Universidad Politécnica de Madrid

Escuela Técnica Superior de Ingenieros de
Telecomunicación

**CONTRIBUCIÓN AL DISEÑO DE SISTEMAS
RESPETUOSOS CON LA PRIVACIDAD USANDO
PATRONES**

TESIS DOCTORAL

Julio César Caiza Ñacato

Ingeniero en Electrónica y Redes de Información

2020

Departamento de Ingeniería de Sistemas Telemáticos

Escuela Técnica Superior de Ingenieros de
Telecomunicación

Universidad Politécnica de Madrid

TESIS DOCTORAL

CONTRIBUCIÓN AL DISEÑO DE SISTEMAS
RESPETUOSOS CON LA PRIVACIDAD USANDO
PATRONES

Autor:

Julio César Caiza Ñacato

Ingeniero en Electrónica y Redes de Información

Director:

Dr. José María Del Álamo Ramiro

Doctor por la Universidad Politécnica de Madrid

Madrid, Septiembre 2020

Tribunal nombrado por el Magnífico y Excelentísimo Sr. Rector de la Universidad Politécnica de Madrid.

Título: Contribución al diseño de sistemas respetuosos con la privacidad usando patrones

Autor: Julio César Caiza Ñacato

Director: José María Del Álamo Ramiro

TRIBUNAL CALIFICADOR

Presidente: Dña. M. Encarnación Pastor Martín

Catedrática de Universidad
Universidad Politécnica de Madrid

Vocales: D. José M. Such

Reader
King's College London

Dña. Natalia Criado

Lecturer
King's College London

D. Rubén Trapero Burgos

Arquitecto de ciberseguridad
Atos Spain

Secretario: D. Félix Cuadrado Latasa

Investigador Distinguido Senior
Universidad Politécnica de Madrid

Suplentes:

Dña. Beatriz San Miguel González

Investigadora Senior
Fujitsu Laboratories of Europe

D. Boni García Gutiérrez

Profesor Ayudante Doctor
Universidad Carlos III de Madrid

Realizado el acto de defensa y lectura de la Tesis Doctoral en Madrid, el día ... de ... de 2020.

El Presidente

Los Vocales

El Secretario

A quienes confiaron y confían en mí

Agradecimientos

Aunque el iniciar esta etapa haya sido decisión propia, el llegar a culminarla ha sido posible gracias a muchas personas geniales que iniciaron o que han aparecido en el camino. De ellas he ido aprendiendo y en ellas me he ido apoyando. Ellas me han hecho reafirmar que lo más importante es la persona. ¡Gracias por estar ahí!

Al grupo de investigación, gracias por el acompañamiento en este período de formación. Chema, gracias por la confianza, por la dirección, por enseñarme, empujarme y motivarme; sobre todo, gracias por el buen trato y por los momentos amenos. Samuel, gracias por tu valiosa colaboración en los diferentes estudios. Gracias por tus μYods sobre lo académico, lo cultural, los lugares, la comida, etc. Gracias por contribuir con ello a que conozca un poco más sobre España. He aprendido muchísimo de ustedes.

A mis grandes amigos, Juan y Danny, gracias por haber sido mi gran apoyo en este camino. Ustedes han hecho más llevadero y ameno el camino. Gracias por las charlas, por las risas, por las reuniones, por los consejos, y si sigo, necesitaré un tomo completo. ¡Gracias!

A mis amigas que siempre han estado cerca, física o espiritualmente, gracias por apoyar a mantenerme sano física y mentalmente. Gracias por la comida, las charlas, por los consejos, por el apoyo. Gracias Nao, Nata, Lila, Nathy y Ale.

Gracias a mis nuevos amigos y a los de siempre, son geniales. Ustedes reafirman que siempre podré encontrar buenos amigos en cualquier lugar. Ustedes hacen que haya muchos lugares, en Ecuador o fuera de él, a los que quiera ir o, de seguro, regresar. No me arriesgo a poner una lista, pero estoy seguro que se reconocerán.

A mis padres y hermanos, gracias porque siempre han estado, porque siempre están, porque siempre me ven con ojos de amor y cariño. Gracias por su alegría Pame, Guiss y Denni. Gracias por devolverme la energía cuando la necesité.

Gracias a todos los que estuvieron y me alentaron. Gracias por haberme empujado cuando dudé en subirme a un avión. Gracias por el abrazo sincero, por el cariño, por las sonrisas. Gracias por las palabras de motivación y aliento, por darme la fortaleza. Gracias por los viajes, por la emoción, por los momentos. Gracias por las reuniones, por los juegos, por las comidas, por las cervezas, por las risas. Gracias por haber compartido felicidad.

No podía olvidarme, gracias a Madrid, la ciudad que me acogió, donde conocí a mucha de la gente genial, y que me encanta. Seguro regresaré.

Gracias a la Escuela Politécnica Nacional y a la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación del Ecuador por el apoyo y auspicio. Voy de regreso a contribuir en mi Ecuador.

¡Gracias a la vida que ha puesto en mi camino a gente tan valiosa!

Resumen

Actualmente, los datos y su procesamiento son inherentes a la sociedad digital en la que vivimos. Los datos han llegado a estar tan vinculados a las personas que, si se quiere proteger su derecho humano a la privacidad, se debe también asegurar la protección de sus datos personales. El escenario es complejo, las organizaciones usan datos para sostener sus modelos de negocio y las personas son cada vez más conscientes y están preocupadas por su privacidad. A esto se han sumado los escándalos de uso masivo e ilícito de dichos datos.

Diferentes mecanismos han sido definidos para tratar de establecer orden ante tal escenario. Por ejemplo, el concepto de la Privacidad desde el Diseño tiene por objetivo crear sistemas de información que desde su concepción velen por la protección de las personas y su derecho a la privacidad. Este concepto ha sido adoptado por legislaciones como el Reglamento General de Protección de Datos Europeo del 2016 que, desde un enfoque más pragmático, obliga a la Protección de Datos desde el Diseño.

Herramientas, técnicas, métodos y teorías propuestas por el área emergente de la Ingeniería de Privacidad buscan cumplir con ese mandato. Aunque el número de estas propuestas ha venido en continuo incremento, aquellas orientadas a la actividad de diseño en el proceso de desarrollo son bastante dispersas y aún necesitan probar su utilidad en la práctica. Sin embargo, dada la necesidad de contar con estos elementos para brindar soporte a los desarrolladores, varios de ellos (p.ej. métodos) ya forman parte de reportes y recomendaciones técnicas realizadas por la Agencia Española de Protección de Datos y la Agencia de la Unión Europea para la Ciberseguridad.

