y protección de datos en el proceso de desarrollo.

6.1 La operativización como método para derivar requisitos de fuentes estándar

Nuestro enfoque para la derivación de requisitos operativos de privacidad y protección de datos bebe de cómo se han tratado los requisitos en otras categorías de requisitos extrafuncionales, en concreto, *cómo se abordan la operativización de los requisitos de accesibilidad*[7] mediante su descomposición jerárquica en plantillas de requisitos sucesivamente más refinados, desde los principios, a las pautas, a los criterios de validación y, finalmente, a técnicas opcionales para su cumplimiento; que terminan por instanciarse en cada proyecto específico, sustituyendo los términos correspondientes de cada requisito-plantilla por los específicos de cada sistema.

En nuestro caso, para sistematizar la operativización de requisitos de privacidad, revisitamos ese enfoque desde el *punto de vista de la ingeniería de la privacidad y protección de datos*, ilustrando cómo puede funcionar este enfoque de **operativización jerárquica** en este campo. En concreto, presentamos los principios de privacidad y protección de datos como nivel más abstracto de requisitos, los roles involucrados en los requisitos de privacidad y protección de datos, y una posible organización por capas de los requisitos de privacidad y protección de datos, utilizando varias fuentes como ejemplo, así como el papel de los niveles de conformidad en los requisitos de privacidad y protección de datos.

[7] La accesibilidad es la posibilidad de que todas las personas puedan acceder a los servicios y productos con independencia de si tienen alguna discapacidad. Aunque pueda parecer una categoría lejana a la privacidad, ambas establecen requisitos extrafuncionales que derivan de Derechos Humanos (y que, como tales, se recogen en leyes, manuales de responsabilidad corporativa, etc.), y requieren de técnicas de diseño y métodos de desarrollo apropiados para abordarlas. El autor ha aprovechado su conocimiento previo para trasladar las mejores prácticas aplicables del área de la accesibilidad a la privacidad y la protección de datos.

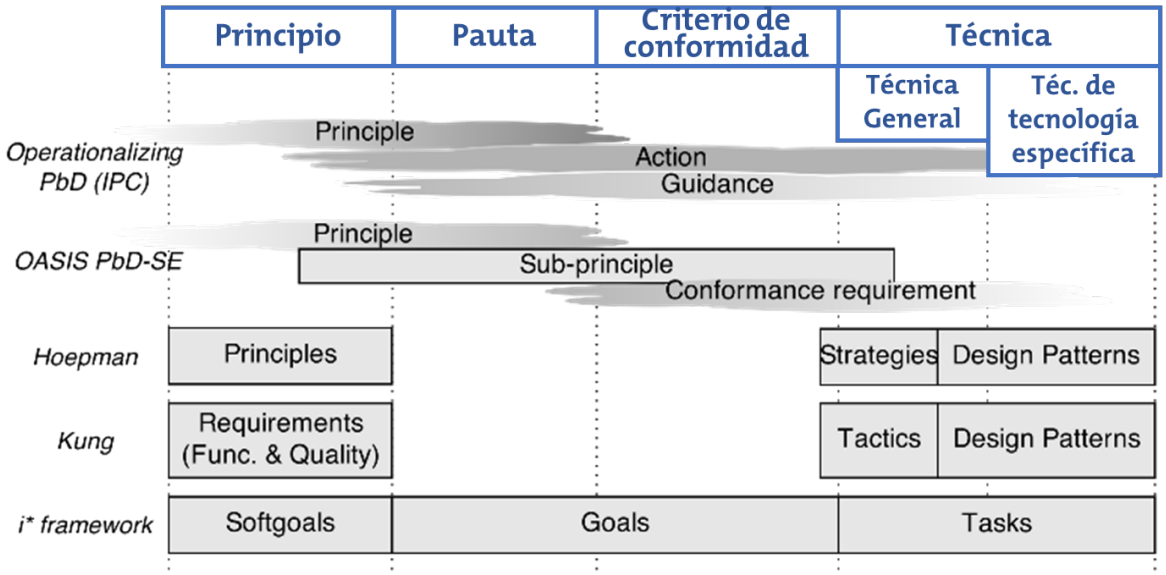


Figura 18. Operativización de requisitos de privacidad mediante su descomposición jerárquica y su correspondencia con distintas fuentes de requisitos.

6.2 Enfoque basado en riesgos y enfoque orientado a metas para el análisis de requisitos de privacidad y protección de datos

Consideramos que la especificación de requisitos de privacidad y protección de datos requiere el **papel complementario de un enfoque orientado a metas** (como el que acabamos de presentar en la sección anterior) **con enfoques basados en el riesgo**, según se refleja en Figura 19. Aunque ambos enfoques tratan de ofrecer una comprensión sobre lo que debe hacer el sistema para atenerse a los principios de privacidad, difieren en cómo abordan esta tarea: el primero se centra en identificar una serie de características que se deben construir, mientras que el segundo se fija en identificar una serie de problemas que deben evitarse.

Nuestra investigación postula que **ambos enfoques son imprescindibles** y *rebate prolijamente los trabajos preexistentes* que pretenden excluir uno de ellos [1, Sec. 6.3]. Por un lado, se dice del RGPD que es una norma 'basada en riesgos', ya que su cumplimiento implica analizar el potencial impacto negativo para el interesado de un tratamiento de datos, y tomar proactivamente medidas en función de dicho impacto [126]–[128]. Pero, por otro, lado esta concepción de los riesgos ha infectado a trabajos sobre privacidad que intentan ver bajo ese prisma lo que debería abordarse como metas.

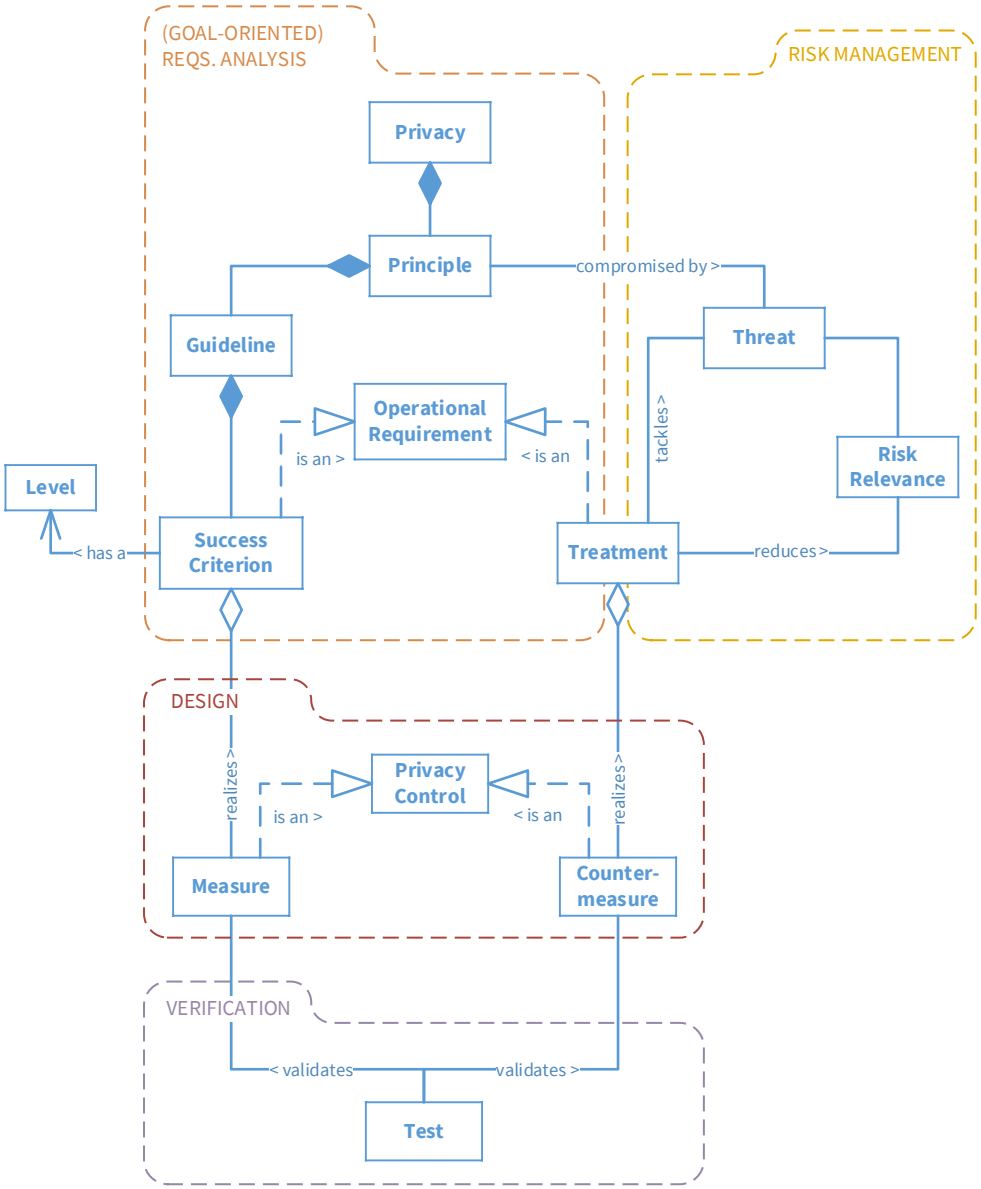


Figura 19. Enfoques de análisis de requisitos en el ciclo de vida de desarrollo: orientado a metas y basado en riesgos.

Un ejemplo: en ocasiones se intentan formular los requisitos de transparencia desde un enfoque basado en riesgos, tales como el *«riesgo de no ofrecer públicamente una política de privacidad clara»* [129], [130]. Sin embargo, en un análisis de riesgos de privacidad y protección de datos, estos deben referirse a riesgos para los interesados (p.ej. usuarios del servicio), para los que existe una incertidumbre que va más allá de las acciones que la organización pueda tomar con su mejor intención (por ejemplo, que alguien no autorizado acceda en el futuro, o intente inferir nuevos datos, etc.) Por el contrario, una organización puede cumplir o incumplir un requisito de transparencia como el del ejemplo, más allá de ninguna incertidumbre que escape a su control. Por otra parte, estos riesgos para los interesados son distintos a los riesgos que podrían afectar a la organización (p.ej. riesgo de que sea multada por incumplir el RGPD). Si, por el contrario, intentamos abusar del análisis de riesgos y aplicarlo a ejemplos como el susodicho '*riesgo*' de no publicar una política, la organización responsable del tratamiento podría concluir que «*no merece la pena poner ninguna política, porque hay poco riesgo de que nos atrapen*», o «*porque obtenemos más beneficio si ocultamos la política*», etc. lo que, a todas luces, va contra el espíritu de la protección de los derechos de los interesados.

6.3 Operativización de los requisitos del RGPD a través del modelado de la norma ISO/IEC 29100 y la correspondencia entre ambos

Hemos aplicado el enfoque de operativización presentado mediante el refinamiento jerárquico de la *especificación de los requisitos de privacidad de un estándar de la industria.* Nuestro enfoque para modelar **requisitos derivados de ISO/IEC 29100** parte del papel complementario que entendemos que deben desempeñar la regulación legal (RGPD) y los estándares técnicos (ISO/IEC 29100) en la especificación de requisitos. No se trata de que estos vengan a sustituir a aquella, sino de que la complementen mediante requisitos descritos en un lenguaje más cercano a la técnica, y que se pueda establecer una trazabilidad desde el estándar técnico a (determinadas partes de) el RGPD.

Para ello, aplicamos el *método de análisis cualitativo de contenidos QCA* (presentado anteriormente en el capítulo 5), para llevar a cabo un análisis del texto de la norma ISO/IEC 29100, cuyo **libro de códigos** para la extracción de conceptos y relaciones se resume a continuación en la Figura 20:

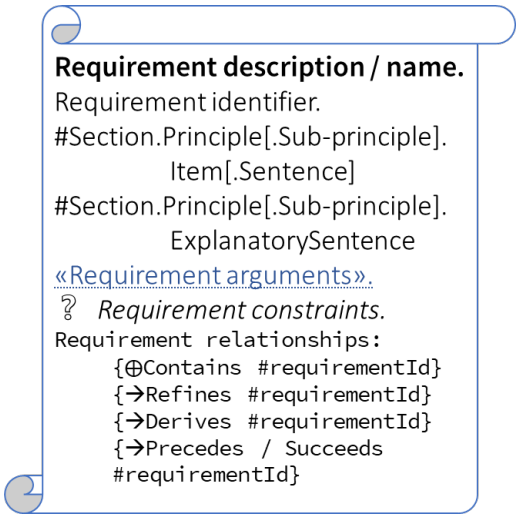


Figura 20. Leyenda del libro de códigos para la extracción mediante QCA de requisitos a partir de la norma ISO/IEC 29100.

La Tabla 8 y la Figura 21 muestran a continuación un **extracto del resultado** de aplicar el método al contenido de dicha norma *ISO/IEC 29100* (en concreto, la parte relativa al principio de consentimiento y elección), cuyo desarrollo en detalle se pueden consultar en la memoria de la tesis doctoral del autor en la que se enmarca esta investigación [1, Sec. 6.4].

Tabla 8. (Página siguiente) Modelado del principio de consentimiento y elección de la ISO/IEC 29100 (resaltados los argumentos de cada requisito que se deben reemplazar al instanciarlos; texto original en inglés de la fuente analizada).

#5.2	**5.2 Consent and choice** **{⊕contains #5.2.I, II}**
#5.2.I	Adhering to the consent principle means: {⊕contains #5.2.I.1, 2, 3, 4, 5}
#5.2.I.1	[«The organization» shall] presenting to the «PII principal» the choice whether or not to allow the «processing» of their «PII» ? except where the «PII principal» cannot freely withhold consent or ? [except] where «applicable law» specifically allows the «processing» of «PII» without the «natural person»'s consent.
#5.2.I.1.i	The «PII principal»'s choice {→refines #5.2.I.1} must be given freely, specific and on a knowledgeable basis
#5.2.I.2	[«The organization» shall] obtaining the opt-in consent {→derives #5.2.I.1} of the «PII principal» for «collecting or otherwise processing» «sensitive PII» ? except where «applicable law» allows the processing of «sensitive PII» without the «natural person»'s consent;
#5.2.I.3	informing «PII principals», before obtaining consent {→precedes #5.2.I.1}, about their rights under the individual participation and access principle {→refines #5.9};

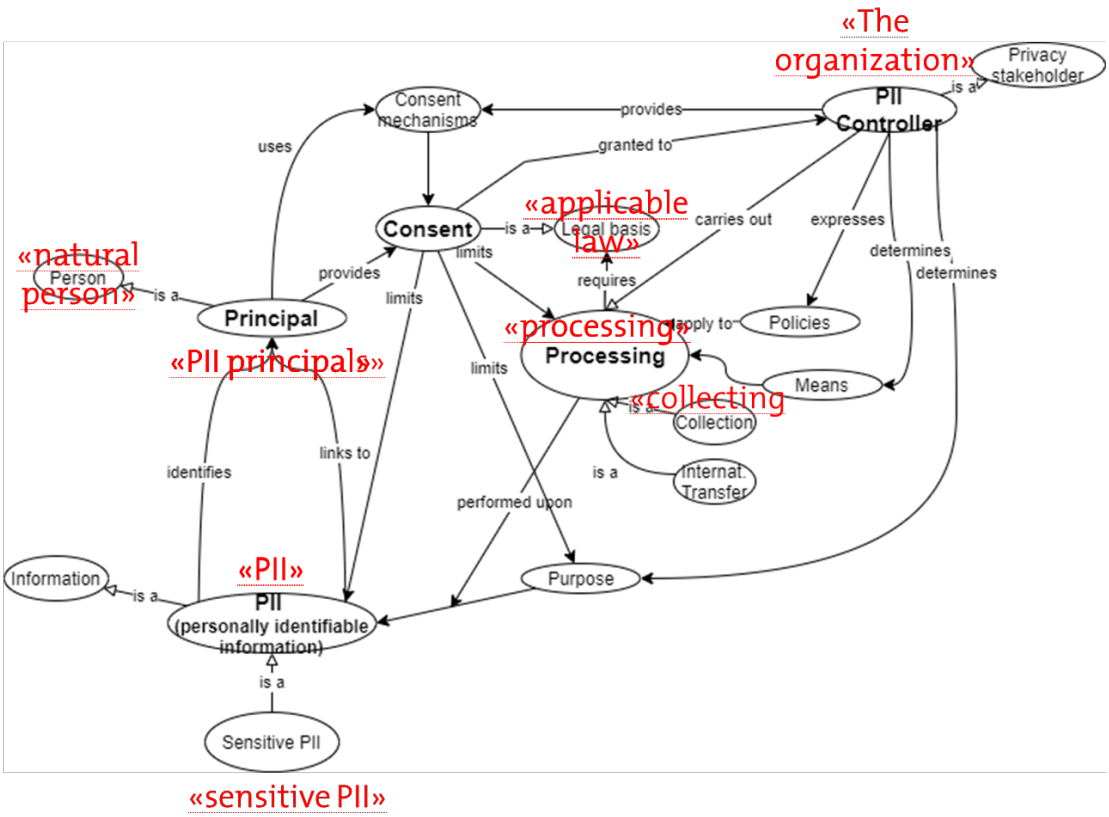


Figura 21. Modelo conceptual de los elementos de la norma ISO/IEC 29100 relacionados con el consentimiento y la elección, se han añadido resaltados en el modelo los argumentos de la tabla anterior.

7 Lenguaje de patrones de diseño para la privacidad y protección de datos

Este capítulo aborda la cuestión de *cómo los ingenieros pueden diseñar sistemas respetuosos con la privacidad y la protección de datos*, ofreciendo como respuesta un 'lenguaje de patrones', es decir, un *conjunto de plantillas de diseño que ofrecen soluciones probadas a problemas recurrentes en contextos habituales* que los ingenieros pueden emplear de manera sistemática sin necesidad de un conocimiento experto en el área de la privacidad y la protección de datos, así como las relaciones entre distintas soluciones.

7.1 Patrones de diseño: motivación, conceptos y método

Para que un sistema cumpla con los requisitos de privacidad y protección de datos (por ejemplo, los establecidos por el RGPD), hay que definir un diseño que los satisfaga. Sin embargo, **traducir los requisitos de privacidad en un diseño** no es para nada algo obvio, ya que implica:

1) explorar el espacio de las posibles soluciones de diseño y, para cada una de ellas,
2) verificar si cumple con los diferentes requisitos en cuestión (la funcionalidad del propio sistema, junto con los requisitos de privacidad y otras categorías de requisitos extrafuncionales),
3) evaluar si el diseño se adapta al entorno (organización, arquitectura, etc.), y
4) analizar cómo impacta en otras restricciones de dominio (tecnología empleada, correcciones disponibles, experticia requerida, etc.)

Sin embargo, cuando un desarrollador aborda la disciplina del diseño de software, en general no necesita dar por sí mismo con la mejor solución de diseño cada vez que se enfrenta a la creación de un nuevo sistema (o a la modificación de uno existente). De ser así, los costes de desarrollo se acabarían disparando, o los ingenieros simplemente acabarían por descartar y desdeñar muchos de estos requisitos. En cambio, los diseñadores de sistemas emplean **soluciones de diseño reutilizables** *que guían el diseño mediante recetas probadas, basadas en la experiencia y el conocimiento acumulado de toda la profesión*, reduciendo la incertidumbre y el coste del proceso de diseño. Aunque cada sistema siempre presenta requisitos y restricciones específicas que requerirán particularizar las soluciones para cada proyecto, siempre habrá aspectos comunes que se pueden reutilizar, *como una especie de plantilla*, en todos aquellos escenarios que comparten una serie de características. Así, con frecuencia, los diseñadores recurren a compilaciones o catálogos creados por diferentes fuentes (foros de estandarización, trabajos académicos, etc.) que recogen y estructuran las *mejores prácticas de la comunidad*, para utilizarlas como un recetario donde pueden encontrar aquellas soluciones más adecuadas para cada necesidad y escenario.

En el caso de la privacidad y la protección de datos, aunque existen abundantes soluciones técnicas para cumplir con los diferentes requisitos aplicables, los diseñadores carecen de una herramienta de este tipo que les permita seleccionar sistemáticamente las más adecuadas para el desarrollo de sistemas respetuosos con la privacidad y protección de datos, en función del problema al que se enfrentan, los requisitos en cuestión y el contexto de aplicación. (De hecho, quizás esta sea una de las causas por las que, como sugeríamos en el párrafo anterior, se desdeña a menudo el diseño de soluciones que cumplan con estos requisitos.) Por ello, en este capítulo, tratamos de llenar este vacío a través de las distintas contribuciones que abordan la creación de un lenguaje para patrones de privacidad.