Esta tesis doctoral busca contribuir al conocimiento sobre el diseño de sistemas respetuosos con la privacidad haciendo avances para su implementación en la práctica. Particularmente centra la atención en los patrones de diseño ya que de manera natural se acoplan a la situación actual de la ingeniería de privacidad. Esto es, el proceso de diseño debe tener en cuenta consideraciones de distintos ámbitos (p.ej. legal, técnico), el nivel de privacidad a alcanzar depende fuertemente del contexto de aplicación del sistema diseñado (p.ej. salud, localización), los desarrolladores no son expertos en privacidad, y los patrones de diseño son elementos bastante conocidos en el proceso de desarrollo de software habitual.

Para cumplir con el objetivo planteado, esta tesis se ha dividido en tres partes. La primera parte, tras realizar un estudio sistemático exhaustivo del estado del arte, presenta un mapa de elementos reusables para el diseño de sistemas de información respetuosos con la privacidad. El mapa muestra claramente que hay un grupo de elementos que captan la mayor atención por parte de los investigadores: los métodos y estrategias, los modelos de referencia, los patrones de privacidad y las herramientas. De estos, los patrones de privacidad se constituyen en un área de investigación de corte transversal al resto y tiene la mayor cantidad de contribuciones.

El análisis del área de investigación en patrones de privacidad muestra que la mayoría de las contribuciones se han centrado en el desarrollo de nuevos patrones, y de métodos y enfoques basados en ellos. Sin embargo, cómo aplicar patrones de privacidad de una manera exitosa ha recibido poca atención. Este tema debe ser abordado de cara a alcanzar la implementación de patrones en la práctica y para probar los beneficios asumidos por las otras propuestas basadas en patrones de privacidad.

Es así que en la segunda parte se analiza en detalle el proceso de aplicación de los patrones de privacidad. Se realiza un estudio experimental exploratorio para determinar la percepción que los desarrolladores tienen sobre los patrones de privacidad usados y para comprobar los beneficios usualmente atribuidos a su aplicación. Para identificar los puntos de mejora, se lleva a cabo un estudio en el que se analiza el área más madura de los patrones de seguridad y se elabora un marco de referencia que organiza los elementos que afectan de manera positiva o negativa la aplicación exitosa de los patrones. Asimismo, se identifican los elementos que más impacto positivo causarían de cara a conseguir dicha aplicación exitosa.

Finalmente, en la tercera parte se desarrollan los elementos que más impacto causan y que son más prioritarios: una taxonomía de tipos de relaciones entre patrones de privacidad y dos sistemas de patrones de privacidad. La taxonomía permite normalizar las relaciones entre los patrones de privacidad y facilita al investigador el establecimiento de las mismas. Además, permitiría que el desarrollador navegue entre los patrones para encontrar aquellos más apropiados para resolver un problema. Respecto a los sistemas de patrones, el primero se enfoca en informar al usuario sobre el procesamiento de sus datos personales, mientras el segundo le permite el control sobre dicho procesamiento. En el estudio se elabora sobre un grupo de patrones, disponibles en el repositorio de patrones de privacidad *privacypatterns.org*, para permitirles llegar al nivel de sistema de patrones.

Las contribuciones realizadas a lo largo de la tesis brindan una panorámica completa de los elementos para el diseño respetuoso con la privacidad, abordan por primera vez las cuestiones de la aplicación de los patrones de privacidad, crean un marco referencial que señala el camino de varias futuras contribuciones, y se elaboran dos de los elementos que más impacto causarían. De este modo, se alcanza el objetivo planteado.

Las contribuciones se han logrado en el contexto de varias colaboraciones con investigadores nacionales e internacionales, algunas en el marco de estancias de investigación en instituciones internacionales de prestigio. Además, las contribuciones se han validado a través de diversas publicaciones en revistas y congresos internacionales de relevancia.

Abstract

Today, data and its processing are inherent in the digital society we live in. Data have become so linked to people that to guarantee their human right to privacy, their personal data must be protected. The scenario is complex. Organizations use data to support their business models, and people are increasingly aware of and concerned about their privacy. Additionally, there have been scandals involving the massive and illicit use of data.

Different mechanisms have been defined to try to establish order in such a scenario. For example, the concept of Privacy by Design aims to create information systems ensuring the protection of people's right to privacy from their conception. This concept has been embraced by legislation such as the European General Data Protection Regulation of 2016, which, in a more pragmatic approach, mandates Data Protection by Design.

Tools, techniques, methods and theories proposed by the emerging field of Privacy Engineering would make it possible to fulfill this mandate. The number of these proposals has been continuously increasing, but those oriented to the design activity in systems development are scattered and still need to prove their usefulness in practice. However, given the necessity of these elements to support developers, several of them (e.g., methods) are already part of reports and technical recommendations made by the Spanish Data Protection Authority and the European Union Agency for Cybersecurity.

This doctoral thesis aims to contribute to the knowledge about the design of privacy-aware systems by making advances towards its implementation in practice. It specifically focuses on design patterns as they naturally fit into the current situation of privacy engineering. That is, the design process must take into account considerations coming from different fields (e.g., legal, technical), the level of privacy to be achieved depends on the application context of a system being designed (e.g., health, location), developers are not privacy experts and design patterns are well-known elements in the software development process.

In order to achieve this objective, this thesis has been divided into three parts. The first part, after an exhaustive systematic study of the state of the art, presents a map of reusable elements for the design of privacy-aware information systems. The map clearly shows a group of the following elements capturing the most attention from researchers: methods and strategies, reference models, privacy patterns and tools. Privacy patterns in particular are a cross-cutting research area and have the most contributions of any of these elements.

The analysis of the research in privacy patterns shows that most of these contributions have focused on the development of new patterns and on methods and approaches based on these new patterns. However, the question of how to apply privacy patterns successfully has received little attention. There is a great need to study this topic to implement patterns in practice and to prove the benefits assumed by other privacy-pattern-based approaches.

Thus, the second part of this thesis discusses in detail the process of applying privacy patterns. An exploratory experimental study is performed to determine developers' perception about privacy patterns and to prove the benefits usually attributed to their application. To identify points of improvement, a study is carried out to analyze the most mature research area of the security patterns and to elaborate a framework that organizes the elements that positively or negatively affect the successful application of

patterns. In addition, those elements that would have the greatest positive impact on achieving a successful pattern application are identified.

Finally, in the third part of this thesis, the most impactful and priority elements are developed: a taxonomy of types of relationships for privacy patterns and two systems of privacy patterns. The taxonomy makes it easy for researchers to establish normalized relationships between privacy patterns. Furthermore, it would allow a developer to navigate through groups of privacy patterns to find the most appropriate pattern to solve a problem. Regarding the pattern systems, the first one focuses on informing users about the processing of their personal data, while the other allows users to control such processing. The study uses a group of privacy patterns from the repository available at *privacypatterns.org* and updates them to allow them to reach the pattern system level.

The contributions made throughout this thesis provide a comprehensive overview of the elements of privacy-aware design, address for the first time the issues regarding the application of privacy patterns, create a framework that points the way for many future contributions and elaborate on two of the most impactful elements. In this way, the stated objective is achieved.

These contributions have been elaborated in the context of several collaborations with national and international researchers. Some collaborations have been materialized during research stays at prestigious institutions. The contributions have been validated through publications in relevant journals and conferences.

Contenido

Capítulo 1: Introducción	1
1.1. Contexto	1
1.2. Problema.....	2
1.3. Objetivo general y preguntas de investigación.....	2
1.4. Ámbito.....	4
1.5. Contribuciones originales y metodología seguida.....	5
1.6. Estructura de la memoria.....	7
Capítulo 2: Antecedentes.....	9
2.1. Privacidad.....	10
2.1.1. Datos personales	12
2.1.2. Protección de la privacidad y de los datos personales.....	13
2.2. Legislación	14
2.3. Privacidad desde el diseño	16
2.4. Privacidad en la ingeniería de sistemas	17
2.4.1. Análisis y diseño de sistemas	17
2.4.2. Ingeniería de privacidad	18
Análisis	19
Diseño	20
2.5. Patrones de privacidad en el diseño de sistemas	21
2.5.1. Patrones	21
Evolución de los patrones.....	22
Ciclo de vida de los patrones.....	23
Conferencias sobre Patrones.....	24
2.5.2. Ingeniería con patrones de privacidad	24
Colecciones existentes	24
Colaboración hacia un repositorio común	26
Métodos usando patrones de privacidad.....	26
Estudios empíricos con patrones de privacidad.....	27
2.6. Resumen	28
Capítulo 3: Mapa de elementos de diseño en la ingeniería de privacidad.....	29
3.1. Definición de los elementos reusables de diseño	30
3.2. Metodología.....	32

3.2.1.	Planificación	33
	Definición del ámbito del estudio.....	33
	Identificación del estudio	34
	Clasificación y extracción de datos	37
3.2.2.	Ejecución	37
3.3.	Mapa de elementos reusables de diseño.....	38
3.3.1.	Madurez	40
3.3.2.	Tendencias.....	41
3.3.3.	Base de conocimiento en privacidad	43
3.3.4.	Métodos	45
3.4.	Patrones de privacidad.....	46
3.4.1.	Madurez	49
3.4.2.	Tendencias	50
3.5.	Resumen del capítulo y contribuciones originales.....	51
Capítulo 4: Estudio exploratorio de la aplicación de los patrones de privacidad...		53
4.1.	Metodología.....	54
4.1.1.	Ámbito	54
4.1.2.	Planificación	54
	Contexto	54
	Preguntas de investigación e hipótesis	55
	Variables.....	55
	Participantes	56
	Tipo de diseño	56
	Instrumentación	56
	Amenazas a la validez	60
4.1.3.	Operación	60
	Preparación	60
	Ejecución	61
4.2.	Resultados	62
4.2.1.	Pregunta de investigación PIX1	62
4.2.2.	Pregunta de investigación PIX2	63
4.3.	Posibilidades de experimentación con patrones de privacidad	63
4.3.1.	Estado actual de los instrumentos para experimentación	64
	Instrumentos faltantes en el estado del arte	64
	Instrumentos existentes en el estado del arte.....	65

4.3.2.	Enfoques de calidad.....	67
4.3.3.	Estudiantes como participantes en experimentos en el área.....	69
4.4.	Consideraciones sobre la aplicación de patrones de privacidad.....	70
4.5.	Resumen del capítulo y contribuciones originales.....	71
Capítulo 5: Marco de referencia para la aplicación de patrones de privacidad.....		73
5.1.	Metodología.....	74
5.1.1.	Exploración del dominio de los patrones de seguridad.....	74
5.1.2.	Elaboración del marco de referencia para la aplicación de patrones..	75
5.1.3.	Exploración del dominio de los patrones de privacidad.....	75
5.2.	Marco de referencia para la aplicación de patrones.....	76
5.2.1.	Elementos del marco.....	76
5.2.2.	Aspectos de aplicación.....	77
	Aspectos en la etapa de redacción.....	77
	Aspectos en la etapa de organización.....	78
	Aspectos en la etapa de revisión.....	79
	Aspectos en la etapa de publicación.....	79
	Aspectos en la etapa de selección.....	79
	Aspectos en la etapa de instanciación.....	79
	Aspectos en la etapa de evaluación de la aplicación.....	80
5.3.	Estado actual de los aspectos de aplicación en los patrones de privacidad....	80
5.4.	Guía para mejorar la instanciación de los patrones de privacidad.....	82
5.4.1.	Priorización de las mejoras.....	83
5.4.2.	Próximos pasos.....	85
5.5.	Resumen del capítulo y contribuciones originales.....	85
Capítulo 6: Relaciones entre patrones de privacidad.....		87
6.1.	Metodología.....	88
6.2.	Taxonomía de relaciones.....	88
6.3.	Análisis de la taxonomía de tipos de relaciones.....	92
6.3.1.	Validación de la taxonomía propuesta en los catálogos individuales	93
6.3.2.	Limitaciones de los esquemas de relación actuales.....	94
6.3.3.	Utilidad de la taxonomía.....	94
6.3.4.	Extensiones futuras.....	95
6.4.	Resumen del capítulo y contribuciones originales.....	96
Capítulo 7: Sistemas de patrones de privacidad.....		99
7.1.	Metodología.....	100