Un **patrón de diseño** es una solución general y reutilizable a un problema que aparece de forma recurrente, y que se puede emplear cada vez que aparece ese problema en un contexto de aplicación especificado, sin necesidad de tener reinventar la solución una y

otra vez. Estos patrones se organizan en forma de **lenguajes de patrones** [131] ap. Coplien, [131]: un catálogo de patrones expresados en un formato uniforme, junto con las relaciones entre ellos, que proporcionan una base consistente para abordar un aspecto en el desarrollo de sistemas de información.

Siguiendo el enfoque de la *privacidad y la protección de datos desde el diseño* (tal como establece el Art. 25 del RGPD), es necesario tomar medidas proactivas que aborden la privacidad durante todo el ciclo de vida del desarrollo de un sistema. Si enfocamos las cuestiones de privacidad y protección de datos como uno de los '*problemas que aparecen de forma recurrente*' de la definición de los patrones presentada más arriba, entenderemos que la definición de un lenguaje patrones de diseño ofrece una respuesta a cómo los ingenieros pueden abordar la privacidad desde el diseño sin necesidad de ser expertos en el campo.

Para construir nuestro lenguaje de patrones, seguimos un **método** para crear tales lenguajes basado en los trabajos de Meszaros y Doble [132] y Noble [133]. Este método prescribe (entre otras actividades) que, para definir un lenguaje de patrones, se debe especificar: (1) la estructura que tendrá cada patrón, (2) las relaciones que pueden existir entre ellos, y (3) el listado preciso de patrones que incluye el lenguaje. Así, este método guiará la construcción de los resultados presentados en el resto del capítulo, donde, siguiendo la analogía lingüística, podemos hablar, respectivamente, de las dimensiones morfológica, sintáctica, y léxica del lenguaje de patrones que creamos.

Por otra parte, hemos realizado un **mapeo sistemático** de la literatura sobre los patrones de privacidad [3], donde hemos analizado en detalle 66 trabajos (filtrados manualmente a partir de una selección previa de 2122) sobre elementos reutilizables para la privacidad en las actividades de diseño de sistemas. Este análisis nos ha permitido valorar el grado de madurez de los trabajos existentes en relación con las distintas etapas en los procesos de creación y utilización de patrones. Los resultados (Figura 22) muestran cómo ha crecido el interés académico por los elementos de diseño reutilizables para el diseño sistemático de sistemas respetuosos con la privacidad; sin embargo, es necesaria una mayor cantidad de estudios empíricamente contrastados. Asimismo, este análisis sistemático nos ha servido para validar nuestros logros, al establecer una correspondencia entre los tipos de relaciones que definimos en nuestro lenguaje y las que se presentan en cada uno de los catálogos de patrones que hemos estudiado.

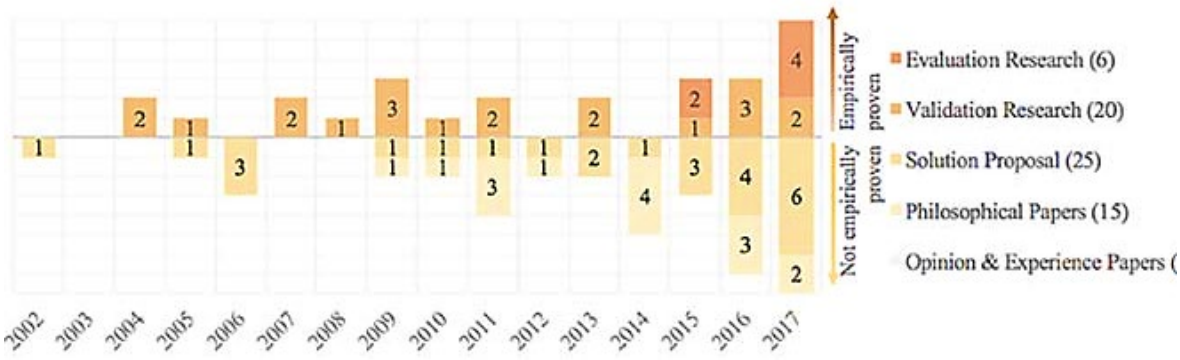


Figura 22. Evolución de la cantidad y madurez de publicaciones académicas sobre elementos reutilizables de diseño para la privacidad [3].

7.2 Lenguaje de patrones de privacidad

En primer lugar, abordamos la dimensión '*morfológica*' de nuestro lenguaje de patrones mediante la especificación de un **metamodelo o plantilla estructural** *que deben seguir los patrones de privacidad* que se definan. La Figura 23 muestra los distintos atributos la estructura de nuestros patrones de privacidad, que se pueden resumir en función de la pregunta a la que responden:

- *cuándo* se aplica (ámbito o contexto),
- *cómo* (descripción detallada del proceso, más ejemplos opcionales y casos de prueba),
- *para qué* (correspondencia con los requisitos de alto nivel a los que el patrón da respuesta),
- *quién* (rol responsable de aplicarlo, a quien se dirige el patrón), y
- *si* se puede aplicar (hasta qué punto está disponible, maduro técnicamente, y respaldado).

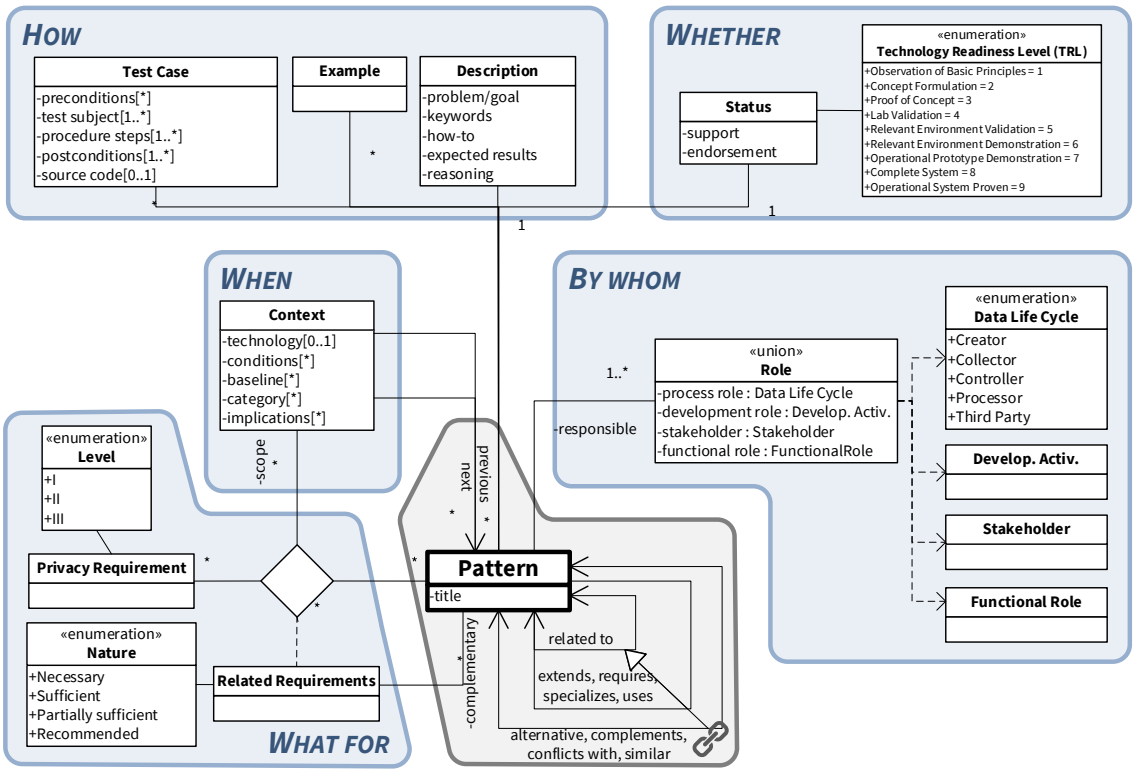


Figura 23. Modelo estructural de los patrones de diseño de ingeniería de privacidad.

En segundo lugar, abordamos la dimensión '*sintáctica*' de nuestro lenguaje de patrones de privacidad, estableciendo un conjunto de **tipos de relaciones** que pueden aparecer y utilizarse para relacionar un patrón con otro, a fin de organizarlos en sistemas y lenguajes de patrones. Por ejemplo, puede suceder que dos patrones ofrezcan soluciones *alternativas* al mismo problema, o que sus respectivas soluciones no se puedan aplicar en un mismo sistema porque entran *en conflicto* entre ellas, o que la solución proporcionada por un patrón genere un nuevo problema que *conduce a* la aplicación de otro patrón, etc. Así, hemos definido siete posibles tipos de relaciones (Alternativa, Complementa, Conflicto, Extiende, Conduce a, Requiere, Similar, Especializa, Usa), cuyas características se resumen en la Tabla 9 a continuación.

Tabla 9. Relaciones posibles entre los patrones de privacidad.

Relación	Problemas	Contextos	Soluciones	Propiedades lógicas
Alternativa	El mismo	Similares	Diferentes	Simétrica
Complementa	Similares o distintos	El mismo o distintos	Mejoradas mutuamente	Simétrica Si (X es alternativa a Y), entonces (X no complementa Y)
Conflicto	El mismo o distintos		Mutuamente excluyentes	Simétrica
Extiende	Extendido	Extendido	Mejorada	Asimétrica Transitiva Acíclica
	(al menos una de las tres)			
Conduce a	El problema generado por el patrón origen implica la solución proporcionada por el patrón destino			Si (X conduce a Y) y (Y conduce a X), entonces (X complementa a Y) [y viceversa]
Requiere		El contexto del patrón destino requiere la solución proporcionada por el patrón origen		Asimétrica Transitiva Acíclica
Similar	Similar		Similar	Simétrica
	(al menos una de las dos)			
Especializa	No disjuntos	No disjuntos	Más específico	Asimétrica Transitiva Acíclica
Usa			La solución del patrón destino se utiliza por la solución del patrón origen	Asimétrica Transitiva Acíclica

A continuación, abordamos la '*componente léxica*' mediante la construcción de un catálogo de patrones y las relaciones entre ellos, conformes con la plantilla estructuras definida más arriba y con las relaciones entre ellos. Nuestro catálogo recoge una lista de patrones que sigue a continuación: *Credenciales basadas en atributos*, *Iconos de privacidad*, *Panel de control de privacidad*, *Diseño en capas de políticas de privacidad*, *Visualización del encaje de políticas*, *Ofuscación de mediciones mediante ruido añadido*, *Pasarela de agregación*, *Plug-in confiable de privacidad*, *Evaluación federada de impacto de privacidad*, *Lista negra anónima basada en reputación*, y *Granularidad dinámica de localización*. Lógicamente, la creación de un catálogo completo requiere de una mayor amplitud asociada a un mayor esfuerzo, para lo que distintos investigadores (incluido el autor) constituyeron la comunidad en línea privacypatterns.org [134], donde se ha ido publicado un catálogo que incluye nuestros patrones y los de otros autores (la Figura 24 muestran capturas de pantalla de los publicados por el autor, y la Figura 25 las relaciones de una selección de ellos, que responden a la estrategia de "Control").

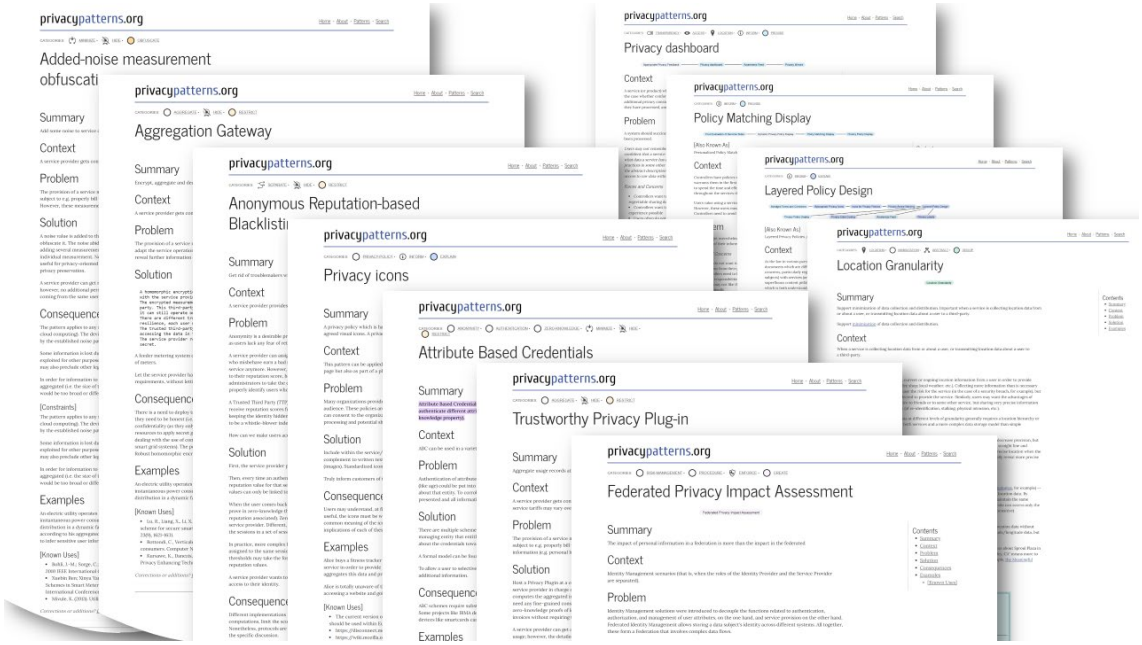


Figura 24. Capturas de pantalla de los patrones publicados por el autor en el portal privacypatterns.org.

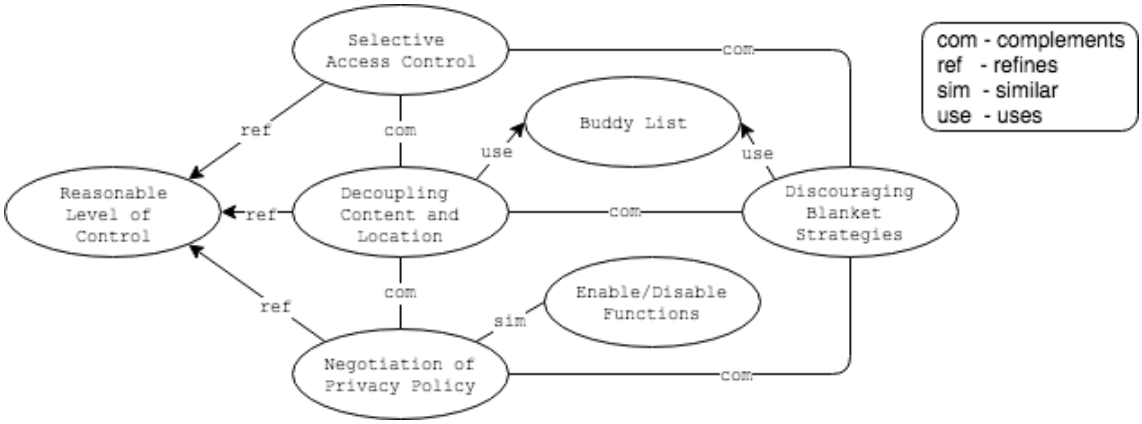


Figura 25. Relaciones entre los patrones que abordan la estrategia de "Control"

Finalmente, sugerimos un método que enmarca la introducción de patrones de privacidad dentro del ciclo global del desarrollo de un sistema, a partir de (1) la selección de los patrones apropiados en función de los problemas y el contexto de aplicación definidos por la especificación del sistema y (2) la instanciación de los patrones seleccionados (es decir, la creación de ejemplares individuales del patrón) a partir de los modelos de diseño del sistema bajo desarrollo.

8 Modelos normativos reutilizables para el aseguramiento de la privacidad y protección de datos

Este capítulo aborda el papel que tiene la *disciplina del aseguramiento (assurance) o garantía de sistemas en el cumplimiento normativo* del marco regulatorio de la privacidad y protección de datos (y, en particular, del RGPD):

- En primer lugar, analizamos la *relación entre el aseguramiento y la privacidad y la protección de datos*, y repasamos los principales tipos de modelos empleados en la disciplina del aseguramiento de sistemas (modelado regulatorio, modelado de patrones de argumentación, y modelado de equivalencias) y el papel que pueden tener para el cumplimiento normativo de la privacidad y la protección de datos.
- A continuación, para cada uno de esos tipos, creamos una serie de *modelos aplicados a las evaluaciones de impacto de protección de datos* (EIPD / DPIA) requeridas por los artículos 35 y 36 del RGPD, a la norma técnica estándar ISO/IEC 29134:2017 (*Tecnología de la información. Técnicas de seguridad. Directrices para la evaluación del impacto de la privacidad*), y a la plantilla para evaluación de impacto de protección de datos en la red eléctrica inteligente y en los contadores inteligentes, definida por la Comisión Europea, así como las equivalencias entre todos ellos.

8.1 El potencial del aseguramiento de sistemas para la privacidad y la protección de datos

Comenzaremos razonando el papel que puede desempeñar el aseguramiento de sistemas en la privacidad y la protección de datos, examinando los diferentes procesos y categorías conceptuales que aparecen en el aseguramiento de sistemas, y trazando su correspondencia con sus equivalentes en el ámbito de la privacidad y la protección de datos, con especial atención al RGPD.

La disciplina del aseguramiento de sistemas se centra en aportar **evidencias que demuestren** que se han llevado a cabo todas las actividades necesarias de acuerdo con un marco normativo (por ejemplo, el RGPD), y aborda la captura sistemática de evidencias, su trazabilidad, y la argumentación para justificar el cumplimiento de la norma. El aseguramiento desempeña un papel crucial en el cumplimiento normativo de requisitos extrafuncionales (es decir, aquellos requisitos de sistema que no especifican qué funciones debe ofrecer este, sino cómo debe ofrecerlas): aunque se había aplicado anteriormente a otros requisitos como la protección de la seguridad física (*safety*), es la primera vez que se aplica para el caso de los requisitos de privacidad y protección de datos.

Pero, más específicamente, en el contexto del RGPD, esta disciplina está estrechamente relacionada con los **principios de responsabilidad proactiva o rendición de cuentas (accountability) y transparencia del RGPD**, así como con otros principios y metas de la privacidad y la protección de datos, como la transparencia y la intervenibilidad. Al igual que el principio de responsabilidad proactiva es transversal a los demás principios, la disciplina de aseguramiento de sistemas discurre en paralelo a todo el Ciclo de Vida del Desarrollo de Sistemas (SDLC) [135], [136], y puede tener en cuenta los resultados de las actividades de otras disciplinas para utilizarlos como evidencias (desde la gestión de riesgos, pasando por los requisitos, hasta el diseño). Por otro lado, el aseguramiento es distinto de las actividades de prueba (*testing*) y validación; está un nivel por encima de ellas: no trata de producir verdades lógicas demostrables mediante pruebas formales, sino de dar soporte a una noción más subjetiva de *confianza justificada* [135], para lo que

puede exigir la realización de distintas actividades (incluyendo las pruebas, pero sin entrar a analizar su resultado).

La Figura 26 presenta las distintas actividades que pueden aparecer en un proceso de aseguramiento. Cabe destacar que existe un conjunto de actividades (en la parte superior de la figura), cuyos resultados se pueden reutilizar en múltiples proyectos, tras haberlas ejecutado en una única ocasión: el modelado regulatorio, el modelado de patrones de argumentación, y el modelado de equivalencias. El resto del capítulo se centrará precisamente en estas tres actividades y el papel de cada una de ellas en relación con la privacidad y la protección de datos (los detalles se pueden consultar en la memoria de la tesis doctoral que recoge en detalle nuestra investigación [1, Sec. 8.1]).

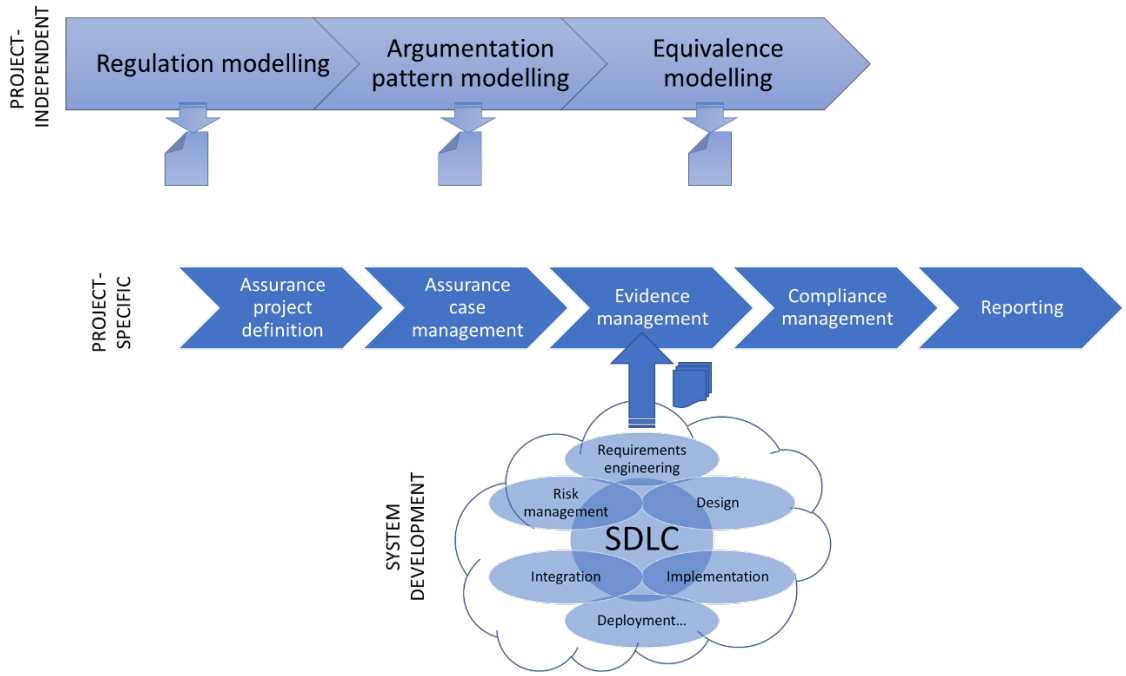


Figura 26. Proceso de alto nivel para el aseguramiento.