7.2.	Conformación de los requisitos para ser sistemas.....	100
7.2.1.	Requisitos existentes (alcanzan para ser un catálogo).....	101
	Base suficiente (R1)	101
	Descripción uniforme (R2).....	101
	Esquema de organización (R4).....	101
	Evolución (R6)	102
7.2.2.	Requisitos cumplimentados para alcanzar el nivel de sistema	102
	Descripción uniforme (R2).....	103
	Guías de implementación (R3).....	103
	Interconexiones (R5)	104
7.3.	Sistema para informar sobre el procesamiento de datos	105
7.4.	Sistema para controlar el procesamiento de datos.....	106
7.5.	Implicaciones en el proceso de aplicación	107
7.5.1.	Mejoras al proceso de aplicación.....	108
7.5.2.	Limitaciones existentes y vías de mejora	109
7.5.3.	Integración en el proceso de diseño con requisitos de privacidad....	110
7.6.	Resumen del capítulo y contribuciones originales	110
Capítulo 8: Divulgación de resultados		113
8.1.	Diseminación de resultados directos de la tesis	114
8.1.1.	Publicaciones en revistas	114
8.1.2.	Publicaciones en congresos	114
8.2.	Publicaciones en el contexto de la tesis.....	115
8.2.1.	Publicaciones en revistas	115
8.3.	Colaboración en proyectos de investigación	116
8.4.	Estancias de investigación.....	117
8.5.	Trabajos fin de titulación.....	117
8.5.1.	Trabajos fin de grado	117
8.5.2.	Trabajos fin de máster	118
8.6.	Conferencias y presentaciones realizadas	118
8.7.	Cursos y seminarios dictados	118
Capítulo 9: Conclusiones y trabajo futuro		119
9.1.	Conclusiones	119
9.1.1.	Versión en español	119
9.1.2.	Versión en inglés	121
9.2.	Trabajos futuros.....	122

Bibliografía	125
Anexos	140

Índice de Figuras

Figura 2.1: Tipos de privacidad acorde a Clarke	11
Figura 2.2: Protección de la privacidad y de los datos personales.....	14
Figura 2.3: Enfoques de elicitación de requisitos de privacidad	19
Figura 2.4 Ejemplo de patrón de privacidad.....	22
Figura 2.5 Ejemplo de implementación de un patrón en una red social	22
Figura 2.6: Etapas del ciclo de vida de un patrón	23
Figura 3.1: Elementos reusables de diseño	30
Figura 3.2: Etapas de un mapeo sistemático.....	33
Figura 3.3: Tareas en la etapa de planificación.....	33
Figura 3.4: Ejecución del mapeo sistemático	38
Figura 3.5: Mapa de elementos reusables de diseño.....	39
Figura 3.6: Madurez de los elementos reusables de diseño	42
Figura 3.7: Cantidad de artículos por año.....	42
Figura 3.8: Tendencias en madurez de elementos reusables de diseño	43
Figura 3.9: Madurez del área de investigación de los patrones de privacidad	49
Figura 3.10: Madurez de categorías de investigación en patrones de privacidad.....	50
Figura 4.1: Proceso de experimentación en ingeniería de software.....	54
Figura 4.2: Proceso de resolución de problema con requisitos de privacidad	57
Figura 4.3: Ejemplo de problema de diseño con requisitos de privacidad	58
Figura 4.4: Ejemplo de patrón de privacidad: Identidad seudónima	58
Figura 4.5: Ejemplo de interfaces del experimentador (superior) y los participantes (inferior)..	59
Figura 4.6: Etapas dentro del estudio.....	62
Figura 4.7: Percepción de los participantes sobre los patrones de privacidad.....	63
Figura 4.8: Relaciones y dependencias entre los instrumentos.....	65
Figura 5.1: Etapas en la construcción del marco de referencia para la aplicación de patrones...	74
Figura 5.2: Marco de referencia para la aplicación de patrones	77
Figura 5.3: Estado de las mejoras a la aplicación de los patrones de privacidad	83
Figura 6.1: Etapas en el desarrollo de la taxonomía de tipos de relaciones.....	88
Figura 6.2: Ejemplo de aplicación de la taxonomía de tipos de relaciones	95
Figura 7.1: Etapas en el desarrollo de los sistemas.....	100
Figura 7.2: Sistema de patrones de privacidad para la estrategia Informar	106
Figura 7.3: Sistema de patrones de privacidad para la estrategia de control	107
Figura 7.4: Repositorio común de patrones de privacidad	109
Figura 8.1: Mapa de contribuciones.....	113

Índice de tablas

Tabla 3.1: Definiciones de los elementos reusables de diseño	31
Tabla 3.2: Cadena de búsqueda	34
Tabla 3.3: Criterios de inclusión y exclusión automatizados	35
Tabla 3.4: Criterios de inclusión y exclusión manuales	36
Tabla 3.5: Tipos de investigación usados como niveles de madurez	37
Tabla 3.6: Elementos reusables dentro de los métodos	41
Tabla 3.7: Investigación en modelos de referencia.....	44
Tabla 3.8: Investigación en métodos de diseño	46
Tabla 3.9: Investigación en patrones de privacidad.....	48
Tabla 4.1: Variables utilizadas en el experimento	56
Tabla 4.2: Conocimiento de los participantes antes del curso	61
Tabla 4.3: Número de participantes asignados a los problemas de diseño	61
Tabla 4.4: Frecuencias de evaluaciones a los problemas de diseño	62
Tabla 4.5: Valores para la prueba U de Mann-Whitney para cada problema.....	63
Tabla 4.6: Características y beneficios atribuidos a los patrones de diseño	67
Tabla 4.7: Enfoques de calidad y métricas para experimentos con patrones de privacidad	69
Tabla 5.1: Cadenas de búsqueda para el dominio de la seguridad.....	74
Tabla 5.2: Cadenas de búsqueda para el dominio de la privacidad	76
Tabla 6.1: Tipos de relaciones para los patrones de privacidad	89
Tabla 6.2: Descripción breve de los patrones de privacidad usados en los ejemplos.....	90
Tabla 6.3: Mapeo de los tipos de relaciones propuestos a los catálogos existentes	93
Tabla 7.1 Plantilla común usada para los patrones dentro de los sistemas.....	103
Tabla A.1: Datos bibliográficos de artículos incluidos en el estudio sistemático	140
Tabla A.2: Artículos sobre la aplicación de patrones de seguridad.....	143
Tabla A.3: Patrones de privacidad de la táctica <i>explicar</i>	144
Tabla A.4: Patrones de privacidad de la táctica <i>notificar</i>	144
Tabla A.5: Patrones de privacidad de la táctica <i>proveer</i>	145
Tabla A.6: Patrones de privacidad de la táctica <i>actualizar</i>	146
Tabla A.7: Patrones de privacidad de la táctica <i>consentir</i>	146
Tabla A.8: Patrones de privacidad de la táctica <i>retirar</i>	146
Tabla A.9: Patrones de privacidad de la táctica consentimiento <i>elegir</i>	147

Acrónimos

ACM	Association for Computing Machinery
AEPD	Agencia Española de Protección de Datos
API	Application Programming Interface
BCP	Base de Conocimiento de Privacidad
CORE	Ranking Computing Research and Education Association
CSA	Cloud Security Alliance
DFD	Data Flow Diagram
DPD	Data Protection Directive
ENISA	European Union Agency for Cybersecurity
EuroPLOP	European Conference on Pattern Languages of Programs
FIPPS	Fair Information Practice Principles
GDPR	General Data Protection Regulation
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
JCR	Journal Citation Reports
JSON	JavaScript Object Notation
NCSRA	National Cyber Security Research Agenda
NSF	National Science Foundation
OECD	Organisation for Economic Cooperation and Development
P4P	Project Patterns for Privacy
PbD	Privacy by Design
PDP4E	Project Methods and tools for GDPR compliance through Privacy and Data Protection Engineering
PETS	Privacy-Enhancing Technologies
PI.lab	Privacy & Identity Lab
PLOP	Pattern Languages of Programs
RAE	Real Academia Española
RGPD	Reglamento General de Protección de Datos
SEMDM	Software Engineering Metamodel for Development Methodologies
SJR	Scimago Journal Rank
SJR	Factor de impacto SCImago Journal Rank
TRUESSEC.eu	Project TRUst-Enhancing certified Solutions for SEcurity and protection of Citizens' rights in digital Europe
UML	Unified Modeling Language
URL	Uniform Resource Locator
XML	Extensible Markup Language
XP	Extreme Programming

Capítulo 1

Introducción

1.1. Contexto

“Los datos personales son el nuevo petróleo de Internet
y la nueva moneda del mundo digital”

Meglana Kuneva

Comisaria Europea de Protección de los Consumidores, 2009

Los datos han llegado a convertirse en la materia prima principal en el mundo digital actual. Se estima que en 2025 el valor de la economía de datos europea será de 829 mil millones de euros, representando un 5,8% de su producto interno bruto [EuroC20]. Esto se debe, según el Director General de Redes de Comunicación, Contenido y Tecnología de la Comisión Europea [Viola18], a la confianza generada que ha permitido aprovechar todo el poder de los datos. Parte fundamental de esta confianza se atribuye al marco sólido de protección de datos personales establecido en el Reglamento General de Protección de Datos (RGPD) [EuroPC16], de obligado cumplimiento desde el 25 de mayo de 2018.

El RGPD trata de poner orden en un escenario complejo en el cual tanto las organizaciones como las personas juegan un rol fundamental. Por un lado, las organizaciones están muy interesadas en explotar los datos personales, a través de diversos modelos de negocio, para ofrecer nuevos productos y servicios. Por el otro, las personas son cada vez más conscientes de sus datos personales, y están más preocupadas de las implicaciones en su privacidad [TNSPS16]. El escándalo de Facebook y Cambridge Analytica [Cadwalladr18] ha servido como un ejemplo real de un uso ilícito y masivo de datos, y de la necesidad de tener una legislación que establezca orden en un escenario así.

Uno de los mecanismos mandados en el RGPD para salvaguardar los datos personales es la protección de datos desde el diseño. En línea con el concepto de Privacidad desde el Diseño [Cavoukian09], busca proteger los datos personales de las personas, y por lo tanto su privacidad, desde las etapas tempranas del proceso de ingeniería de los sistemas. Precisamente, la disciplina emergente de la Ingeniería de Privacidad se ha ocupado de proponer soluciones que ayuden a los ingenieros a incluir cuestiones de privacidad a lo largo de todas las actividades del proceso de desarrollo de los sistemas de información.

Entre estas soluciones se incluyen herramientas, técnicas, métodos y teorías para su uso en las diferentes áreas de la ingeniería de privacidad: análisis de riesgos y amenazas [De16], análisis de requisitos [Kalloniatis09][Anthonysamy17], diseño respetuoso con la privacidad [Hoepman14] [Lenhard17], evaluación de la privacidad [Guaman20], tecnologías de soporte a la privacidad (Privacy-enhancing technologies – PETS) [Shen11][Heurix15], entre otros.

A pesar de la disponibilidad de estas propuestas, estudios de campo recientes [Ayalon17][Bednar18][Hadar18][Senarath18][Senarath19] muestran que los desarrolladores aún no las adoptan. Incluso cuando usarlas podría ayudarles a enfrentar

varios obstáculos en el desarrollo de sus sistemas. Estos obstáculos incluyen la falta de conocimiento y entendimiento de la privacidad, además de la dificultad de pasar de un alto nivel de abstracción de los principios y consideraciones de privacidad hacia las decisiones técnicas de diseño e implementación. En este escenario, los desarrolladores usan su propia opinión, creencias y experiencia respecto a la privacidad para realizar implementaciones en un sistema. Los resultados muestran que las propuestas serían mejor recibidas si las mismas son percibidas como útiles y se acoplan al esquema de trabajo al cual están acostumbrados los desarrolladores.

En un contexto en el cual hay un número creciente de contribuciones en la ingeniería de privacidad, con una adopción aún no conseguida, se tiene además un desequilibrio en la atención que se ha dado a las distintas áreas. Esto es, hay áreas más trabajadas y consolidadas que otras. Áreas como el análisis de riesgo y amenazas han llegado a definir guías bastante claras en estándares internacionales [ISO/IEC29134] [ISO/IEC29151]; relacionado a esto, el área de análisis de requisitos ha llegado a consolidar dos enfoques, uno basado en riesgos y uno orientado a metas [Notario15]; asimismo, las PETs se han venido investigando antes del año 2000 [Goldberg97] y actualmente existen soluciones ampliamente conocidas y utilizadas (P. ej. Tor para enrutamiento anónimo¹). Sin embargo, áreas como el diseño de sistemas aún son incipientes y muestran un problema adicional: una dispersión desordenada. Varios tipos de propuestas han aparecido para brindar soporte al diseño: estrategias [Hoepman14], tácticas [Colesky16], patrones [Schümmer04][Fischer-Hübner10][Lenhard17]), métodos [Iachello05][Kalloniatis8][Moncrieff09][Fhom11][Aljohani16][Notario15], modelos [Karjoth02][Gope16] y herramientas [Drozd16a].

1.2. Problema

La investigación en el área del diseño respetuoso con la privacidad está en un contexto en el cual hay un número creciente y disperso de contribuciones, cuya adopción aún no se ha logrado. Para avanzar en este campo, es urgente encarar los siguientes problemas:

- La gran dispersión de propuestas en un entorno donde se tiene una creciente cantidad de contribuciones. Además, es necesario conocer su madurez de cara a saber cuán cerca están de poder ser usadas en la práctica, alejándose de ser meramente soluciones filosóficas.
- La falta de soporte adecuado a aquellos desarrolladores que no tienen una alta formación previa en privacidad. Asimismo, de cara a aumentar las posibilidades de adopción, deben acoplarse a los modelos a los cuales los desarrolladores están acostumbrados.
- La falta de evidencia de que las propuestas pueden ser usadas en la práctica y que son útiles. Esto es, deberían demostrar que proveen beneficios en el proceso de desarrollo de los sistemas de cara a fomentar su adopción por los desarrolladores.