El **modelado regulatorio** traduce la normativa y los estándares cuyo cumplimiento se persigue en un modelo orientado al proceso de desarrollo, compuesto de una serie de actividades, roles, artefactos, y relaciones entre ellos (más otros elementos auxiliares, como requisitos formales, restricciones, o niveles de aplicabilidad, v. Figura 27). En relación con el RGPD podemos encontrar, p.ej. actividades de tratamiento de datos, actividades de protección de datos durante el desarrollo, y actividades de protección durante su operación, así como relaciones de precedencia o de composición entre algunas de ellas.

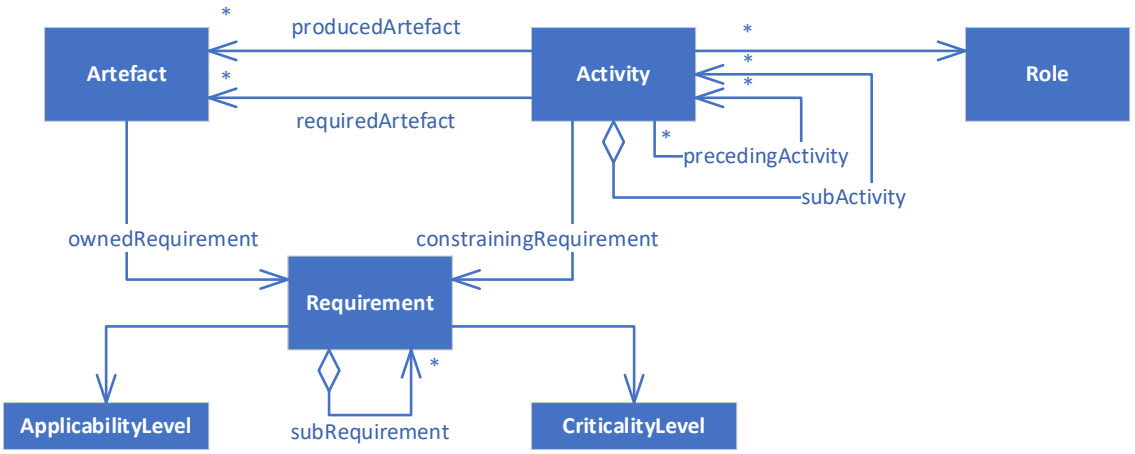


Figura 27. Representación simplificada de los elementos del metamodelo de los marcos de referencia.

El **modelado de patrones de argumentación**, que extrae de la normativa una serie de plantillas (patrones) de predicados lógicos (argumentaciones) que se pueden utilizar para establecer relaciones lógicas entre una norma y las evidencias que se usen para reivindicar la conformidad con ella. Los patrones de argumentación proporcionan una forma conveniente de respaldar a la justificación de las reivindicaciones de conformidad con una norma, pero, como patrones que son, no son la única solución posible, sino (ni más ni menos), una solución factible y probada en aquellos contextos a los que se refiere cada patrón. En relación con el RGPD, encontramos que estos patrones de argumentación se derivan de dos posibles orígenes: cláusulas condicionales, y controles de seguridad y de protección de datos.

El **modelado de equivalencias** establece una correspondencia entre los elementos de los modelos normativos (actividades, roles, artefactos…) de diferentes estándares o regulaciones, hasta el punto de que se puedan considerar parcialmente equivalentes. De este modo, se puede justificar el cumplimiento parcial de una norma a partir del cumplimiento de otra. En nuestro caso, planteamos la equivalencia entre estándares técnicos industriales de privacidad y el marco legal definido por el RGPD) así como entre ambos y las pautas para sectores específicos (red eléctrica inteligente o '*smart grid*').

Por último, la disciplina del aseguramiento puede servir para respaldar a la integración de la privacidad y la protección de datos a lo largo de la **cadena de suministro**, en aquellos escenarios donde el tratamiento de datos personales involucra a varios actores (p. ej. responsables conjuntos, encargados, o grupos de empresas.

8.2 Escenario de aplicación: método y resultados

El análisis planteado en la sección anterior se ha aplicado al RGPD y a estándares de privacidad para crear artefactos independientes del proyecto (tales como los resultantes de las tareas presentadas en la sección anterior) que pueden ser luego reutilizados en proyectos de aseguramiento específicos, tanto por el gestor del proceso de aseguramiento como por los ingenieros. En primer lugar, planteamos un método sistemático basado aplicar de nuevo el método QCA (ya aplicado en capítulos anteriores), para lo que hemos creado la **guía de codificación (codebook)**, cuya leyenda se muestra resumida en la Figura 28.

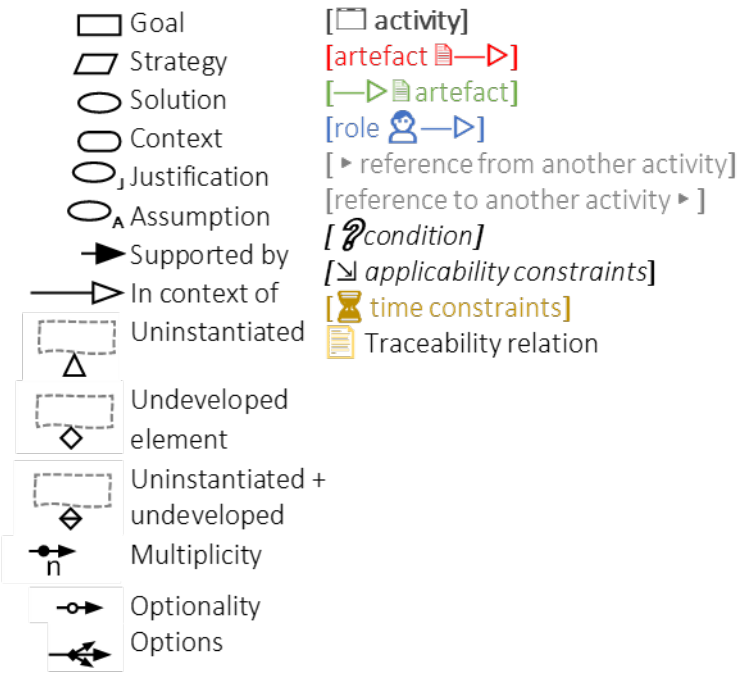


Figura 28. Elementos del código para la elaboración de modelos de marco regulatorio y patrones de argumentación.

Esta guía de codificación incluye reglas para extraer los elementos de los modelos de aseguramiento a partir del análisis sintáctico-semántico del texto normativo (en particular, nos valemos de los conceptos lingüísticos de los *roles theta y las relaciones temáticas*). Por ejemplo, identificamos un artefacto de entrada a partir del sintagma que ocupa el *tema* de locuciones verbales como 'tener en cuenta', o el complemento directo de un verbo de comunicación o de transferencia; igualmente, un artefacto de salida se identifica a partir del *tema* o *paciente* de 'resultar en', o de 'contener', o el complemento directo de otros verbos; asimismo, los roles involucrados en una actividad aparecen como el *agente*, o como complemento indirecto, o mencionado como 'asesor'; etc.

Hemos aplicado ese método para:

- *Modelar en términos de aseguramiento una EIPD a partir del RGPD* (complementado por documentos interpretativos como las directrices del Grupo "Protección de Datos" del Artículo 29 y el Comité Europeo de Protección de Datos CEPD/EDPB), en forma de un marco de referencia de aseguramiento y un conjunto de patrones de argumentación.
- Modelar, de manera análoga, la norma técnica *ISO/IEC 29134 de Evaluación de Impacto de la Privacidad* [137].
- Modelar la *correspondencia de equivalencias de ISO/IEC 29134 a las EIPD del RGPD (GDPR DPIA)*, como ejemplo de un estándar orientado a la ingeniería que puede poner en práctica las regulaciones legales de privacidad y protección de datos.
- Modelar la plantilla aprobada por la UE para las *EIPD en la red eléctrica inteligente* (*smart grid*) [138], y su *correspondencia* con las EIPD del RGPD y de la norma ISO/IEC 29134, como ejemplo del mapeo de un estándar específico del sector para operativizar los contenidos del RGPD.

Como **ejemplo y extracto de dichos modelos**, se presentan a continuación (Figura 29, Figura 30, Figura 31) sendos diagramas del modelo del marco de referencia y de un patrón de argumentación para la determinación de la necesidad de llevar a cabo una EIPD a partir del Art. 35.1 y 35.2 del RGPD, así como el modelo del marco de referencia para la susodicha plantilla de EIPD para *smart grid*. Se remite al lector a la memoria de la tesis doctoral para consultar los modelos completos, junto con la explicación detallada de cada uno de ellos [1, Secs 8.2, 8.4].

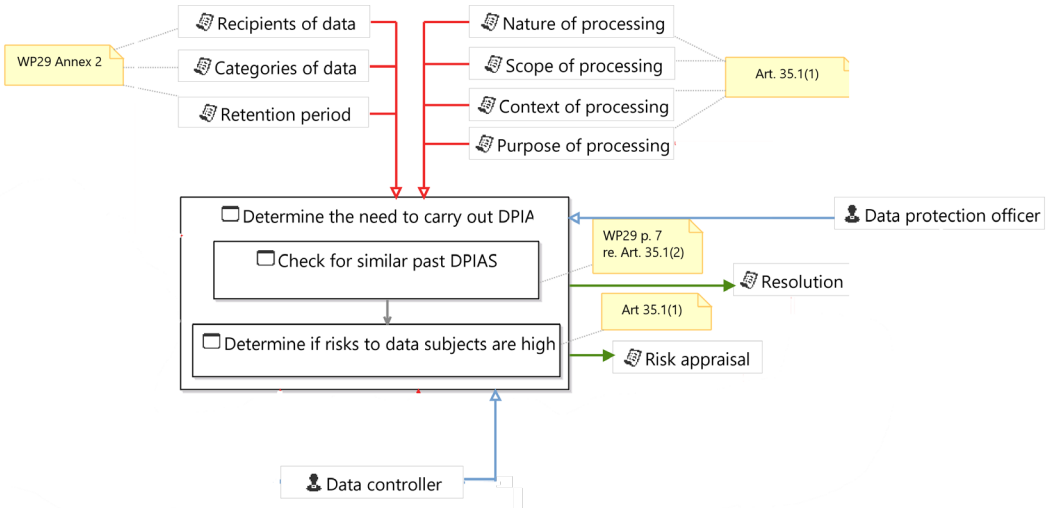


Figura 29. Modelo del marco de referencia del Art. 35.1 y 35.2 del RGPD.

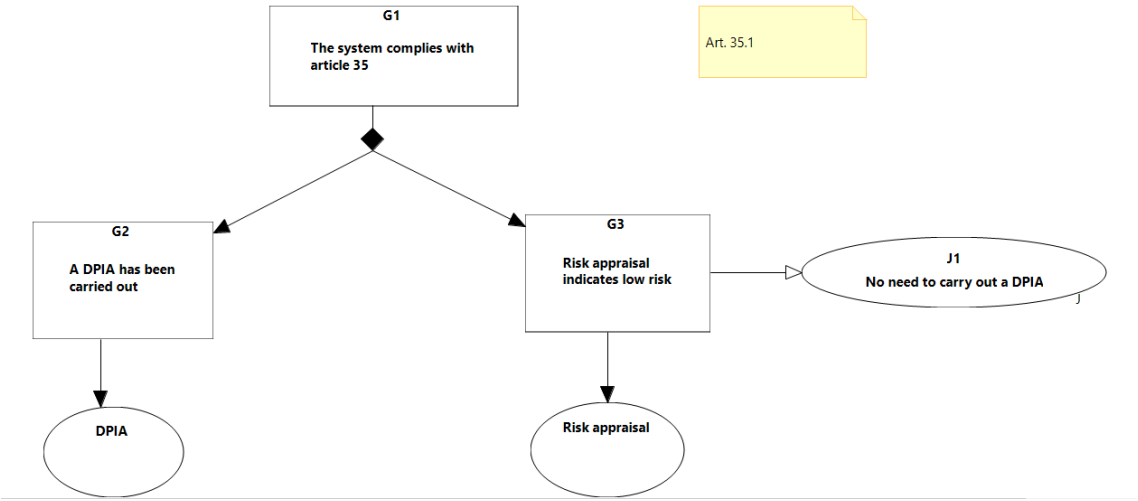


Figura 30. Patrón de argumentación para el artículo 35.1 del RGPD.

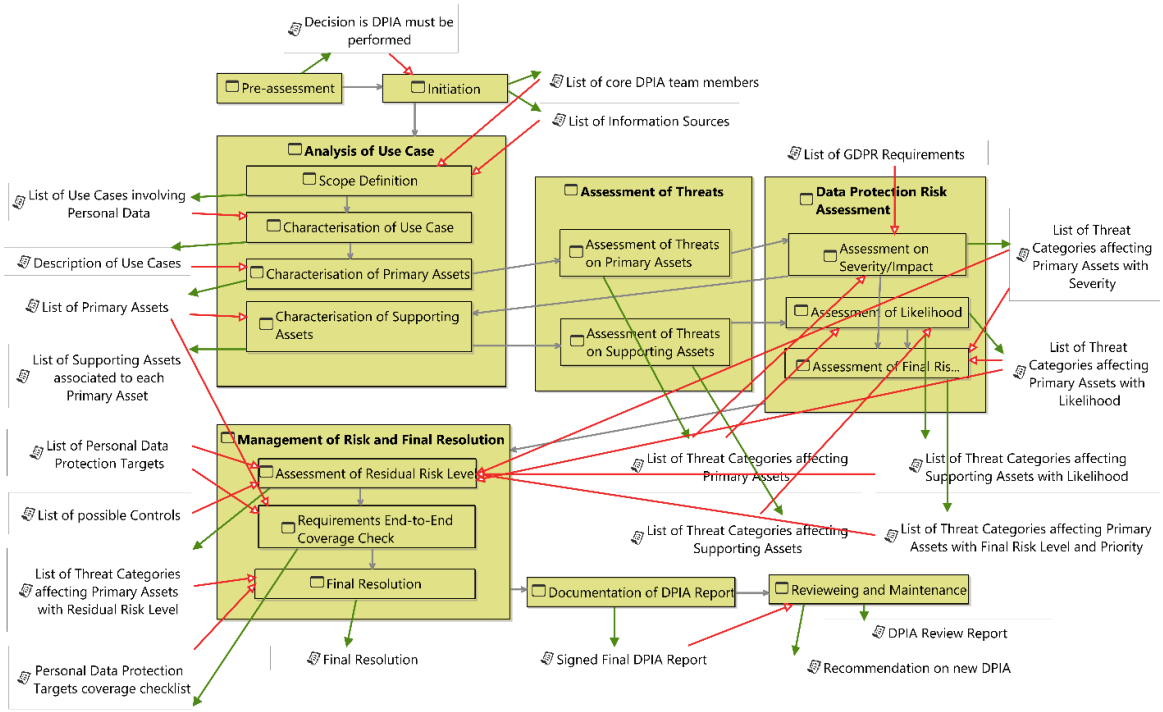


Figura 31. Modelo del marco de referencia obtenido a partir de la plantilla de EIPDs para red eléctrica inteligente (smart grid).

9 Ingeniería de métodos de privacidad y protección de datos

Este capítulo describe nuestras contribuciones al abordaje de la privacidad y la protección de datos desde la perspectiva de la disciplina de la ingeniería de métodos, que permite *ensamblar las actividades del resto de disciplinas*.

- En primer lugar, se presenta un *método para diseñar métodos de ingeniería*, basado en el enfoque denominado '*ingeniería situacional de métodos*', que emplearemos en nuestra investigación.
- A continuación, presentaremos nuestro *metamodelo de métodos de ingeniería de privacidad*, es decir, una serie de elementos metodológicos que aparecerán en los distintos métodos de ingeniería de privacidad que se pueden crear, y cuya definición facilita la integración entre distintos métodos de ingeniería de privacidad, y entre estos y los métodos de ingeniería de software y sistemas en general.
- Por último, presentamos la *descripción funcional de los métodos de ingeniería de privacidad propuestos propios de cada una del resto de las disciplinas* abordadas en nuestra investigación (gestión de riesgos, ingeniería de requisitos, diseño de sistemas, y aseguramiento de sistemas), así como una serie de casos de uso abstractos que engloban a todos esos métodos. Finalmente, introducimos una propuesta de arquitectura de una cadena herramientas software para ingenieros, que puedan ayudar a la implementación de estos métodos, y las relaciones entre ellas (que, como se verá en el último capítulo, ha servido de inspiración para herramientas software desarrolladas por terceros).

9.1 Construcción de métodos mediante la ingeniería situacional de métodos

La ingeniería de privacidad y protección de datos requiere la aplicación de *métodos sistemáticos para capturar y abordar los problemas de privacidad y protección de datos* a lo largo del proceso de desarrollo. Sin embargo, la diversidad de enfoques, tanto desde el lado de la privacidad como en la práctica de la ingeniería, junto con el contexto específico y el alcance de cada proyecto, han generado una gran cantidad de métodos de ingeniería de privacidad. Para hacer frente a este panorama, se puede recurrir a la **ingeniería de métodos**, disciplina que se encarga de "*diseñar, construir y adaptar métodos, técnicas y herramientas para el desarrollo de sistemas de información*" [139]; abordando el diseño de métodos para situaciones específicas (por ejemplo, una organización específica, un proyecto específico) en lugar de recurrir a metodologías rígidas. Así, esta disciplina permite describir los métodos existentes en términos de una variedad limitada de *elementos metodológicos* (y eventualmente permitir su recombinación en métodos nuevos y personalizados) y aplicarlos a distintos modelos de procesos de desarrollo.

Existen distintos estándares de facto para esta disciplina (ISO/IEC 24744 SEMDM [140], OMG SPEM 2.0 [141] y Open OPF [142]). Cada uno de ellos define un **metamodelo de metodologías**, aunque en todos los casos encontramos conceptos similares y comunes a toda la disciplina de la ingeniería situacional de métodos: unidades de trabajo (p. ej. actividades y procesos), productores, productos del trabajo, etapas, y recursos externos (p.ej. lenguajes, notaciones, pautas y restricciones). Con estos elementos, se puede llevar a cabo un proceso de **construcción de métodos** [143]. Bajo este paradigma, un método no se consideran una receta prescriptiva e inflexible cuyos elementos deban tomarse como un todo, sino que más bien se interpreta como el resultado de *combinar e integrar adecuadamente partes del método* que pueden haberse definido en distintos momentos y lugares. Por supuesto, para que dos partes del método se puedan combinar, deberán haberse definido de manera compatible. Esto es análogo a cómo se construye un sistema software como resultado de combinar componentes listos para usar (por

ejemplo, bibliotecas, sistemas operativos) junto con fragmentos desarrollados internamente: todos deben interoperar utilizando un lenguaje común y deben cumplir con las interfaces acordadas. Del mismo modo, en la ingeniería de métodos, todos los componentes que se integran deben haber sido definidos en el mismo metamodelo y ofrecer puntos de contacto compatibles. Un ejemplo concreto ayudará a entenderlo: un método de ingeniería de privacidad puede definir un paso donde se miden los riesgos y otro paso siguiente donde se clasifican (priorizan) los riesgos. Esto no excluye que algunos riesgos también se obtengan y midan a través de otro método y luego se introduzcan en el primero para su priorización, siempre que ambos utilicen una especificación compatible de riesgos (por ejemplo, un riesgo en ambos casos se define como: amenaza, vulnerabilidad, probabilidad e impacto).