1.3. Objetivo general y preguntas de investigación

El objetivo general de esta tesis doctoral es contribuir al conocimiento sobre el diseño de sistemas respetuosos con la privacidad haciendo avances para su implementación en la práctica.

---

¹ Tor Project | Anonymity Online <https://www.torproject.org/>

El enfoque que se adopta en esta tesis tiene tres partes:

- Un análisis de la situación actual de las diferentes propuestas para el diseño de sistemas respetuosos con la privacidad. A través de ello se establece una organización de las mismas y se identifican las áreas más prometedoras de cara a realizar contribuciones que causen impacto.
- Un estudio más específico sobre el tipo de propuesta más reportado para la actividad de diseño, los patrones de privacidad, y sobre su proceso de aplicación. Por su naturaleza de proveer conocimiento experto reusable para resolver problemas recurrentes, pueden llegar a ser aplicados por desarrolladores novatos en el dominio de la privacidad.
- Finalmente, el desarrollo de mejoras para los patrones de privacidad de cara a favorecer un proceso de aplicación exitoso.

Para cumplir con el objetivo general, las siguientes preguntas de investigación marcan la pauta del desarrollo de esta tesis.

- *P1. ¿Cuáles son las propuestas existentes para el diseño de sistemas respetuosos con la privacidad?*
El estudio de las propuestas existentes permite proveer un mapa del estado de la investigación en el área, ayudando a determinar el contexto en el cual se desarrolla la presente tesis.
- *P1.1. ¿Cuáles con las tendencias y madurez de la investigación en los patrones de privacidad?*
Teniendo el contexto de las propuestas existentes, es necesario hacer un acercamiento al área de los patrones de privacidad, siendo que estos han sido mayoritariamente reportados. Esto permite identificar los temas de investigación que se han abordado y se abordan actualmente, para impulsar aquellos que más impacto pueden causar.
- *P2. ¿Los patrones de privacidad actuales reúnen las características suficientes para su aplicación?*
Aunque la investigación en patrones de privacidad esté en crecimiento, interesa saber si los patrones de privacidad, en su estado actual, pueden ser aplicados por los desarrolladores para diseñar sistemas respetuosos con la privacidad. Más aún, se desea conocer si su uso provee beneficios en el proceso de diseño.
- *P3. ¿Cuáles son las cuestiones que afectan la aplicación de los patrones de privacidad?*
La aplicación de un patrón de privacidad no se reduce a una tarea aislada de instanciación; esta tarea puede tener afectaciones desde la definición del patrón. Por ende, es menester identificar las afectaciones y las medidas que se pueden tomar para maximizar las oportunidades de una aplicación exitosa.
- *P4. ¿Qué medidas ya se han implementado para mejorar la aplicación de los patrones de privacidad?*
Conociendo las medidas que se podrían tomar para favorecer la aplicación de los patrones de privacidad, es necesario saber cuáles ya han sido implementadas para evitar ‘reinventar la rueda’.
- *P5. ¿Qué medidas deberían implementarse primero para causar más impacto?*

El conocer las medidas que actualmente existen desvela aquellas que necesitan trabajarse. Debido a las limitaciones de recurso humano y de tiempo, es necesario establecer un orden y una prioridad.

- *P6. ¿Se pueden implementar las medidas prioritarias?*

La respuesta a esta pregunta es directamente la implementación de las medidas previamente seleccionadas. Su importancia es implícita, debido a que necesitamos implementar la medida para poder luego usarla o validarla.

- *P6.1 ¿Cómo se pueden validar estas medidas?*

La implementación de las medidas no es suficiente, se necesita llegar a validarla para demostrar su utilidad.

1.4. **Ámbito**

El objetivo de esta tesis aporta a la ingeniería de privacidad proveyendo contribuciones que permiten llevar a cabo lo establecido en legislaciones de protección de datos como el RGPD, para finalmente proteger a las personas y garantizar medidas que permiten a las organizaciones mantener sus modelos de negocio. Esto se logra haciendo contribuciones a la base de conocimiento en la ingeniería de privacidad, específicamente al diseño de sistemas respetuosos con la privacidad. La importancia de avanzar en la investigación en ingeniería de privacidad ha quedado plasmada, entre otros, en los objetivos estratégicos españoles y europeos, y en los reportes y recomendaciones por agencias internacionales en el ámbito de la ciberseguridad.

Respecto a los objetivos estratégicos a nivel europeo, la Agenda Digital Europea del 2010 [EuroC10] incluyó como campo de actuación a la confianza y seguridad (en su Sección 2.3), haciendo referencia explícita a la “privacidad a través del diseño” como un mecanismo de protección a la intimidad y a los datos personales. Asimismo, la Estrategia para el Mercado Único Digital de Europa del 2015 [EuroC15], para proveer un entorno digital seguro para el mercado único en línea, considera como uno de sus pilares el refuerzo de la confianza y seguridad para el tratamiento de los datos personales (en su Sección 3.4).

En línea con estas estrategias, la Agenda Digital para España del 2013 [GobEsp13] definió como uno de sus objetivos (objetivo 4) el reforzar la confianza digital, y reconoció la protección de la privacidad como un mecanismo de mucho impacto para llevarlo a cabo. Asimismo, el Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020 [GobEsp17], dentro del Reto 7, fija como prioridades la investigación y el desarrollo tecnológico e innovación en ciberseguridad e identidad digital, estando esto directamente relacionado con los datos personales.

Aunque se han venido realizando contribuciones importantes a la ingeniería de privacidad, aún se requiere más trabajo. La Agencia de la Unión Europea para la Ciberseguridad (European Union Agency for Cybersecurity –ENISA) elaboró un reporte [Danezis14] sobre los avances realizados (p. ej. métodos y tecnologías) para materializar la privacidad desde el diseño. Aunque reconoce el creciente interés en el área, llama por esfuerzos para lograr llevar las contribuciones desde la investigación a la práctica. Asimismo, menciona que una inclusión de aspectos de privacidad y protección de datos en los estándares correspondientes podría impulsar la privacidad desde el diseño. Más adelante, el Supervisor Europeo de Protección de Datos en su informe 2018 sobre la privacidad desde el diseño [EuroDPS18] reconoció las diferentes iniciativas e investigación en la disciplina de la ingeniería de privacidad, e indica que aún se requieren más esfuerzos, por lo que recomienda brindar apoyo a las iniciativas de investigación en este dominio. En este escenario, el Programa H2020 de la Unión

Europea considera la privacidad y protección de datos, dentro del área de Sociedades Seguras [EuroC18], como sus prioridades estratégicas para el periodo 2018-2020.