Así, una de nuestras contribuciones en relación con la ingeniería de métodos es una *descripción funcional de la integración de diferentes métodos en las cuatro disciplinas previamente tratadas de la ingeniería de software y sistemas* (gestión de riesgos, ingeniería de requisitos, diseño y aseguramiento). Desde una perspectiva centrada en cada disciplina, aquí abordamos cómo las actividades y pasos tradicionales de cada disciplina se pueden adaptar al contexto de la privacidad y la protección de datos (para que resulten completamente nuevos para los no expertos en privacidad). Para ello, emplearemos los **modelos de proceso SIPOC** (que utilizaremos en la sección 9.3 para modelar las funciones expuestas por los métodos de cada una de nuestras disciplinas de ingeniería), provenientes de la metodología Six Sigma [144]. Los diagramas SIPOC ofrecen información en cinco columnas, respectivamente sobre los *Proveedores*, *Entradas*, *Procesos*, *Salidas* y *Consumidores* de cada proceso, así como información sobre las interacciones entre un subsistema y el exterior, y la relación de los procesos que implementan esas funciones con agentes externos, por lo tanto, ilustran indirectamente las dependencias potenciales entre las funciones.

9.2 Metamodelo de ingeniería de privacidad

Siguiendo el enfoque de la ingeniería de métodos situacionales, hemos contribuido con un marco conceptual que ordena los diferentes conceptos que generalmente subyacen a las diversas contribuciones acogidas bajo el paraguas de la Ingeniería de privacidad, generalizando y sistematizando los elementos de los métodos de ingeniería de privacidad para que puedan ser sometidos a comparación e integración entre ellos. Este marco se ha construido como un *metamodelo de ingeniería de privacidad que extiende el metamodelo SEMDM / ISO/IEC 24744 para metodologías de desarrollo de software y sistemas*, y proporciona un **vocabulario controlado de elementos metodológicos** de ingeniería de privacidad y un **conjunto normalizado de puntos de conexión y relaciones** para organizar esos elementos. Por lo tanto, permite la descripción de elementos de diferentes métodos existentes de la ingeniería de privacidad en términos comparables, para que puedan ser catalogados y evaluados. Ese metamodelo puede considerarse como si fuera las estanterías de un almacén etiquetadas, en cada uno de cuyos compartimentos se pueden anclar las distintas contribuciones sobre la ingeniería de la privacidad. Además, al enriquecer la descripción de los elementos del método con puntos de conexión bien definidos, se fomenta su reutilización e integración.

Los principales **elementos que hemos definido** en este metamodelo de métodos responden a la ingeniería de la privacidad desde distintas perspectivas:

- *Ontológica*: *Modelo Conceptual de Privacidad*, descripción conceptual de lo que se entiende por '*privacidad*' en el contexto de la teoría de la privacidad en la que se sustente un método de privacidad (teniendo en cuenta que se trata de un concepto esencialmente controvertido cf. Sec. 2.2 supra.). Frecuentemente puede ser descompuesto en *Unidades Conceptuales de Privacidad* (p.ej. principios).
- *Deontológica*: *Marco Normativo de Privacidad*, requisitos normativos, vinculantes o no, que debe cumplir un método de ingeniería de privacidad, incluyendo restricciones existenciales (elementos que debe incluir el método), temporales

(orden de los elementos en el método) y requisitos de proyecto (aplicables a los productos que se generen aplicando el método).

- *Situacional*: *Código de Ingeniería de Privacidad*, que clarifica la aplicación del marco normativo en determinados contextos. Se conocen en ocasiones como '*códigos de conducta*' o '*códigos de práctica*' y pueden estar constituidos por varias *Pautas*.
- *Epistémica*: *Base de Conocimientos de Privacidad*, que pueden reutilizarse en los proyectos. No se trata de recursos externos, sino de plantilla de elementos que se pueden instanciar en los proyectos que sigan el método dado (como, por ejemplo, los patrones de diseño o patrones de argumentación mencionados).
- Además, se añade un tipo abstracto de *rol* (*Rol de Ingeniero de Privacidad*) que agrupa las responsabilidades comunes que se pueden esperar de un ingeniero de privacidad.

La Figura 32 y Figura 33 muestran (sombreados en gris) las relaciones entre estos nuevos elementos, y entre ellos y los ya definidos por SEMDM.

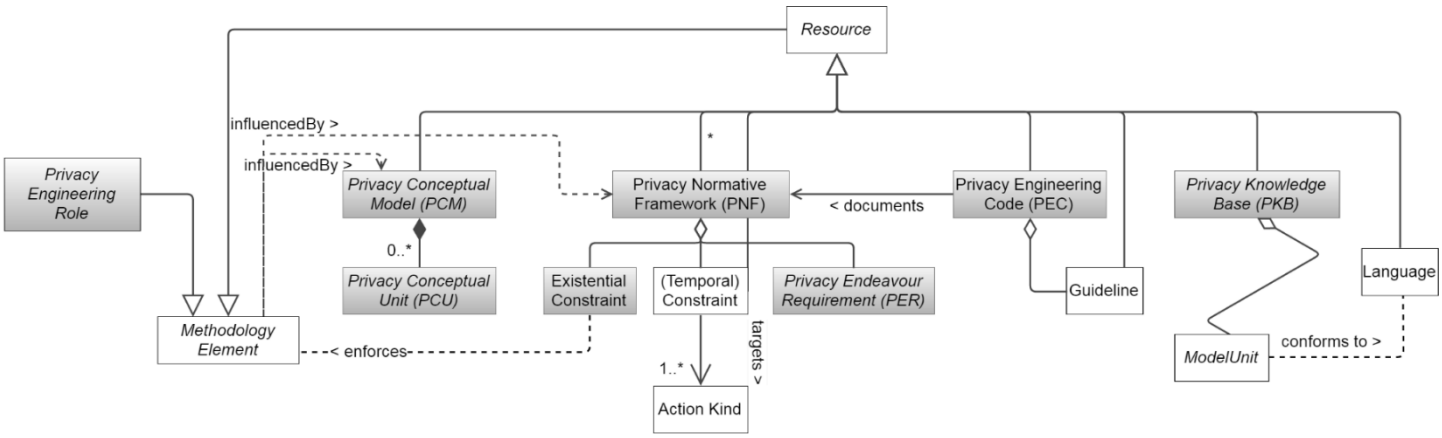


Figura 32. Metamodelo estructural de las extensiones a SEMDM para el marco de ingeniería de privacidad.

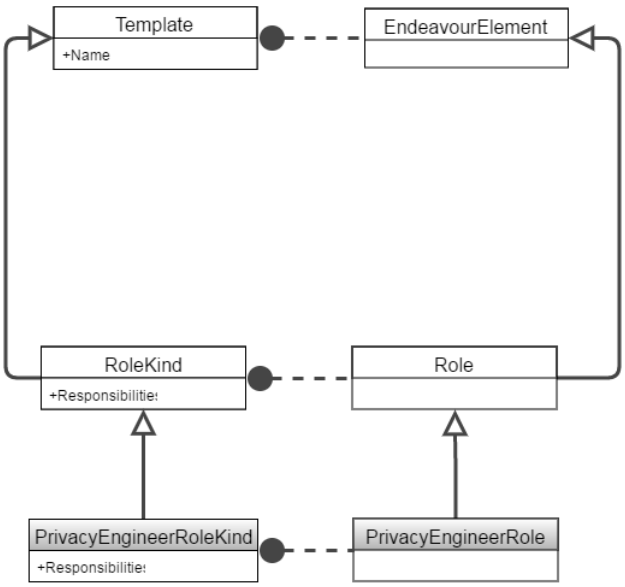


Figura 33. Rol / Tipo de Rol "Ingeniero de Privacidad".

Este metamodelo puede utilizarse para ayudar a especificar de manera formal cualquier método de ingeniería de privacidad. En nuestro caso, lo hemos aplicado, por ejemplo, a un método de ingeniería de privacidad ampliamente reconocido (LINDDUN).[1, Sec. 9.1.3]

9.3 Descripción funcional de métodos de ingeniería de privacidad

Siguiendo la metodología de modelado de procesos de Six Sigma (SIPOC) presentada más arriba, ofrecemos una propuesta de **modelos de proceso de actividades de ingeniería de privacidad** introducidas en las disciplinas respectivas:

1. *Gestión de riesgos*: las actividades tradicionales de gestión de riesgos se abordan desde la perspectiva de la privacidad y la protección de datos (por ejemplo, EIPDs). Por lo tanto, en este método, se capturan amenazas y vulnerabilidades de privacidad y protección de datos, se analizan la probabilidad y el impacto de los riesgos de privacidad y protección de datos, y se priorizan y se tratan los riesgos mediante controles de seguridad y privacidad. La gestión de riesgos se puede aprovechar para la selección de proveedores (es decir, encargados del tratamiento de datos) dependiendo de los controles que implementen.
2. *Ingeniería de requisitos*: se especifican, analizan y rastrean los requisitos de privacidad y protección de datos. Los requisitos de privacidad y protección de datos incluyen los derivados directamente de textos normativos y otros de la aplicación de otros métodos de obtención de requisitos; todos ellos particularizados a las especificidades de cada proyecto.
3. *Diseño de sistemas basado en modelos*: varias vistas de los modelos de sistemas se anotan con atributos relacionados con la privacidad y la protección de datos (por ejemplo, categorías de datos personales, actividades de tratamiento de datos, ámbitos de responsable y encargado del tratamiento), se analizan frente restricciones específicas de privacidad y protección de datos, y se transforman de acuerdo con estrategias predefinidas para mejorar el cumplimiento del sistema con los atributos de privacidad y protección de datos. Los almacenes de datos personales estructurados y no estructurados existentes y los archivos de código fuente se pueden analizar para crear las anotaciones mencionadas.
4. *Aseguramiento*: se recogen pruebas y, sobre ellas, se emiten reivindicaciones de cumplimiento para demostrar el cumplimiento de un marco normativo (que ha sido modelado adecuadamente de antemano).

En las siguientes figuras (Figura 34, Figura 35, Figura 36 y Figura 37), se resumen dichos procesos en forma de diagramas SIPOC (para la explicación detallada de cada proceso, se remite al lector a la memoria de la tesis doctoral en la que se enmarca esta investigación [1, Ch. 9]).

Más allá de la secuencia de pasos de cada disciplina y las entradas y salidas que comparten, podemos ver que los distintos métodos siguen una estructura general subyacente con puntos en común. Por ello, y desde una perspectiva transversal a las distintas disciplinas, hemos definido un conjunto de **casos de uso abstractos comunes**, para permitir organizar los métodos de acuerdo con una estructura común. Estos casos de uso (presentados en la Figura 38, y detallados según una plantilla formal en el texto de la tesis doctoral mencionada [1, Ch. 9]), responden a la siguiente secuencia:

1. Se modela un marco de privacidad y protección de datos (por ejemplo, RGPD, u otros).
2. Ese marco se adapta para seleccionar aquellas partes que son aplicables a las necesidades de un proyecto u organización.
3. Aparte de ello, se modelan una o más vistas del sistema, incluidas las propiedades relevantes con respecto a la privacidad y la protección de datos.
4. El marco de privacidad y protección de datos se instancia con los valores específicos de los parámetros del marco en un proyecto determinado.
5. Las vistas del modelo del sistema se transforman, con el fin de abordar las propiedades de privacidad y protección de datos y mejorar la calidad del sistema desde la perspectiva de la privacidad y la protección de datos.
6. El sistema (o sus modelos) se evalúan en función del cumplimiento del marco de privacidad y protección de datos dado.
7. Los proveedores externos también son evaluados con respecto al cumplimiento.

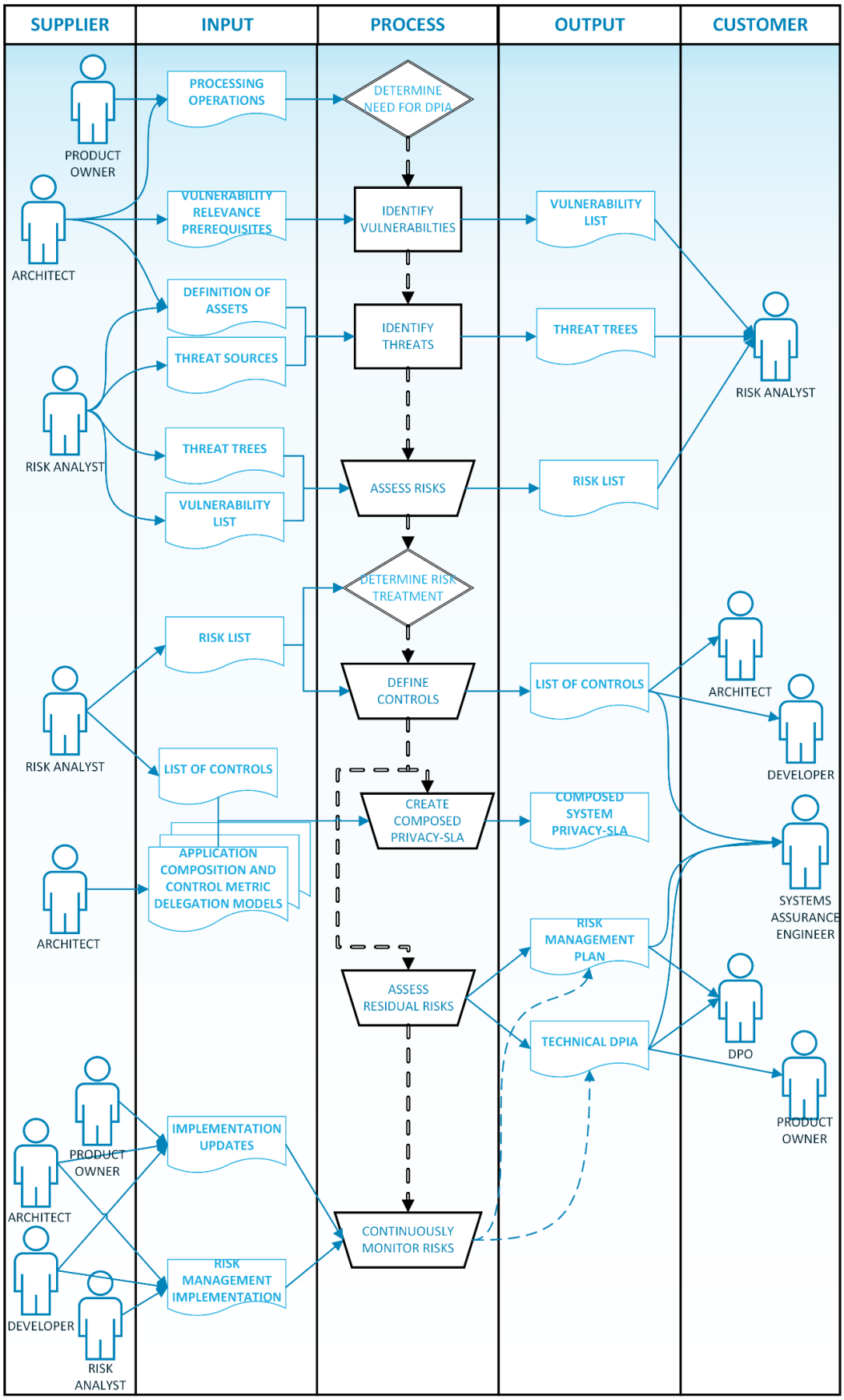


Figura 34. Diagrama SIPOC de las funciones de gestion de riesgos.

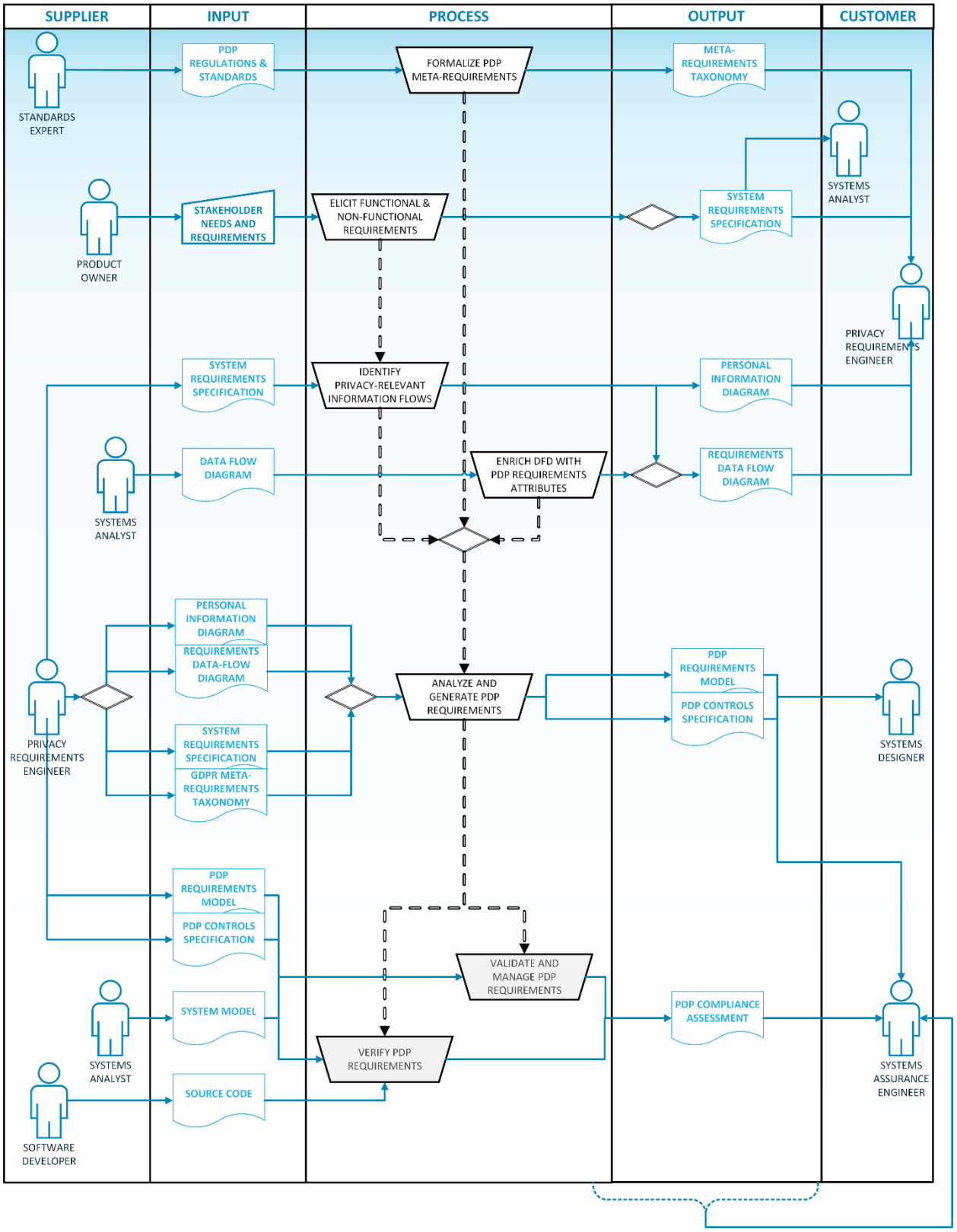


Figura 35. Diagrama SIPOC de las funciones de ingeniería de requisitos

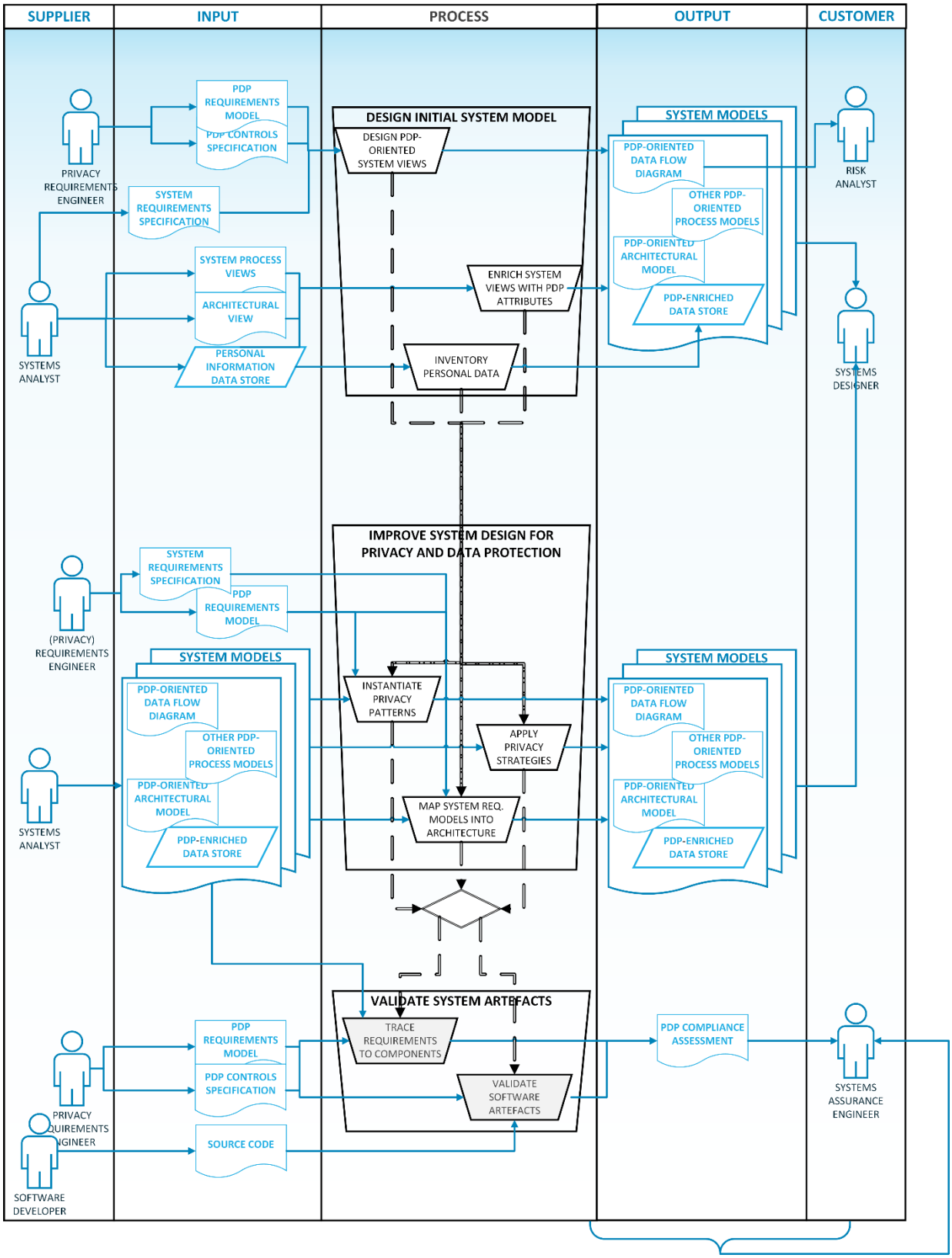


Figura 36. Diagrama SIPOC de las funciones de diseño dirigido por modelos.

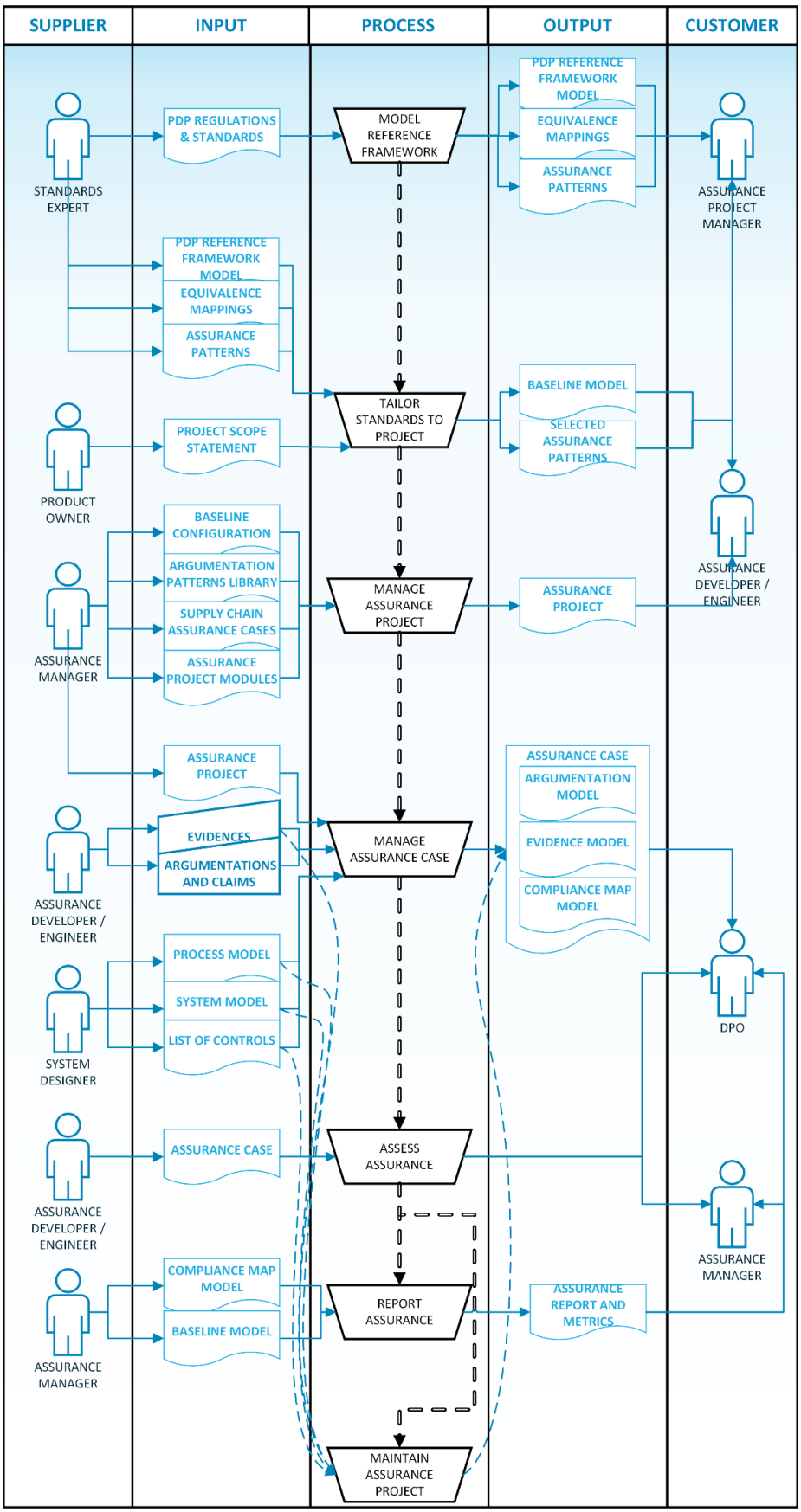


Figura 37. Diagrama SIPOC de las actividades de aseguramiento del sistema.

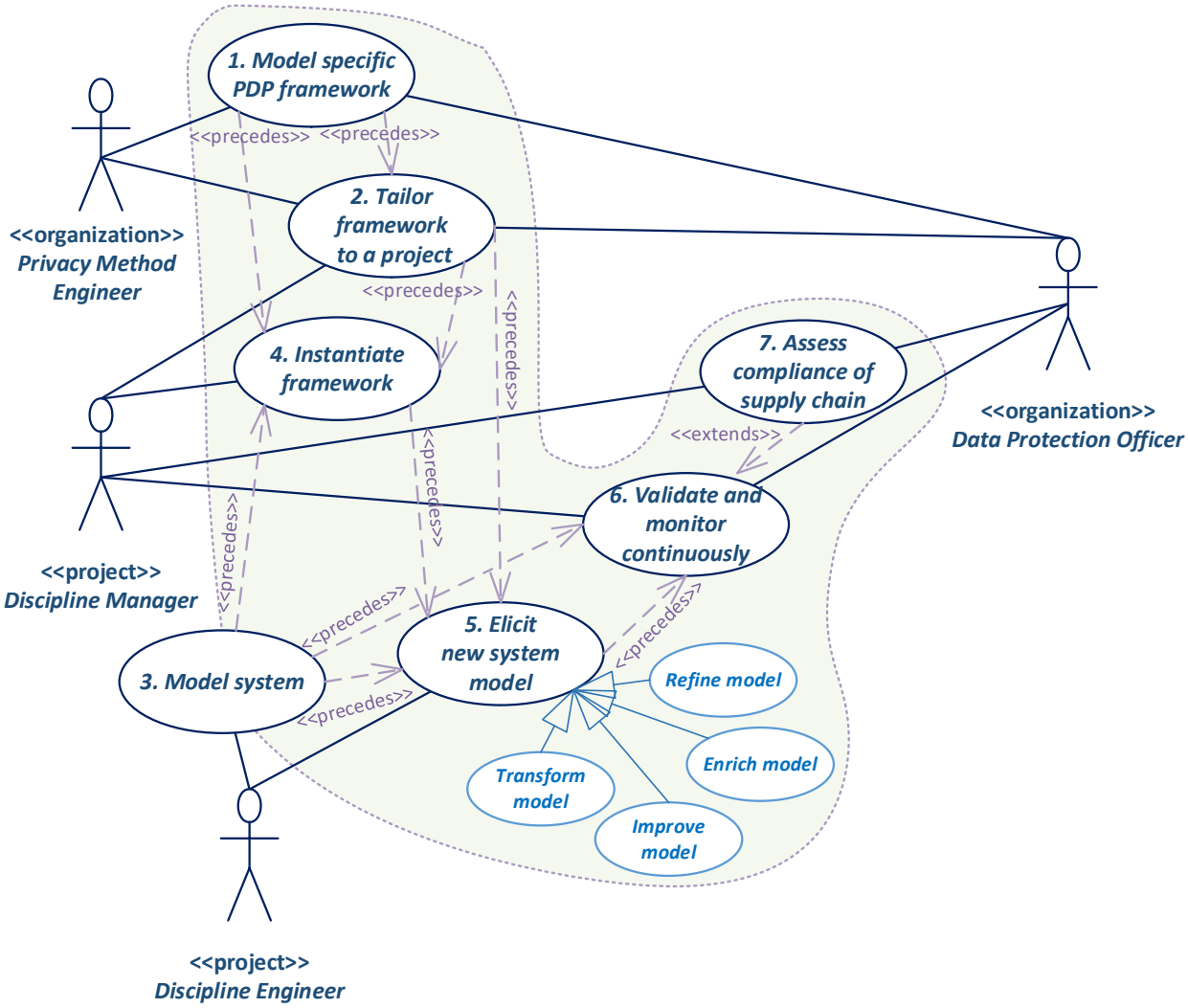


Figura 38. Casos de uso abstractos en los métodos de ingeniería de privacidad.

Por último, presentamos una propuesta de una **arquitectura de una cadena de herramientas de software** para respaldar los métodos de ingeniería de privacidad y protección de datos, que soporte los métodos presentados y fomente su automatización, incluida una descripción de los componentes de software en los que consistiría esta cadena de herramientas y el enfoque general de integración entre ellos, incluidos los puntos de integración entre sí y las posibles interacciones entre las herramientas, según se muestran en la Figura 39.

La razón de esta propuesta es que *los métodos como tales no resultan de mucha utilidad a menos que estén respaldados por herramientas de software* que guíen a los desarrolladores en el proceso de desarrollo y les ayuden a agilizarlo, automatizando las actividades apropiadas, proporcionando orientación semiautomática (p.ej. asistentes, etc.), fomentando la reutilización de las soluciones y conocimiento disponibles, proporcionando un fácil acceso a los recursos y la documentación; y, en definitiva, reduciendo el tiempo, el coste, la complejidad y la carga del desarrollo, sobre todo si se integran con otras herramientas que ya están acostumbrados a emplear en su trabajo cotidiano. Hablamos de una *'cadena' de herramientas* porque entendemos que diferentes roles utilizarán diferentes herramientas para diferentes funciones (y, posiblemente, estarán más familiarizados con las interfaces de usuario típicas en su respectiva disciplina), por lo que no tiene sentido tener una herramienta monolítica que soporte toda la metodología, pero sí que los distintos ingenieros cooperar en aspectos de privacidad y protección de datos (al igual que cooperan en sus tareas a lo largo del ciclo de vida del desarrollo). De ahí que propongamos una *integración ligera* (funcional y metodológica) entre las herramientas, para garantizar el flujo fluido de la información relevante de privacidad y protección de datos a lo largo del ciclo de vida del desarrollo.

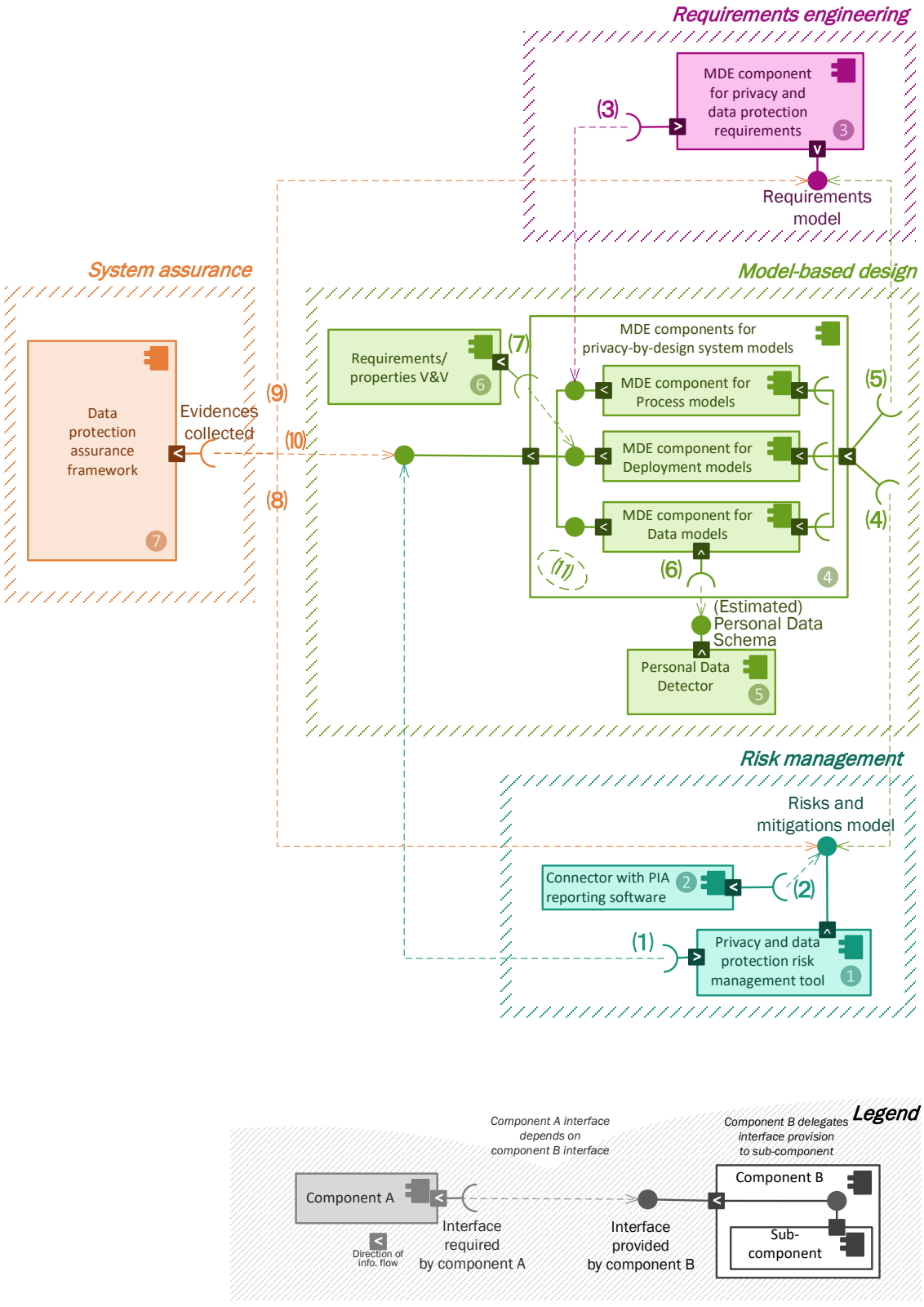


Figura 39. Arquitectura propuesta para una cadena de herramientas software para el apoyo a los métodos de ingeniería de privacidad y protección de datos.

Esta propuesta de arquitectura, como mostraremos en la sección 10.3, ha servido como guía para la creación por parte de otras organizaciones de una cadena de herramientas de código abierto que implementan los métodos propuestos, logrando así la transferencia de los resultados de esta investigación al sector productivo.

10 Valoración, trabajo futuro y conclusiones

En este documento hemos presentado varias contribuciones originales para abordar las cuestiones de privacidad y protección de datos desde las distintas disciplinas de la ingeniería de software y sistemas. Este último capítulo evalúa los resultados obtenidos en nuestra investigación —incluyendo una serie de evidencias que demuestran su impacto— y, a partir de ellos, propone una mirada al futuro con diferentes acciones que llevar a cabo para que el trabajo de los ingenieros redunde, en última instancia, en mejorar el respeto a la privacidad y la protección de datos.

10.1 Resumen de resultados obtenidos y alineación con la AEPD

Las **contribuciones originales** presentadas a lo largo de este documento cubren los objetivos de investigación planteados al inicio y presentados en la introducción. La Tabla 10 lista los principales resultados de investigación, identificados por el número de capítulo donde aparecen.

Tabla 10. Contribuciones originales del trabajo de investigación.

C2a. Una **definición operativa del concepto de privacidad,** desde una perspectiva multidimensional (aplicando el marco teórico de los '*conceptos esencialmente controvertidos*', tras haber analizado en profundidad las definiciones y descomposiciones predominantes de la privacidad).
C2b. Un **modelo de adversario** para el ámbito de la ingeniería de privacidad y protección de datos que tiene en cuenta la interacción entre el ingeniero *'honesto pero despreocupado'*, el custodio de los datos personales, y los intermediarios de datos sin escrúpulos, considerando los privilegios, metas y capacidades de cada uno de ellos.
C2c. Un **análisis de los paradigmas dominantes** actualmente en el ámbito de la privacidad y la protección de datos, y cómo sus debilidades motivan el abordaje desde un nuevo enfoque complementario: la *Ingeniería de Privacidad.*
C3. Un **análisis de las necesidades** a que afrontan los ingenieros para implementar la privacidad y la protección de datos en los sistemas que desarrollan y, en particular, para poner en práctica el RGPD; las **oportunidades** para resolver esas necesidades con el apoyo de las disciplinas de ingeniería de software y sistemas (gestión de riesgos, ingeniería de requisitos, diseño de sistemas basado en modelos, aseguramiento de sistemas e ingeniería de métodos); y las **deficiencias** que quedan por cubrir por las herramientas actuales.
C4. Un **lenguaje específico de aspectos del dominio** para el modelado de las cuestiones de privacidad y protección de datos, que abarca múltiples puntos de vista (requisitos, riesgos, diseño de estructuras de datos, flujos de datos y arquitectura de despliegue, aseguramiento de sistemas, y métodos de desarrollo), y que puede utilizarse para enriquecer los modelos de sistemas definidos en lenguajes de modelado convencionales, independientemente de su implementación.
C5a. Una **definición operativa de la minimización de datos** en múltiples dimensiones cuantificables (que abarcan los propios datos personales, los interesados a los que se refieren, otros datos inferidos, y los usos, tiempos y lugares donde se tratan).
C5b. Su aplicación a una serie de **métricas de riesgos de privacidad** de ámbito organizacional, basadas en la teoría de la información, junto con un método de decisión operativa para comparar los riesgos asociados a diferentes controles o proveedores de servicios.

C6a. Un enfoque dual para la operativización de los requisitos de privacidad y protección de datos que combina la elicitación basada en riesgos con los procesos de refinamiento orientados a metas.
C6b. Un modelo de requisitos derivados de los estándares de privacidad y su correspondencia con los requisitos legales.
C7. Un modelo de un lenguaje de patrones de diseño para la privacidad y la protección de datos, incluida la definición de la estructura de los patrones de privacidad, las relaciones entre ellos, ejemplos de patrones y su aplicabilidad.
C8a. Una aplicación de procesos de aseguramiento para apoyar los principios de responsabilidad y transparencia del RGPD, incluido un modelo de los aspectos de EIPD (DPIA) del RGPD como marco de referencia de aseguramiento y un conjunto de patrones de argumentación.
C8b. Modelos de un estándar industrial para EIPD y de pautas para la industria de redes eléctricas inteligentes (*smart grid*).
C8c. Una correspondencia entre los modelos de ambos.
C9a. Un metamodelo de métodos de ingeniería de privacidad y una descripción de casos de uso abstractos de ingeniería de privacidad.
C9b. Su aplicación a través de una descripción integral de la introducción de métodos de ingeniería de privacidad en el SDLC.
C9c. Una propuesta de la arquitectura de una cadena de herramientas para implementar esos métodos a través de herramientas de modelado convencionales.
C10. La *validación y transferencia* externa de nuestros resultados a través de proyectos de investigación industrial, y propuestas para el futuro, incluidas recomendaciones de políticas y una agenda de investigación.