Dada la urgencia por hallar soluciones en la ingeniería de la privacidad, incluso los incipientes esfuerzos han sido agrupados [Danezis14], y recomendados por agencias de protección de datos; por ejemplo, la Agencia Española de Protección de Datos publicó en 2019 una Guía de Privacidad desde el Diseño [AEPD19]. Asimismo, aquellos esfuerzos más estables han sido incluidos en el reciente estándar ISO/IEC sobre Ingeniería de Privacidad para los Procesos del Ciclo de Vida de los Sistemas [ISO/IEC27550]. Sin embargo, lo que se ha logrado no es suficiente; además de crear nuevas propuestas, aún se requieren esfuerzos para avanzar las existentes hacia su implementación en la práctica.

1.5. Contribuciones originales y metodología seguida

Con el afán de contribuir a la investigación en el diseño de los sistemas respetuosos con la privacidad, se ha definido un marco de trabajo delimitado por el objetivo general y las preguntas de investigación. Las diferentes contribuciones de esta tesis ayudan a alcanzar dicho objetivo y a responder las preguntas de investigación.

Debido a que las contribuciones son de distinto tipo (o naturaleza), las metodologías usadas en los estudios llevados a cabo son diferentes. A continuación, se introducen las contribuciones originales principales de esta tesis junto con las metodologías empleadas en los estudios para su elaboración.

- **Un mapa de las propuestas existentes para el diseño de sistemas de información respetuosos con la privacidad.** Este mapa muestra una vista panorámica de las propuestas existentes que han sido reportadas en la literatura. Además, las organiza y muestra su madurez, lo que permite identificar aquellas que mayor atención captan y las tendencias dentro de estas. Además, se hace un acercamiento en los patrones de privacidad al ser las propuestas más reportadas, mostrando así el estado actual en los diferentes temas investigados dentro de esta área. Este mapa ayuda a responder las preguntas *PI* y *PI.1*.

Para construir el mapa se adopta un enfoque basado en evidencias [Kitchenham04], y se lleva a cabo un estudio del tipo *mapeo sistemático de literatura*. Para ello se utiliza las guías, ampliamente aceptadas por los investigadores, propuestas en [Petersen08] y que han sido actualizadas en [Petersen15].

El estudio muestra que la investigación en el área de los patrones de privacidad reporta el desarrollo de nuevos patrones, colecciones de patrones, e incluso, procesos o métodos usándolos. Estas propuestas usualmente asumen que los patrones están listos para aplicarse y que dicha aplicación genera beneficios; sin embargo, esto no ha sido comprobado empíricamente; en su lugar, se han generalizado los beneficios de los patrones en otros dominios más maduros como el de la orientación a objetos [Gamma95].

- **Un análisis de los instrumentos y oportunidades de experimentación en la aplicación de los patrones de privacidad.** Se presenta el estado actual de los instrumentos (incluyendo los patrones de privacidad) y las posibilidades de experimentación basado en los resultados de un estudio experimental exploratorio en el que se aplica un grupo de patrones de privacidad para

determinar la percepción de los desarrolladores y los beneficios de su aplicación. Además, se elaboran instrumentos que podrían ser reutilizados en estudios futuros, que incluyen un conjunto de métricas, una herramienta de soporte y un ejemplo de problema de diseño con requisitos de privacidad. El análisis permite contestar la pregunta *P2*.

En este caso se adopta una postura positivista [Easterbrook08] de cara a llevar a cabo un estudio para comprobar con hechos observables las creencias sobre los patrones de privacidad. En este sentido se lleva a cabo un experimento controlado siguiendo las guías para experimentación en ingeniería de software provistas por Wohlin et al. en el 2012 [Wohlin12].

Durante el estudio, se observa que el proceso de aplicación de los patrones de privacidad no se reduce únicamente a una mera instanciación, sino que, esta podría verse afectada desde etapas tempranas como su definición. Por ejemplo, el uso de un vocabulario normalizado ayudaría a entender un patrón para poder aplicarlo. Se hace necesario entonces localizar aquellos elementos que pueden afectar la instanciación de un patrón; y, por lo tanto, al proceso de aplicación.

- **Un marco de referencia para la aplicación de los patrones de privacidad.**

Este marco define las cuestiones que, a lo largo del ciclo de vida de los patrones, afectan su aplicación. Se presentan dentro del marco los defectos (que afectan negativamente) y las mejoras (que afectan positivamente) la instanciación de los patrones. Posteriormente, se usa el marco para identificar las mejoras existentes, y aquellas faltantes y de gran impacto dentro del área de los patrones de privacidad. Todo esto permite contestar las preguntas *P3*, *P4* y *P5*.

El estudio llevado a cabo adopta una postura crítica [Easterbrook08] y un enfoque basado en evidencia de cara a demostrar que la aplicación de patrones no es una simple tarea de instanciación. Este estudio se realizó siguiendo una metodología mixta: una búsqueda sistemática en bases de datos, siguiendo pasos de las guías propuestas por Petersen et al. [Petersen15] y la teoría fundamentada [Glaser17] para la construcción de nuevo conocimiento.

El marco ha permitido definir las mejoras que pueden ayudar a instanciar exitosamente un patrón. De estas, las que más impacto podrían causar son el establecimiento de relaciones y los sistemas de patrones de privacidad.

- **Una taxonomía de tipos de relaciones para definir conexiones entre patrones de privacidad.** La taxonomía propone un conjunto de tipos de relaciones que pueden ser usados para establecer las conexiones entre patrones de privacidad. Esta es validada a través de un proceso de mapeo y aplicación en catálogos individuales y en un repositorio común de patrones de privacidad localizado en *privacypatterns.org*. Esta contribución responde directamente a las preguntas *P6* y *P6.1*.

Se adopta una postura práctica de cara a implementar las mejoras definidas y priorizadas dentro del marco establecido anteriormente. Para desarrollar la taxonomía se usó un enfoque de triangulación usando estudios relevantes sobre patrones de diseño en dominios más maduros [Meszaros97][Noble98] y un mapeo hacia colecciones de patrones de privacidad.

La taxonomía permite crear relaciones bien definidas entre patrones de privacidad; lo que conjuntamente con otras características de normalización y

categorización, ayudan a los patrones del repositorio común a llegar a un nivel de sistema.