Cabe mencionar, asimismo, que estos resultados están plenamente *alineados y desarrollan las recomendaciones emitidas por los supervisores de protección de datos y las agencias responsables* de la protección de ciberseguridad. Por ejemplo, la AEPD, en su guía de *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales* [145], aborda la necesidad del cumplimiento normativo como precondición para la gestión del riesgo, y la diferencia entre los riesgos para los derechos (en el sentido del RGPD) y para la organización: ambas reflexiones están completamente alineados con nuestro análisis respecto a la complementariedad de los enfoques orientados a metas y dirigidos por riesgos (v. Sec. 6.2), sin que los segundos deban sustituir a los primeros. En ese mismo documento aparece un análisis de factores de riesgo, muchos de los cuales pueden encontrar correspondencia en alguna de las dimensiones de minimización que abordamos (v. Sec. 5.2). Por último, el documento lleva a cabo una descripción detallada del proceso de análisis de riesgos para las EIPDs, que está alineado con nuestro modelo del marco regulatorio de EIPDs para el aseguramiento (v. Sec. 8.2). Asimismo, la guía de ENISA sobre *Ingeniería de la Protección de Datos De la teoría a la práctica* [146] aborda la diferencia entre los principios de privacidad y protección de datos desde el diseño (PbD / DPbD) y la necesidad de sistematizar los métodos y herramientas de la Ingeniería de Protección de datos (que abordamos en este trabajo, y que discutimos profusamente en las Secs. 1.2 y 2.3). Más aún, la propia *Guía de Privacidad desde el Diseño de la AEPD* [147], (aparte de utilizar las estrategias de minimización que utilizamos como fuente en nuestra Sec. 5.2) referencia varios de nuestros propios trabajos y contribuciones aquí reseñadas (como los catálogos de patrones o las guías metodológicas de privacidad de PRIPARE).

10.2 Orientación de políticas para la CE (*policy brief*)

Uno de los resultados de nuestro trabajo de investigación ha sido un **documento de orientación de políticas presentado a la Comisión Europea**, cuyo objetivo es abordar la ingeniería de privacidad y protección de datos desde la perspectiva legal e institucional. Consideramos que la adopción de las prácticas de ingeniería de privacidad y protección de datos será más probable si va de la mano de la persuasión del marco institucional mediante medidas coercitivas e incentivadoras.

Más concretamente, analizamos *dos necesidades* principales que deberá abordar el marco regulatorio e institucional:

1) Desplazar el enfoque de las leyes de privacidad y protección de datos desde la mera operación del servicio a las etapas iniciales del ciclo de desarrollo.
2) Simplificar la introducción de la privacidad y la protección de datos en el desarrollo de software, sistemas y servicios por parte de ingenieros no expertos en privacidad.

Para plantear la respuesta que se puede ofrecer a estas cuestiones desde un punto de vista legal, hemos llevado a cabo un **análisis jurídico comparado** sobre cómo la legislación de la UE ha abordado cuestiones similares para tratar otros requisitos extrafuncionales (seguridad laboral, seguridad de productos de consumo, accesibilidad tecnológica, etc.) Como resultado, hemos emitido **cuatro recomendaciones** que puedan orientar el desarrollo legislativo y las políticas institucionales en la privacidad y protección de datos, resumidas a continuación:

Recomendación #1: Dirigir las leyes de privacidad y protección de datos de la UE explícitamente a las organizaciones que desarrollan los productos, servicios e infraestructuras (que implementan los tratamientos de datos personales), *ampliando el ámbito subjetivo del requisito de privacidad y protección de datos desde el diseño (PbD / DPbD) a las organizaciones que desarrollan* dichos productos, servicios e infraestructuras, en tanto que determinan en la práctica las características de los tratamientos de datos posibles y las medidas de protección que incluirán.

Recomendación #2: *Requerir desde la legislación el cumplimiento vinculante de estándares para la ingeniería de privacidad y protección de datos*. Somos conscientes de que la legislación se debe ceñir a establecer requisitos (el *qué*), mientras que debe ser neutral en cuanto a la solución tecnológica adoptada (el *cómo*); sin embargo, la legislación puede contemplar la inclusión de mecanismos legales para requerir el cumplimiento de estándares técnicos (estándares armonizados, especificaciones técnicas elegibles, mandatos de elaboración de estándares, mejores técnicas disponibles), como han hecho anteriormente, por ejemplo, la Directiva general de seguridad de productos [148], la Directiva sobre la accesibilidad de sitios web y aplicaciones móviles [149], el Acta europea de accesibilidad[150], el Reglamento de dispositivos Médicos [151], el Reglamento de equipos de protección individual [152], la Directiva de emisiones industriales [153]; así como aparece en las más recientes Directiva de seguridad de las redes y sistemas de información 'NIS2' [154], Reglamento sobre servicios digitales 'DSA' [155] y del Reglamento sobre la resiliencia operativa digital 'DORA' [156]. Otros mecanismos legales, como los actos delegados y de ejecución, se han utilizado también en algunas de estas normas para habilitar a la Comisión Europea para establecer requisitos detallados sin necesidad de una (lenta) reforma de la norma habilitante.

Recomendación #3: *Facilitar el desarrollo sencillo de soluciones respetuosas con la privacidad, fomentando la amplia disponibilidad de soluciones reutilizables, probadas y consensuadas*. Los ingenieros no están preparados para trasladar directamente el contenido de los textos legales a su práctica habitual, pero tampoco están capacitados para desarrollar soluciones específicas para la privacidad y la protección de datos si no son expertos en el tema. Por el contrario, necesitan disponer de soluciones reutilizables, listas para 'arrastrar y soltar', cuyo uso no

requiera de una formación o conocimientos avanzados (igual que los ingenieros tampoco se preocupan de los detalles técnicos del cifrado cuando crean un sitio web cuyas comunicaciones están securizadas mediante HTTPS). Aunque hay instituciones e iniciativas (IPEN del CEPD/EDPB, ENISA, el CyBoK) que han venido recogiendo catálogos parciales de soluciones en los últimos años, se necesita un empuje para ofrecer un auténtico cuerpo de conocimiento compartido y ampliamente reconocido por la comunidad de práctica de la ingeniería en su conjunto.

Recomendación #4: *Ampliar el ámbito de las certificaciones de privacidad, fomentando su aplicación en la adquisición de servicios, sistemas, componentes de software y hardware* respetuosos con la privacidad y la protección de datos. Consideramos que la difusión de las certificaciones puede ser clave para fomentar la adopción de soluciones respetuosas con la privacidad: si los ingenieros pueden confiar en que todos los componentes involucrados en la provisión de un servicio respetan la privacidad y la protección de datos, si pueden apoyarse en pautas ampliamente aceptadas sobre cómo proceder en sus desarrollos, y si tienen a mano soluciones fáciles de usar, entonces les resultará mucho más directo asegurarse de que los servicios que creen son a su vez respetuosos con la privacidad.

10.3 Transferencia tecnológica

Desde una perspectiva de transferencia tecnológica, nuestras *contribuciones han sido transferidas a y validadas por los socios industriales del contexto del* **proyecto europeo PDP4E** (un proyecto de investigación europeo del programa *Horizonte 2020*, de 3 años de duración y financiado con 3,4 millones de euros). La misión de PDP4E es muy similar a la de nuestro trabajo de investigación que hemos presentado en este documento (integrar el conocimiento en privacidad y protección de datos con la práctica de la ingeniería de software y sistemas). Sin embargo, PDP4E ha ido un paso más allá al introducir características de privacidad en herramientas de ingeniería del software de uso habitual, para que los ingenieros no necesiten aprender herramientas que les resultan ajenas cuando aplican los métodos de ingeniería de privacidad y protección de datos.

El autor de este documento ha sido *líder científico técnico* del proyecto desde su preparación y durante toda su ejecución y líder del Paquete de Trabajo 2 sobre especificación de necesidades y arquitectura. Nuestra propuesta metodológica integrada (presentada en el capítulo 9) se ha desarrollado en características de distintas **herramientas software de código abierto para ingenieros**, dentro del ecosistema de herramientas Eclipse (MUSA DST, Papyrus y OpenCert), que soportan estos métodos. Así, los métodos que hemos propuesto se han trasladado a herramientas software que los respaldan, demostrando y validando la viabilidad de nuestro enfoque. Para guiar la arquitectura de la cadena de herramientas que los socios PDP4E han implementado se ha seguido un enfoque basado en modelos (presentado en el capítulo 4), que ha resultado clave para garantizar su adaptabilidad, flexibilidad, reusabilidad e interoperabilidad.

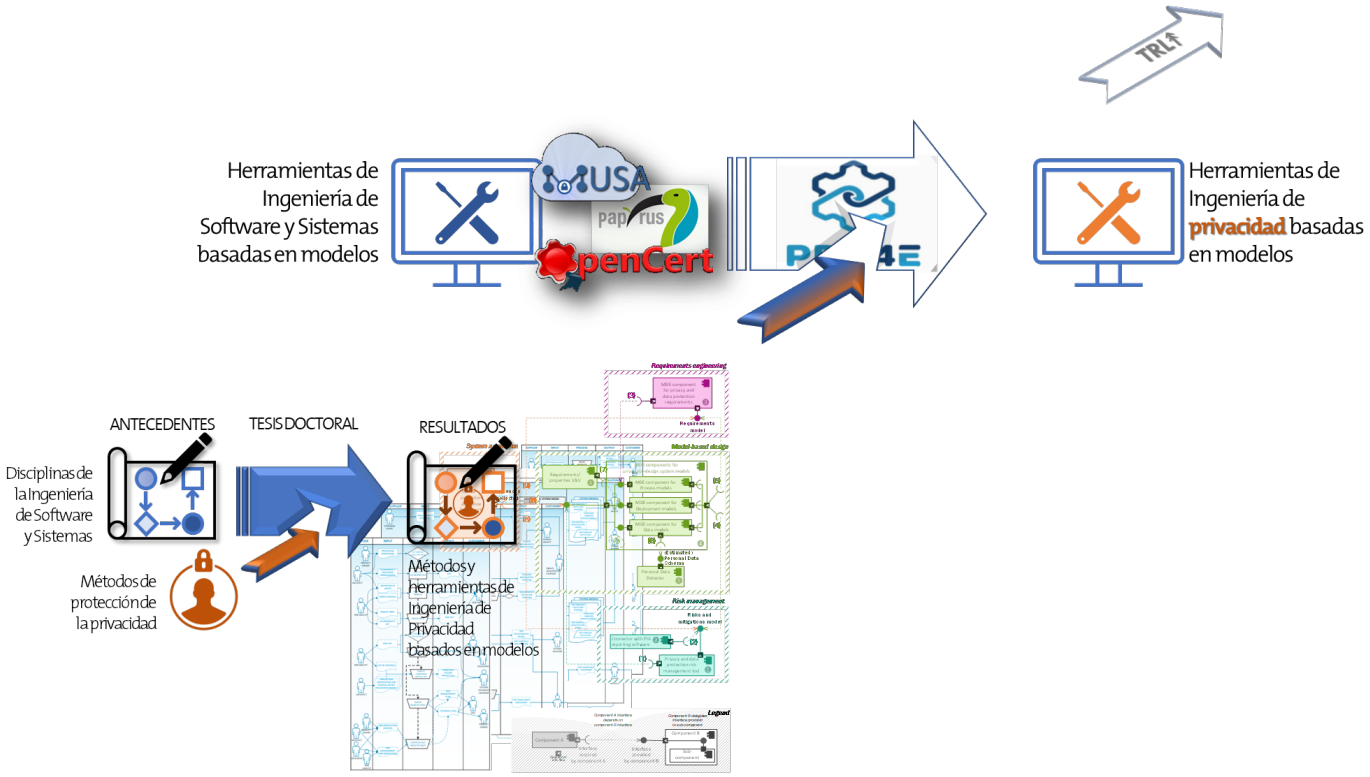


Figura 40. Transferencia tecnológica de los resultados a través del proyecto PDP4E.

En la actualidad *continuamos nuestro trabajo en dos iniciativas*:

- La constitución de un **Grupo de Interés en la Eclipse Foundation** enfocado en modelos abiertos de protección de la privacidad, que «*desarrollará una serie de modelos abiertos y reutilizables de ingeniería para la protección de la privacidad que se puedan utilizar en aplicaciones de consumo, Internet de las Cosas y procesamiento de datos*», aprovechando la amplia comunidad de desarrolladores que alberga esta fundación de proyectos de código abierto.
- Un ítem de trabajo preliminar sobre modelos de privacidad (**ISO/IEC PWI 27564 Privacy models**) aceptado en el grupo ISO/IEC JTC1 SC27/WG5, del que el autor es co-editor.

10.4 Agenda de investigación

Si repasamos los retos que quedan abiertos en nuestros resultados, encontramos cuatro vías futuras que pueden derivarse y para los que proponemos una **agenda de investigación**.

En primer lugar, se puede aumentar la **amplitud del alcance, replicando nuestro enfoque** a más *regulaciones de privacidad y protección de datos*: cubrir todo el ámbito del RGPD, acudir a normativas secundarias (certificaciones, códigos de conducta, pautas del EDPB/WP29, cláusulas contractuales, políticas vinculantes), otras normativas generales relacionadas (directivas ePrivacy, NIS, Law Enforcement Data Protection, DSA, DMA, etc.), normativas sectoriales o de industrias específicas, abordar más normas ISO (p.ej. ISO/IEC CD 27557, ISO/IED WD TS 27561 'POMME'), legislación de otros países, etc.

En segundo lugar, buscando también ampliar el alcance, pero ahora en una dimensión ortogonal, planteamos la aplicación a **más disciplinas de ingeniería**, y, en particular, a las actividades de *prueba (testing) y verificación*, que no hemos abordado en nuestro trabajo: creemos que, siguiendo nuestro enfoque podrían ser útil la verificación de modelos, la validación de modelos y las pruebas basadas en modelos, apoyándose además en métodos modernos de generación de casos de prueba (como la ejecución simbólica, el *fuzzing*, las pruebas metamórficas, o las pruebas de mutación).

En tercer lugar, existen **ámbitos tecnológicos** de gran extensión en la actualidad para las que nuestro enfoque podría presentar obstáculos y para los que habría que analizar su adaptabilidad.

- Por una parte, el uso de *datos inherentemente no estructurados* presenta dificultades para abordar la privacidad y protección de datos desde un enfoque puramente basado en modelos. Por ejemplo, nuestra propuesta de métricas de riesgos asume que existe una base de datos relacional que contiene datos personales. Incluso aunque se extendiera mediante ingeniería inversa a bases de datos NoSQL, ¿qué pasaría cuando los datos personales están almacenados en documentos escritos en lenguaje natural?, ¿y si son imágenes o grabaciones de vídeo? Parece claro que estos datos no pueden ser tratados desde un enfoque basado en modelos. En este sentido, los recientes desarrollos en grandes modelos de lenguaje natural (LLM – *Large Language Models*) podrían llegar en el futuro a resultar útiles para poder aplicar reglas de privacidad, codificadas explícitamente como modelos, a la información personal descrita en lenguaje natural (contenida en mensajes, documentos, multimedia, etc.)
- Otra área que puede resultar incierta es el aprendizaje automático (*machine learning*), dentro del campo de la inteligencia artificial, ya que los 'modelos' de aprendizaje automático describen los *datos de manera estadística*, mientras que los modelos de MBSSE prescriben su estructura de manera determinista. Además, el carácter opaco de los modelos de aprendizaje automático dificulta los análisis que planteamos aquí (dentro de un contexto más amplio de los problemas de la falta de explicabilidad de los modelos de datos personales en relación con la protección de datos personales).

Por último, se debe plantear una mejora en la **preparación de los resultados** para ampliar su aceptación y adopción. Aunque los resultados presentados han demostrado la validez técnica de nuestro enfoque, sigue quedando trabajo por hacer de cara a asegurar su adopción. El uso de un enfoque de ingeniería de software y sistemas basada en modelos puede plantear algunas preocupaciones potenciales en cuanto a su idoneidad, ya que existen grupos de desarrolladores reacios a las prácticas que exige el modelado, al considerarlas complejas y contrarias a la productividad. Especialmente, entre las *comunidades de desarrollo denominado 'ágil'*, se discute la idoneidad general del uso de enfoques basados en modelos. No obstante, frente a esa idea de que los métodos basados en modelos resultan esotéricos, son obsoletos, o inapropiados para las tecnologías actuales (APIs REST, NoSQL, pilas MEAN/MERN, etc.) y las metodologías ágiles y *lean*, encontramos que, al contrario, las herramientas basadas en modelos gozan de salud bajo otros nombres como '*generadores de esquemas*', '*herramientas de bajo código*' (*low-code*), '*lenguajes específicos de dominio*' (DSL), o '*gemelos digitales*', y, precisamente, pueden ayudar a limitar las consecuencias más perniciosas de unos métodos 'ágiles' mal empleados, como la deuda técnica de requisitos extrafuncionales (incluidos los de privacidad y protección de datos). Por ello, consideramos que es necesaria una labor de *comunicación y evangelización*, para hacer entender que nuestro enfoque es compatible con estos otros enfoques actuales de desarrollo.

10.5 Conclusión

Para concluir, debemos recordar la relevancia y actualidad de la privacidad y protección de datos y cómo, al ser cuestiones inherentemente multidisciplinares, necesitan contar tanto con el ámbito legal como con el de la ingeniería (entre otros), para que nuestros ingenieros '*honestos pero despreocupados*', sin conocimientos legales ni pericia en las tecnologías de mejora de la privacidad, dispongan de métodos y herramientas que les hagan la vida más fácil al abordar la privacidad y la protección de datos.

Como se expuso en los primeros capítulos, la privacidad y la protección de datos plantean cuestiones que involucran —al menos— aspectos psicológicos, sociológicos, legales, comerciales, y tecnológicos, y, por lo tanto, solo se pueden abordar adecuadamente *desde una perspectiva interdisciplinar*. Sin embargo, en muchas ocasiones, estas disciplinas se ignoran entre ellas y trabajan en direcciones separadas, por lo que los investigadores y los profesionales deben tratar de crear puentes entre sí: así, en nuestra investigación, nos hemos esforzado por responder desde la perspectiva

de la ingeniería a algunos de los problemas que conlleva la introducción de los requisitos de privacidad y protección de datos en el desarrollo de software y sistemas, pero siempre teniendo en mente esas otras dimensiones. De hecho, a pesar de que nuestra formación proviene de la Ingeniería de Telecomunicación, hemos tratado de introducir puntos de vista multidisciplinares en nuestro propio trabajo, combinando métodos tradicionales en ciencias sociales y humanidades (análisis de contenido cualitativo), desarrollo de productos (modelos mentales) y enfoques de ingeniería de software y sistemas (ingeniería basada en modelos, teoría de la información). Hemos llevado a cabo análisis legales detallados, tanto desde la perspectiva interna del RGPD (junto con normas auxiliares, por ejemplo, las pautas y recomendaciones del Comité Europeo de Protección de datos o EDPB y el Grupo de Trabajo del Artículo 29 o WP29), como desde su relación con estándares industriales (ISO/IEC 29100, ISO/IEC 27550 e ISO/IEC 29134), y desde una perspectiva comparada de la legislación de la UE para otras categorías de requisitos. Asimismo, dentro de la ingeniería de software y sistemas específicamente, hemos abordado las cuestiones de privacidad y protección de datos en cinco subdisciplinas diferentes (gestión de riesgos, ingeniería de requisitos, diseño de sistemas basado en modelos, aseguramiento de sistemas e ingeniería de métodos). Y, algo no menos importante, distintas organizaciones participantes en el proyecto europeo PDP4E han trasladado varios de nuestros resultados a herramientas software que explotan comercialmente.

Podemos concluir que *la privacidad y la protección de datos han llegado para quedarse.* Los consumidores y las sociedades en su conjunto han dejado atrás una era de ingenuidad tecnológica, donde cualquier avance en los servicios de Internet se abrazaba abiertamente con una mezcla de la curiosidad infantil por un nuevo juguete y la esperanza adulta en una sociedad más equitativa; y, en su lugar, ahora se reciben a menudo con una desconfianza que vacila en entre la sospecha por defecto frente a cualquier nuevo servicio y la aquiescencia sumisa para utilizarlos a pesar de la convicción de que '*algo huele a podrido*' en Internet. La prueba de que entre el público hay una insatisfacción que se debe abordar es que, como se señalaba al comienzo de este trabajo, los gigantes tecnológicos han comenzado a hacer valer la privacidad y la protección de datos como valores diferenciales en la venta de sus servicios.

No hay duda de que dependemos más que nunca de los servicios de información que hacen uso de nuestros datos personales, tanto en el trabajo como en el ocio; sin embargo, las actitudes hacia muchos negocios de Internet *han pasado de la admiración a la desconsideración*. Tal vez el público no haya cambiado sus prácticas por cuestiones de privacidad (según la '*paradoja de la privacidad*', tan a menudo citada, y discutida hasta la saciedad), pero sí lo han hecho sus puntos de vista: la satisfacción parece haber dado paso al desprecio. En fechas recientes, cuando Facebook anunció el cambio del nombre de su empresa matriz a Meta, para centrarse en el '*metaverso*' (un mundo de realidad virtual habitado por avatares de los usuarios), surgieron críticas inmediatas sobre posibles problemas de privacidad. Si comparamos eso con el discurso público que dominó durante la anterior moda de los mundos virtuales, en 2007 (en el pico del uso de Second Life), podemos concluir que, efectivamente, *algo ha cambiado*.

En este contexto, somos conscientes de que siempre habrá individuos dispuestos a beneficiarse dañando la privacidad de los demás; y estamos seguros de que la sociedad seguirá lamentándose en este tira y afloja tan antiguo como la tecnología —o, al menos, desde los 130 años que han transcurrido desde el artículo de Warren y Brandeis—, y cuyo resultado final nadie conoce. No obstante, también sabemos que muchos otros individuos quieren *hacer negocios mientras respetan la privacidad*, ya sea por los efectos disuasorios del marco legal, o por la aspiración de mantener una buena reputación comercial, o simplemente por razones éticas (que también existen); pero quienes, a la vez, carecen del tiempo necesario, el conocimiento adecuado, o los recursos económicos para abordar adecuadamente estas preocupaciones. Es a ellos, gerentes, ingenieros y otros interesados, *a quienes dirigimos nuestras contribuciones*. (O, en los términos del modelo adversario que introdujimos en el Capítulo 2, apuntamos al ingeniero '*honesto*

pero despreocupado', independientemente de lo que el '*intermediario de datos sin escrúpulos'* esté dispuesto a hacer por su parte.)

Tradicionalmente, la ingeniería de privacidad y de protección de datos se había considerado como una disciplina menor, subordinada, en su caso (y según a quién se preguntase), al derecho de la privacidad o a la ingeniería de ciberseguridad. Pero el progreso del en las tecnologías de mejora de la privacidad (PET), junto con la madurez en la ingeniería de software y sistemas en general, implican que la ingeniería de privacidad y protección de datos ya está en condiciones de proporcionar su contribución *en pie de igualdad con otras disciplinas*. Ya era hora de que la ingeniería de privacidad y protección de datos consiguiera el papel que merece y, en nuestro trabajo, hemos dado un paso adelante con nuestras aportaciones para mostrar cómo está *alcanzando su mayoría de edad*: hemos operativizado conceptos abstractos y que a los oídos de los ingenieros suenan a jerga legal críptica, y los hemos hecho corresponder a otros términos habituales en las diferentes disciplinas de la ingeniería de software y sistemas, tales como requisitos, entropía de la información, índices de riesgo, clases, atributos, patrones, nodos de implementación, casos de aseguramiento, actividades o, en general, *modelos*. De hecho, prevemos que el uso de modelos bien diseñados desempeñará un papel relevante en el futuro de la ingeniería de privacidad y protección de datos y allanará el camino para una integración más fácil de las PET existentes y venideras en la práctica convencional de la ingeniería (incluso si los ingenieros no son muy conscientes de '*qué implica eso del modelado*'). Esperamos que nuestras iniciativas para abrir nuestro trabajo en el futuro a una comunidad más amplia de ingenieros den fe de ello.

Reiteramos que entendemos la privacidad y la protección de datos como cuestiones que se deben abordar desde una perspectiva interdisciplinar. Por ello, *lejos de caer en un solucionismo tecnológico*, avalamos la colaboración de investigadores y profesionales de todas las disciplinas involucradas, que pueden beneficiarse de la experiencia, el conocimiento y la sabiduría de los demás (como hemos hecho nosotros mismos aplicando métodos de investigación cualitativa y profundizando en textos jurídicos en una tesis doctoral de ingeniería). Por nuestra parte, intentamos contribuir a cerrar esa brecha desde el lado de la ingeniería de la privacidad, de modo que, de la mano de otras disciplinas, contribuya a un mundo donde la privacidad sea una preocupación del pasado y podamos '*superarla*', pero no porque esté muerta como decían algunos dirigentes corporativos, sino porque se pueda encontrar materializada allá donde uno mire.

Referencias

[1] Y. S. Martín García, 'Contributions to Privacy and Data Protection from a Multi-Disciplinary Model-Based Software and Systems Engineering Approach', Universidad Politécnica de Madrid, 2021. doi: 10.20868/UPM.thesis.69940.

[2] J. M. del Alamo, R. Trapero, Y. S. Martin, J. C. Yelmo, and N. Suri, 'Assessing privacy capabilities of cloud service providers', *IEEE Latin America Transactions*, vol. 13, no. 11, pp. 3634–3641, Nov. 2015, doi: 10.1109/TLA.2015.7387942.

[3] J. C. Caiza, Y. S. Martín, D. S. Guamán, J. M. Del Alamo, and J. C. Yelmo, 'Reusable Elements for the Systematic Design of Privacy-Friendly Information Systems: A Mapping Study', *IEEE Access*, vol. 7, pp. 66512–66535, 2019, doi: 10.1109/ACCESS.2019.2918003.

[4] Y. S. Martin, J. M. del Alamo, and J. C. Yelmo, 'Engineering Privacy Requirements: Valuable Lessons from Another Realm', in *1st International Workshop on Evolving Security and Privacy Requirements Engineering - ESPRE2014*, 2014. doi: 10.1109/ESPRE.2014.6890523.

[5] N. Notario *et al.*, 'Pripare: A new vision on engineering privacy and security by design', in *Cyber Security and Privacy Forum CSP 2014 (Communications in Computer and Information Science vol. 470)*, 2014. doi: 10.1007/978-3-319-12574-9_6.

[6] Y. S. Martín García, J. M. del Álamo Ramiro, and J. C. Yelmo Garcia, 'Toward a pattern language for privacy enhancing design techniques', in *Memorias del VII Congreso Iberoamericano de Telemática CITA2015*, Á. Rendón Gallón, J. C. Yelmo García, and G. Ramirez-Gonzalez, Eds., 2015, pp. 173–180. [Online]. Available: http://web.archive.org/web/20160415104125/http://cita2015.co/wp-content/uploads/2015/07/LibroCITA2015.pdf

[7] Y. S. Martín and J. M. del Álamo, 'A metamodel for privacy engineering methods', in *Proceedings of the 3rd International Workshop on Privacy Engineering co-located with 38th IEEE Symposium on Security and Privacy (S&P 2017) (CEUR Workshop Proceedings Vol-1873)*, 2017, pp. 41–48. Accessed: Nov. 16, 2021. [Online]. Available: http://ceur-ws.org/Vol-1873/IWPE17_paper_24.pdf

[8] N. Notario *et al.*, 'PRIPARE: Integrating privacy best practices into a privacy engineering methodology', in *Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015*, 2015, pp. 151–158. doi: 10.1109/SPW.2015.22.

[9] J. M. del Alamo, Y. S. Martín, and J. C. Caiza, 'Towards Organizing the Growing Knowledge on Privacy Engineering', in *IFIP International Summer School on Privacy and Identity Management Privacy and Identity 2017: Privacy and Identity Management. The Smart Revolution (IFIP Advances in Information and Communication Technology vol. 526)*, Springer, Cham, Sep. 2017, pp. 15–24. doi: 10.1007/978-3-319-92925-5_2.

[10] J. C. Caiza, Y.-S. Martín, J. M. del Alamo, and D. S. Guamán, 'Organizing Design Patterns for Privacy: A Taxonomy of Types of Relationships', in *EuroPLoP '17: Proceedings of the 22nd European Conference on Pattern Languages of Programs*, Association for Computing Machinery, Jul. 2017, p. 1. doi: 10.1145/3147704.3147739.

[11] A. Kung *et al.*, 'A Privacy Engineering Framework for the Internet of Things', in *Data Protection and Privacy: (In)visibilities and Infrastructures - (Computers, Privacy and Data Protection 2016)*, Springer, Cham, 2017, pp. 163–202. doi: 10.1007/978-3-319-50796-5_7.

[12] M. Colesky, J. C. Caiza, J. M. del Lamo, J. H. Hoepman, and Y. S. Martín, ‘A system of privacy patterns for user control’, in *SAC ’18: Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, New York, NY, USA: Association for Computing Machinery, Apr. 2018, pp. 1150–1156. doi: 10.1145/3167132.3167257.

[13] Y. S. Martin and A. Kung, ‘Methods and Tools for GDPR Compliance Through Privacy and Data Protection Engineering’, *Proceedings - 3rd IEEE European Symposium on Security and Privacy Workshops, EURO S and PW 2018*, pp. 108–111, Jul. 2018, doi: 10.1109/EUROSPW.2018.00021.

[14] R. M. de Carvalho *et al.*, ‘Protecting Citizens’ Personal Data and Privacy: Joint Effort from GDPR EU Cluster Research Projects’, *SN Computer Science 2020 1:4*, vol. 1, no. 4, pp. 1–16, Jun. 2020, doi: 10.1007/S42979-020-00218-8.

[15] G. Pedroza, V. Muntes-Mulero, Y. S. Martin, and G. Mockly, ‘A Model-based Approach to Realize Privacy and Data Protection by Design’, *2021 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pp. 332–339, Sep. 2021, doi: 10.1109/EUROSPW54576.2021.00042.

[16] A. Ruiz *et al.*, ‘Modeling ecosystems of reference frameworks for assurance: a case on privacy impact assessment regulation and guidelines’, *Softw Syst Model*, Nov. 2022, doi: 10.1007/s10270-022-01061-6.

[17] M. Andreessen, ‘Why Software Is Eating The World’, *Wall Street Journal*, pp. 1–5, 2011, [Online]. Available: http://online.wsj.com/article/SB10001424053111903480904576512250915629460.html

[18] M. Palmer, ‘Data is the New Oil’, ANA Marketing Maestros. Accessed: Sep. 05, 2021. [Online]. Available: https://ana.blogs.com/maestros/2006/11/data_is_the_new.html

[19] J. Thorp, ‘Big Data Is Not the New Oil’, *Havard Business Review*, Nov. 30, 2012. Accessed: Oct. 20, 2021. [Online]. Available: https://hbr.org/2012/11/data-humans-and-the-new-oil

[20] R. Serra and C. F. Schoolman, ‘Television Delivers People’, 1973. Accessed: Oct. 20, 2021. [Online]. Available: https://www.youtube.com/watch?v=LvZYwaQlJsg

[21] G. O’Toole, ‘You’re Not the Customer; You’re the Product’, Quote Investigator. Accessed: Oct. 20, 2021. [Online]. Available: https://quoteinvestigator.com/2017/07/16/product/#note-16480-1

[22] P. Sprenger, ‘Sun on Privacy: “Get Over It”’, *Wired*, Jan. 26, 1999. Accessed: Oct. 20, 2021. [Online]. Available: https://www.wired.com/1999/01/sun-on-privacy-get-over-it/

[23] B. Johnson, ‘Privacy no longer a social norm, says Facebook founder’, *The Guardian*, 2010. Accessed: Oct. 20, 2021. [Online]. Available: https://www.theguardian.com/technology/2010/jan/11/facebook-privacy

[24] *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. European Union: European Parliament, European Council, 2016, p. L119/1-L119/88. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj

[25] L. Lessig, ‘Code Is Law - On Liberty and Cyberspace’, *Hardvard Magazine*, no. 1, Jan. 2000, Accessed: Oct. 21, 2021. [Online]. Available: https://www.harvardmagazine.com/2000/01/code-is-law-html

[26] R. Balebako, A. Marsh, J. Lin, J. Hong, and L. Faith Cranor, 'The Privacy and Security Behaviors of Smartphone App Developers', in *Proceedings 2014 Workshop on Usable Security*, 2014. doi: 10.14722/usec.2014.23006.

[27] R. Balebako and L. Cranor, 'Improving App Privacy: Nudging App Developers to Protect User Privacy', *IEEE Secur Priv*, vol. 12, no. 4, pp. 55–58, Jul. 2014, doi: 10.1109/MSP.2014.70.

[28] S. Jain and J. Lindqvist, 'Should I Protect You? Understanding Developers' Behavior to Privacy-Preserving APIs', in *Proceedings 2014 Workshop on Usable Security*, 2014. doi: 10.14722/usec.2014.23045.

[29] Y. S. van der Sype and W. Maalej, 'On lawful disclosure of personal user data: What should app developers do?', in *2014 IEEE 7th International Workshop on Requirements Engineering and Law, RELAW 2014 - Proceedings*, 2014, pp. 25–34. doi: 10.1109/RELAW.2014.6893479.

[30] S. Gürses and J. M. del Alamo, 'Privacy Engineering: Shaping an Emerging Field of Research and Practice', *IEEE Secur Priv*, vol. 14, no. 2, pp. 40–46, Mar. 2016, doi: 10.1109/MSP.2016.37.

[31] G. Danezis, J. Domingo-Ferrer, R. Tirtea, and S. Schiffner, 'Privacy and Data Protection by Design – from policy to engineering', Dec. 2014. doi: 10.2824/38623.

[32] Article 29 Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679', Madrid, Apr. 2017. [Online]. Available: ec.europa.eu/newsroom/document.cfm?doc_id=44137%0A

[33] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, 'Systematic mapping studies in software engineering', in *12th International Conference on Evaluation and Assessment in Software Engineering, EASE 2008*, 2008. doi: 10.14236/ewic/ease2008.8.

[34] K. Petersen, S. Vakkalanka, and L. Kuzniarz, 'Guidelines for conducting systematic mapping studies in software engineering: An update', in *Information and Software Technology*, 2015. doi: 10.1016/j.infsof.2015.03.007.

[35] A. R. Hevner, S. T. March, J. Park, and S. Ram, 'Design science in information systems research', *MIS Q*, vol. 28, no. 1, 2004, doi: 10.2307/25148625.

[36] S. D. Warren and L. D. Brandeis, 'The Right to Privacy', *Harv Law Rev*, vol. 4, no. 5, p. 193, Dec. 1890, doi: 10.2307/1321160.

[37] A. F. Westin, *Privacy and freedom*, 1st ed. Atheneum, 1967. Accessed: Oct. 31, 2021. [Online]. Available: https://openlibrary.org/works/OL4083191W/Privacy_and_freedom?edition=privacyfreedom00west

[38] Irwin. Altman, *The environment and social behavior : privacy, personal space, territory, crowding*. Brooks/Cole Pub. Co, 1975.

[39] I. Altman, 'Toward a Transactional Perspective', in *Environment and Behavior Studies*, 1990. doi: 10.1007/978-1-4684-7944-7_10.

[40] H. Nissenbaum, 'Privacy as contextual integrity', *Washington Law Review*, vol. 79, no. 1, pp. 101–139, 2004, doi: 10.1109/SP.2006.32.

[41] H. Nissenbaum, *Privacy in Context*. Stanford University Press, 2009. doi: 10.1515/9780804772891.

[42] R. Clarke, 'Data Surveillance: Theory, Practice & Policy.', 1995. [Online]. Available: http://www.rogerclarke.com/DV/PhD.html

[43] R. Clarke, 'Introduction to Dataveillance and Information Privacy, and Definitions of Terms', Jul. 2016. [Online]. Available: http://www.rogerclarke.com/DV/Intro.html

[44] R. L. Finn, D. Wright, and M. Friedewald, ‘Seven types of privacy’, in *European Data Protection: Coming of Age*, 2013. doi: 10.1007/978-94-007-5170-5_1.

[45] International Security Trust and Privacy Alliance (ISTPA), ‘ISTPA Privacy Framework v. 1.1’, 2002. [Online]. Available: http://istpa.org/pdfs/ISTPAPrivacyFrameworkV1.1.pdf

[46] T. & P. A. International Security, ‘Privacy Management Reference Model A framework for resolving privacy policy requirements into operational privacy services and functions - Version 2.0’, 2009. Accessed: Oct. 31, 2021. [Online]. Available: http://xml.coverpages.org/ISTPA-PrivacyManagementReferenceModelV20.pdf

[47] J. Sabo, M. Drgon, and G. Magnuson, ‘Privacy Management Reference Model and Methodology (PMRM) Version 1.0 - Committee Specification 02’, Jul. 2016. [Online]. Available: http://docs.oasis-open.org/pmrm/PMRM/v1.0/PMRM-v1.0.html

[48] National Institute of Standards and Technology (NIST), ‘NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management, Version 1.0’, Gaithersburg, MD, Jan. 2020. doi: 10.6028/NIST.CSWP.01162020.

[49] A. S. Patrick and S. Kenny, ‘From Privacy Legislation to Interface Design: Implementing Information Privacy in Human-Computer Interactions’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 2760, pp. 107–124, 2003, doi: 10.1007/978-3-540-40956-4_8.

[50] D. N. Jutla and P. Bodorik, ‘Sociotechnical architecture for online privacy’, *IEEE Security and Privacy*, vol. 3. pp. 29–39, 2005. doi: 10.1109/MSP.2005.50.

[51] C. di Iorio *et al.*, ‘Cross-border flow of health information: is “privacy by design” enough? Privacy performance assessment in EUBIROD’, *Eur J Public Health*, vol. 23, no. 2, pp. 247–253, Apr. 2013, doi: 10.1093/EURPUB/CKS043.

[52] A. Cavoukian *et al.*, ‘Privacy by Design Documentation for Software Engineers Version 1.0 - Committee Specification Draft 01’, Jun. 2014. [Online]. Available: http://docs.oasis-open.org/pbd-se/pbd-se/v1.0/csd01/pbd-se-v1.0-csd01.html

[53] A. Cavoukian *et al.*, ‘Annex Guide to Privacy by Design Documentation for Software Engineers Version 1.0 - Committee Note Draft 01’, Jun. 2014. [Online]. Available: http://docs.oasis-open.org/pbd-se/pbd-se-annex/v1.0/pbd-se-annex-v1.0.html

[54] W. L. Prosser, ‘Privacy’, *Calif. L. Rev.. California Law Review*, vol. 48, no. 3, pp. 383–423, Aug. 1960, [Online]. Available: http://lawcat.berkeley.edu/record/1109651

[55] D. J. Solove, ‘A Taxonomy of Privacy’, *Univ PA Law Rev*, vol. 154, no. 3, pp. 477–560, 2006, doi: 10.2307/40041279.

[56] D. Solove, *Understanding Privacy*. London (United Kingdom): Hardvard University Press, 2008.

[57] D. K. Citron and D. J. Solove, ‘Privacy Harms’, *Boston University Law Review*, vol. 102, no. [In press], 2022, doi: 10.2139/SSRN.3782222.

[58] A. Pfitzmann and M. Hansen, ‘A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management’, *Technical University Dresden*, pp. 1–98, 2010, doi: 10.1.1.154.635.

[59] K. Wuyts and W. Joosen, ‘LINDDUN privacy threat modeling: a tutorial’. Department of Computer Science, KU Leuven, 2015.

[60] ISO/IEC JTC 1/SC 27, ‘ISO/IEC 29100:2011 - Information technology — Security techniques — Privacy framework’, Geneva (CH), 2011. [Online]. Available: http://standards.iso.org/ittf/PubliclyAvailableStandards/c045123_ISO_IEC_29100_2011.zip

[61] European Parliament and Council of the European Union, *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Data Protection Directive)*. European Union: European Parliament, European Council, 1995, p. L 281/31-L 281/50. [Online]. Available: https://eur-lex.europa.eu/eli/dir/1995/46/oj

[62] S. T. Margulis, 'Three Theories of Privacy: An Overview', in *Privacy Online*, 2011. doi: 10.1007/978-3-642-21521-6_2.

[63] J. W. DeCew and A. L. Allen, 'Uneasy Access: Privacy for Women in a Free Society.', *Philos Rev*, vol. 101, no. 3, 1992, doi: 10.2307/2186088.

[64] D. K. Mulligan, C. Koopman, and N. Doty, 'Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2083, p. 20160118, 2016, doi: 10.1098/rsta.2016.0118.

[65] W. B. Gallie, 'IX.—Essentially Contested Concepts', *Proceedings of the Aristotelian Society*, vol. 56, no. 1, Jun. 1956, doi: 10.1093/aristotelian/56.1.167.

[66] Q. Do, B. Martini, and K. K. R. Choo, 'The role of the adversary model in applied security research', *Comput Secur*, vol. 81, pp. 156–181, Mar. 2019, doi: 10.1016/J.COSE.2018.12.002.

[67] European Parliament and Council of the European Union, *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)* . 2016, p. L 119/1-L 119/88. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj

[68] A. E. Waldman, 'Privacy Law's False Promise', *SSRN Electronic Journal*, 2019, doi: 10.2139/ssrn.3339372.

[69] A. E. Waldman, 'Privacy's Law of Design', *UC Irvine Law Review*, vol. 9, no. 5, pp. 1239–1288, Jul. 2019, [Online]. Available: https://scholarship.law.uci.edu/ucilr/vol9/iss5/8/

[70] A. Cavoukian, 'Privacy by Design The 7 Foundational Principles', Toronto, Ontario (Canada), 2009. Accessed: May 29, 2014. [Online]. Available: http://www.privacybydesign.ca/content/uploads/2009/08/7foundationalprinciples.pdf

[71] A. Cavoukian *et al.*, 'Resolution on Privacy by Design', in *The 32nd International Conference of Data Protection and Privacy Commissioners*, Jerusalem (Israel), Oct. 2010. Accessed: Oct. 21, 2021. [Online]. Available: https://globalprivacyassembly.org/wp-content/uploads/2015/02/32-Conference-Israel-resolution-on-Privacy-by-Design.pdf

[72] L. Fritsch, 'State of the Art of Privacy-enhancing Technology (PET) - Deliverable D2.1 of the PETweb project', Nov. 2007. Accessed: Oct. 22, 2021. [Online]. Available: https://projects.nr.no/en/nrpublication?query=/file/4589/Fritsch_-_State_of_the_Art_of_Privacy-enhancing_Technology.pdf

[73] S. Fischer-Hübner, C. Hoofnagle, K. Rannenberg, M. Waidner, I. Krontiris, and M. Marhöfer, 'Online Privacy: Towards Informational Self-Determination on the Internet Edited by Executive Summary', *Tagungsübersicht*, vol. 1, no. 2, 2011.

[74] I. S. Rubinstein and N. Good, 'The trouble with Article 25 (and how to fix it): the future of data protection by design and default', *International Data Privacy Law*, vol. 10, no. 1, pp. 37–56, Feb. 2020, doi: 10.1093/IDPL/IPZ019.

[75] I. S. Rubinstein, 'Regulating Privacy By Design.', *Berkeley Technol Law J*, vol. 26, no. 3, 2011.

[76] M. Hildebrandt and M. Meints, ‘D7.7: RFID, Profiling, and AmI’, Aug. 2006. Accessed: Oct. 22, 2021. [Online]. Available: http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp7-del7.7.RFID_Profiling_AMI.pdf

[77] C. Zimmermann, ‘A Categorization of Transparency-Enhancing Technologies’, Jul. 2015. Accessed: Oct. 22, 2021. [Online]. Available: https://arxiv.org/abs/1507.04914v2

[78] H. Hedbom, ‘A Survey on Transparency Tools for Enhancing Privacy’, in *IFIP Summer School on the Future of Identity in the Information Society Privacy and Identity 2008: IFIP Advances in Information and Communication Technology*, Springer, Berlin, Heidelberg, 2008, pp. 67–82. doi: 10.1007/978-3-642-03315-5_5.

[79] F. Karegar, ‘Towards Improving Transparency, Intervenability, and Consent in HCI’, 2018. Accessed: Oct. 23, 2021. [Online]. Available: http://kau.diva-portal.org/smash/record.jsf?pid=diva2%3A1180397&dswid=-9000

[80] M. Hansen, ‘Marrying Transparency Tools with User-Controlled Identity Management’, in *IFIP International Summer School on the Future of Identity in the Information Society Privacy and Identity 2007: IFIP International Federation for Information Processing*, Springer, Boston, MA, 2007, pp. 199–220. doi: 10.1007/978-0-387-79026-8_14.

[81] S. Spiekermann and L. F. Cranor, ‘Engineering privacy’, *IEEE Transactions on Software Engineering*, vol. 35, no. 1, pp. 67–82, 2009, doi: 10.1109/TSE.2008.88.

[82] International Association of Privacy Professionals (IAPP), ‘IAPP Privacy Tech Vendor Report’, International Association of Privacy Professionals (IAPP). Accessed: Oct. 23, 2021. [Online]. Available: https://iapp.org/resources/article/privacy-tech-vendor-report/

[83] A. Cavoukian, S. Shapiro, and R. J. Cronk, ‘Privacy Engineering: Proactively Embedding Privacy by Design’, Toronto, 2014. Accessed: Jan. 27, 2018. [Online]. Available: https://www.ipc.on.ca/wp-content/uploads/Resources/pbd-priv-engineering.pdf

[84] S. Gurses, ‘Can You Engineer Privacy?’, *Commun ACM*, vol. 57, no. 8, pp. 20–23, 2014, doi: 10.1145/2633029.

[85] M. F. Dennedy, J. Fox, and T. R. Finneran, *The privacy engineer’s manifesto: Getting from policy to code to QA to value*. 2014. doi: 10.1007/978-1-4302-6356-2.

[86] D. I. Oliver, *Privacy Engineering: A Dataflow and Ontological Approach*, 1st ed. North Charleston, SC, USA: CreateSpace Independent Publishing Platform, 2014.

[87] S. Gürses, C. Gonzalez Troncoso, and C. Diaz, ‘Engineering privacy by design’, *Computers, Privacy & Data Protection*, 2011.

[88] S. Gürses, C. Troncoso, and C. Diaz, ‘Engineering privacy by design reloaded’, in *Amsterdam Privacy Conference*, 2015.

[89] M. Hansen, M. Jensen, and M. Rost, ‘Protection goals for privacy engineering’, in *Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015*, 2015, pp. 159–166. doi: 10.1109/SPW.2015.13.

[90] The MITRE Corporation, ‘MITRE Privacy Engineering Framework and Lifecycle Adaptation Guide MITRE Privacy Engineering - Version 2’, Sep. 2019. [Online]. Available: https://www.mitre.org/sites/default/files/publications/pr-19-00598-5-privacy-engineering-framework-v2.pdf

[91] ‘IEEE Position Statement In Support of Privacy Engineering’, Nov. 2018. doi: 10.1109/CloudCom.2014.123.

[92] N. van Dijk, A. Tanas, K. Rommetveit, and C. Raab, ‘Right engineering? The redesign of privacy and personal data protection’,

https://doi.org/10.1080/13600869.2018.1457002, vol. 32, no. 2–3, pp. 230–256, Sep. 2018, doi: 10.1080/13600869.2018.1457002.

[93] Interaction Design Foundation (IxDF), 'A Very Useful Work of Fiction – Mental Models in Design'. Accessed: Sep. 06, 2021. [Online]. Available: https://www.interaction-design.org/literature/article/a-very-useful-work-of-fiction-mental-models-in-design

[94] Interaction Design Foundation (IxDF), '23. Mental models', *The Glossary of Human Computer Interaction*. 2017. Accessed: Sep. 06, 2021. [Online]. Available: https://www.interaction-design.org/literature/book/the-glossary-of-human-computer-interaction/mental-models

[95] 'Mental Models', *Business Dictionary*. WebFinance, Inc., 2007. Accessed: Sep. 06, 2021. [Online]. Available: http://www.businessdictionary.com/definition/mental-models.html

[96] I. Young, *Mental Models: Aligning Design Strategy with Human Behavior*. Rosenfield Media, 2008.

[97] '2021 Privacy Tech Vendor Report', 2021. [Online]. Available: https://iapp.org/media/pdf/resource_center/2021TechVendorReport.pdf

[98] A. Rodrigues Da Silva, 'Model-driven engineering: A survey supported by the unified conceptual model', *Comput Lang Syst Struct*, vol. 43, pp. 139–155, Oct. 2015, doi: 10.1016/J.CL.2015.06.001.

[99] 'MDA Foundation Model - Cambridge MA AB Edits draft, clean - OMG Document -- ormsc/10-09-06', Sep. 2010. Accessed: Sep. 08, 2021. [Online]. Available: https://www.omg.org/cgi-bin/doc?ormsc/10-09-06

[100] INCOSE, 'Systems Engineering Vision 2020, INCOSE-TP-2004-004-02, Version 2.03', Sep. 2007.

[101] 'P24641_D2_CD, Jun 2020 - ISO/IEC/IEEE Draft International Standard - Systems and software engineering - Methods and Tools for Model-based Systems and Software Engineering', Jun. 2020, Accessed: Sep. 09, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9123713

[102] J. Fabry, T. Dinkelaker, J. Noyé, and É. Tanter, 'A Taxonomy of Domain-Specific Aspect Languages', *ACM Computing Surveys (CSUR)*, vol. 47, no. 3, Feb. 2015, doi: 10.1145/2685028.

[103] Y. Zhang and B. M. Wildemuth, 'Qualitative Analysis of Content', *Applications of Social Research Methods to Questions in Information and Library Science*, 2009, doi: 10.1002/hbm.20661.

[104] P. Mayring, 'Qualitative Content Analysis', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 1, no. 2, p. Art. 20, Jun. 2000, doi: 10.17169/FQS-1.2.1089.

[105] M. Bengtsson, 'How to plan and perform a qualitative study using content analysis', *NursingPlus Open*, vol. 2, pp. 8–14, Jan. 2016, doi: 10.1016/J.NPLS.2016.01.001.

[106] D. R. Thomas, 'A General Inductive Approach for Analyzing Qualitative Evaluation Data':, *http://dx.doi.org/10.1177/1098214005283748*, vol. 27, no. 2, pp. 237–246, Jun. 2016, doi: 10.1177/1098214005283748.

[107] K. Williamson, L. M. Given, and P. Scifleet, 'Qualitative data analysis', in *Research Methods: Information, Systems, and Contexts: Second Edition*, Chandos Publishing, 2018, pp. 453–476. doi: 10.1016/B978-0-08-102220-7.00019-4.

[108] H. F. Hsieh and S. E. Shannon, 'Three approaches to qualitative content analysis', *Qual Health Res*, vol. 15, no. 9, pp. 1277–1288, Nov. 2005, doi: 10.1177/1049732305276687.

[109] B. L. Berg and H. Lune, *Qualitative research methods for the social sciences*, 9th-Global ed., vol. 7, no. 4. Pearson, 2017.

[110] S. Elo and H. Kyngäs, 'The qualitative content analysis process', *J Adv Nurs*, vol. 62, no. 1, pp. 107–115, Apr. 2008, doi: 10.1111/j.1365-2648.2007.04569.x.

[111] J. H. Hoepman, 'Privacy design strategies', *arXiv preprint arXiv:1210.6621*, p. 12, 2012, [Online]. Available: http://arxiv.org/abs/1210.6621

[112] J. H. Hoepman, 'Privacy design strategies', in *ICT Systems Security and Privacy Protection. SEC 2014. IFIP Advances in Information and Communication Technology, vol 428.*, N. Cuppens-Boulahia, F. Cuppens, S. Jajodia, A. Abou El Kalam, and T. Sans, Eds., Springer, 2014, pp. 446–459. doi: 10.1007/978-3-642-55415-5_38.

[113] J.-H. Hoepman, 'Privacy Design Strategies (The Little Blue Book)', 2019. Accessed: Jul. 23, 2019. [Online]. Available: https://www.cs.ru.nl/~jhh/publications/pds-booklet.pdf

[114] ISO/IEC JTC 1/SC 27 Information security cybersecurity and privacy protection, 'ISO/IEC TR 27550:2019 - Information technology — Security techniques — Privacy engineering for system life cycle processes', Sep. 2019. [Online]. Available: https://www.iso.org/standard/72024.html

[115] A. Cooper, M. Hansen, R. Smith, and H. Tschofenig, 'Privacy Terminology and Concepts', Accessed: Jul. 23, 2019. [Online]. Available: https://tools.ietf.org/html/draft-iab-privacy-terminology-01

[116] A. Cooper *et al.*, 'Privacy Considerations for Internet Protocols- RFC 6973', Jul. 2013. Accessed: Nov. 09, 2021. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc6973

[117] 'Common Criteria for Information Technology Security Evaluation Part 2: Security functional components', 2017. Accessed: Jul. 23, 2019. [Online]. Available: https://www.commoncriteriaportal.org/files/ccfiles/CCPART2V3.1R5.pdf

[118] ISO/IEC JTC 1/SC 27 Information security cybersecurity and privacy protection, 'ISO - ISO/IEC TS 19608:2018 - Guidance for developing security and privacy functional requirements based on ISO/IEC 15408', Oct. 2018. [Online]. Available: https://www.iso.org/standard/65459.html

[119] A. Cavoukian and CACM Staff, 'Operationalizing privacy by design', *Commun ACM*, vol. 55, no. 9, p. 7, Sep. 2012, doi: 10.1145/2330667.2330669.

[120] A. Cavoukian, 'Operationalizing Privacy by Design: A Guide to Implementing Strong Privacy Practices', Toronto, Ontario (Canada), Dec. 2012. Accessed: Nov. 14, 2021. [Online]. Available: https://collections.ola.org/mon/26012/320221.pdf

[121] Y. Y. Haimes, 'Risk of Extreme Events and the Fallacy of Expected Value', in *Risk Modeling, Assessment, and Management*, John Wiley & Sons, Ltd, 2009, pp. 325–374. doi: 10.1002/9780470422489.CH8.

[122] Y. Y. Haimes, J. H. Lambert, and D. Li, 'RISK OF EXTREME EVENTS IN A MULTIOBJECTIVE FRAMEWORK', *JAWRA Journal of the American Water Resources Association*, vol. 28, no. 1, 1992, doi: 10.1111/j.1752-1688.1992.tb03165.x.

[123] T. L. Saaty, 'How to make a decision: The analytic hierarchy process', *Eur J Oper Res*, vol. 48, no. 1, pp. 9–26, Sep. 1990, doi: 10.1016/0377-2217(90)90057-I.

[124] E. H. Forman and S. I. Gass, 'The Analytic Hierarchy Process—An Exposition', *https://doi.org/10.1287/opre.49.4.469.11231*, vol. 49, no. 4, pp. 469–486, Aug. 2001, doi: 10.1287/OPRE.49.4.469.11231.

[125] J. M. del Álamo Ramiro, R. Trapero Burgos, Y. S. Martín García, J. C. Yelmo Garcia, and N. Suri, 'Quantitative assessment and comparison of cloud service

providers' privacy practices', in *Memorias del VII Congreso Iberoamericano de Telemática CITA2015*, Á. Rendón Gallón, J. C. Yelmo García, and G. Ramirez-Gonzalez, Eds., 2015, pp. 165–172. [Online]. Available: https://oa.upm.es/41943/

[126] C. Quelle, 'Enhancing compliance under the general data protection regulation: The risky upshot of the accountability- A nd risk-based approach', *European Journal of Risk Regulation*, vol. 9, no. 3, pp. 502–526, Sep. 2018, doi: 10.1017/ERR.2018.47.

[127] R. Gellert, 'Why the GDPR risk-based approach is about compliance risk, and why it's not a bad thing', in *Trends and Communities of legal informatics: IRIS 2017 - Proceedings of the 20th International Legal Informatics Symposium*, E. , Schweighofer, F. Kummer, and C. Sorge, Eds., Österreichische Computer Gesellschaft, 2017, pp. 527–532.

[128] R. Gellert, 'Why the GDPR risk-based approach is about compliance risk, and why it's not a bad thing', *Jusletter IT*, no. February, 2017.

[129] Florian Stahl *et al.*, 'OWASP Top 10 Privacy Risks'. Accessed: Sep. 06, 2021. [Online]. Available: https://owasp.org/www-project-top-10-privacy-risks/

[130] M. C. Oetzel and S. Spiekermann, 'A systematic methodology for privacy impact assessments: a design science approach', *European Journal of Information Systems*, vol. 23, no. 2, pp. 126–150, 2014, doi: 10.1057/ejis.2013.18.

[131] B. Appleton, 'Patterns and Software: Essential Concepts and Terminology', Feb. 2000. Accessed: Nov. 21, 2021. [Online]. Available: http://www.bradapp.com/docs/patterns-intro.html

[132] G. Meszaros and J. Doble, 'A pattern language for pattern writing', in *Pattern languages of program design 3*, Addison-Wesley Longman Publishing Co., Inc., 1997, pp. 529–574. doi: 10.5555/273448.273487.

[133] J. Noble, 'Classifying relationships between object-oriented design patterns', in *Proceedings - 1998 Australian Software Engineering Conference, ASWEC 1998*, Institute of Electrical and Electronics Engineers Inc., 1998, pp. 98–107. doi: 10.1109/ASWEC.1998.730917.

[134] 'privacypatterns.eu - collecting patterns for better privacy'. Accessed: Feb. 19, 2017. [Online]. Available: https://privacypatterns.eu/

[135] 'NATO Standard AEP-67 Engineering for System Assurance in NATO Programmes. Edition B Version 1', 2017.

[136] '15026-4-2021 - ISO/IEC/IEEE International Standard - Systems and software engineering--Systems and software assurance -- Part 4 Assurance in the life cycle.', IEEE, 2021. doi: 10.1109/IEEESTD.2021.9444258.

[137] cybersecurity and privacy protection ISO/IEC JTC 1/SC 27 Information security, 'ISO/IEC 29134:2017 - Information technology -- Security techniques -- Guidelines for privacy impact assessment'. International Organization for Standardization (ISO), p. 43, 2017. Accessed: Jan. 27, 2018. [Online]. Available: https://www.iso.org/standard/62289.html

[138] D. P. and C.-S. in the S. G. E. Smart Grid Task Force 2012-14 - Expert Group 2: Regulatory Recommendations for Privacy, 'Data Protection Impact Assessment Template for Smart Grid and Smart Metering systems', 2018. [Online]. Available: https://ec.europa.eu/energy/sites/default/files/documents/dpia_for_publication_2018.pdf

[139] S. Brinkkemper, 'Method Engineering: Engineering of Information Systems Development Methods and Tools', *Inf Softw Technol*, vol. 38, no. 4, pp. 275–280, 1996, doi: 10.1016/0950-5849(95)01059-9.

[140] ISO/IEC JTC 1/SC 27, 'ISO/IEC 24744:2014. Software engineering -- Metamodel for development methodologies', Geneva (CH), 2014.

[141] I. Jacobson, G. Booch, and J. Rumbaugh, *The unified software development process*. Reading (Massachusetts, U.S.): Addison-Wesley, 1999.

[142] OPEN Process Framework Repository Organization (OPFRO), 'OPEN Process Framework Repository Organization (OPFRO)'. [Online]. Available: https://www.opfro.org/

[143] B. Henderson-Sellers, J. Ralyté, P. J. Ågerfalk, and M. Rossi, 'Processes for Creating a Methodology from Method Parts', *Situational Method Engineering*, pp. 133–167, 2014, doi: 10.1007/978-3-642-41467-1_6.

[144] ISO/TC 69/SC 7 Applications of statistical and related techniques for the implementation of Six Sigma, 'ISO 13053-2:2011 - Quantitative methods in process improvement — Six Sigma — Part 2: Tools and techniques', Sep. 2011. Accessed: Nov. 02, 2021. [Online]. Available: https://www.iso.org/standard/52902.html

[145] Agencia Española de Protección de Datos, 'Gestión del riesgo y evaluación de impacto en tratamientos de datos personales', 2021. [Online]. Available: https://www.aepd.es/documento/gestion-riesgo-y-evaluacion-impacto-en-tratamientos-datos-personales.pdf

[146] Prokopios Drogkaris and Monika Adamczyk, 'Ingeniería de la Protección de Datos - De la teoría a la práctica', 2023. [Online]. Available: https://www.aepd.es/documento/enisa-ingenieria-de-la-proteccion-de-datos.pdf

[147] Agencia Española de Protección de Datos, 'Guía de Privacidad desde el Diseño', 2019. [Online]. Available: https://www.aepd.es/documento/guia-privacidad-desde-diseno.pdf

[148] European Parliament and Council of the European Communities, *Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety*. 2002, p. L 11/4-L 11/17. [Online]. Available: http://data.europa.eu/eli/dir/2001/95/oj

[149] European Parliament and Council of the European Union, *Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies*. 2016, p. L 327/1-L 327/15. [Online]. Available: https://eur-lex.europa.eu/eli/dir/2016/2102/oj

[150] European Parliament and Council of the European Communities, *Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services*. 2019, p. L 151/70-L 151/115. [Online]. Available: http://data.europa.eu/eli/dir/2019/882/oj

[151] European Parliament and Council of the European Union, *Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC*. 2017, p. L 117/1-L 117/175. [Online]. Available: http://data.europa.eu/eli/reg/2017/745/oj

[152] European Parliament and Council of the European Union, *Regulation (EU) 2016/425 of the European Parliament and of the Council of 9 March 2016 on personal protective equipment and repealing Council Directive 89/686/EEC (Text with EEA relevance)*. 2016, p. L 81/51-L 81/98. [Online]. Available: http://data.europa.eu/eli/reg/2016/425/oj

[153] European Parliament and Council of the European Union, *Directive 2010/75/EU of the European Parliament and of the Council of 24 November 2010 on industrial emissions (integrated pollution prevention and control)*. 2010, p. L 334/17-L 334/119. [Online]. Available: http://data.europa.eu/eli/dir/2010/75/oj

[154] European Parliament and Council of the European Union, *Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)*. European Union, 2022, p. L333/80-L333/152. [Online]. Available: http://data.europa.eu/eli/dir/2022/2555/oj

[155] European Parliament and Council of the European Union, *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)*. European Union, 2022, p. L277/1-L277/102. [Online]. Available: http://data.europa.eu/eli/reg/2022/2065/oj

[156] European Parliament and Council of the European Union, *Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011*. European Union, 2022, p. L333/1-L333/79. [Online]. Available: http://data.europa.eu/eli/reg/2022/2554/oj

[157] P. Foitzik, 'What you must know about "third parties" under GDPR and CCPA', International Association of Privacy Professionals (IAPP). Accessed: Nov. 10, 2021. [Online]. Available: https://iapp.org/news/a/what-you-must-know-about-third-parties-under-the-gdpr-ccpa/