- **Dos sistemas de patrones de privacidad.** Estos sistemas son desarrollados, el uno para informarles a los usuarios sobre el procesamiento de sus datos personales, y el otro para permitirles controlar el proceso. De este modo se evolucionan los patrones de privacidad pertenecientes al repositorio común en *privacypatterns.org* al nivel de sistema. Estas contribuciones también permiten responder las preguntas *P6* y *P6.1*.

Al igual que el caso anterior, se adopta una postura práctica y se implementan los sistemas como mejoras dentro del marco establecido anteriormente. El desarrollo de los sistemas de patrones se llevó a cabo siguiendo los lineamientos y características enunciadas en [Buschmann96].

1.6. Estructura de la memoria

Esta memoria de tesis doctoral se ha estructurado del siguiente modo:

En el Capítulo 2 se revisan los *antecedentes* y se plantea el estado actual de la investigación en el dominio de esta tesis, brindando así el conocimiento previo requerido para entender el resto de capítulos. Este capítulo contiene dos bloques principales, el uno que explica el marco teórico de la privacidad, donde se revisan las características de la privacidad, su importancia, y las legislaciones que abordan los datos personales. El otro bloque está enfocado en el proceso de ingeniería, que aborda el paso del nivel abstracto de la legislación hacia las actividades de la ingeniería de sistemas, enfocándose en el diseño.

En el Capítulo 3 se describe un *mapa de las propuestas existentes para el diseño de sistemas de información respetuosos con la privacidad*. En la primera parte se describe en detalle la metodología seguida para llevar a cabo un mapeo sistemático de literatura. A continuación, se muestran las propuestas encontradas agrupadas por tipo, se indica su madurez y sus tendencias. Posteriormente, se muestra el estado actual y las tendencias en los temas de investigación dentro del área de los patrones de privacidad.

El Capítulo 4 presenta *un análisis de los instrumentos y oportunidades de experimentación en la aplicación de los patrones de privacidad*. Se describe el contexto de los patrones de privacidad a evaluar, el diseño del experimento y la ejecución. A continuación, se muestran los resultados preliminares respecto a la percepción y a los beneficios de los patrones de privacidad. Posteriormente, se presenta un análisis en profundidad del estado de los instrumentos para la experimentación en patrones de privacidad, con lo cual se presenta las posibilidades actuales de experimentación. Además, se elabora un análisis sobre las implicaciones en el proceso de aplicación de los patrones de privacidad.

En el Capítulo 5 se expone *un marco de referencia para la aplicación de los patrones de privacidad*. Se inicia con la descripción de la metodología seguida para desarrollar el marco. A continuación, se describe en detalle el marco y sus elementos constitutivos: los defectos y las mejoras, agrupados en aspectos y ordenados acorde a las etapas por las que pasa un patrón (p. ej. escritura, organización, publicación, etc.). Posteriormente, se detallan las mejoras existentes en cada etapa y aquellas que aún se necesitan en el ámbito de los patrones de privacidad. Además, se ordenan y priorizan las mejoras de cara a determinar aquellas que más impacto tendrían.

El Capítulo 6 describe *una taxonomía de tipos de relaciones para definir conexiones entre patrones de privacidad*. Se presenta la importancia y los beneficios de tener una taxonomía como medio de normalización de las relaciones. A continuación, se describen las experiencias previas en dominios más maduros. Posteriormente, se presenta la taxonomía, describiendo cada uno de los tipos de relaciones definidos y brindando ejemplos de uso. Como medio de validación, se muestran el mapeo hacia catálogos existentes de patrones de privacidad.

En el Capítulo 7 se desarrollan *dos sistemas de patrones de privacidad*. En primer lugar, se presentan los requisitos que deben cumplir los conjuntos de patrones para llegar a ser considerados sistemas de patrones. A continuación, se describen los cambios hechos de cara a cumplimentar los requisitos; y posteriormente, se exponen los dos sistemas desarrollados, uno para informar al usuario sobre el procesamiento de sus datos personales, y otro para permitir el control de ese procesamiento. Finalmente, se presenta un análisis de las implicaciones al proceso de aplicación de los patrones de privacidad.

En el Capítulo 8 se listan las actividades relacionadas con la disseminación de los resultados obtenidos a lo largo de la realización de la tesis, incluyendo: publicaciones en revistas y congresos de relevancia, exposiciones en conferencias, colaboraciones con proyectos de investigación, realizaciones de estancias de investigación, direcciones de trabajos de titulación de máster y de pregrado, y capacitaciones en cursos específicos.

Finalmente, en el Capítulo 9 se presentan las conclusiones y los trabajos futuros de esta tesis.

Capítulo 2

Antecedentes

Este capítulo plantea el estado actual de la investigación en el dominio de esta tesis, y pone de manifiesto los problemas y las contribuciones que aportan a solucionarlos. De este modo, provee el conocimiento previo requerido para entender los siguientes capítulos. Se puede distinguir dos bloques principales: uno que explica el marco teórico de la privacidad y otro enfocado en el proceso de ingeniería. El primero permite entender y diferenciar los conceptos de privacidad, y el marco legislativo que los soporta. El segundo se ocupa de explicar el paso del nivel abstracto de la legislación hacia la ingeniería de sistemas, enfocado en la actividad de diseño.

El primer bloque inicia con la explicación del concepto de privacidad, mostrando sus diferentes características y tipos. Se abordan los datos personales y su relación con la privacidad. A continuación, se explica la diferencia entre la protección de la privacidad y la de los datos personales.

Luego, se presenta una breve cronología de las legislaciones, principalmente concentradas en proteger los datos personales, hasta llegar al Reglamento General de Protección de Datos en Europa (RGPD), mandatorio desde el 25 de mayo de 2018.

El segundo bloque inicia con un tema que sirve como punto de conexión entre lo abstracto de la legislación y lo operacional de la ingeniería: la Privacidad desde el Diseño.

A continuación, se describe como se aborda la privacidad en el desarrollo de los sistemas de información, tema principal del área de investigación de la Ingeniería de Privacidad. Se abordan de manera específica las actividades de análisis y diseño que son las primeras en el proceso de desarrollo.

Posteriormente, se describen los patrones, uno de los tipos de elementos que más relevancia han adquirido en el diseño de sistemas respetuosos con la privacidad. Luego de explicar sus características, formas de evolución y ciclo de vida, se abordan aquellos patrones enfocados en el dominio de la privacidad.

Se describen los patrones de privacidad como elementos en plena evolución y que están siendo usados como herramientas o dentro de métodos para el diseño de sistemas respetuosos con la privacidad. Se detalla el estado actual y los problemas encontrados de cara a su aplicación.

Finalmente, se presenta un resumen y las conclusiones.

2.1. Privacidad

La privacidad es un concepto difícil de definir [Finn13], incluso se ha llegado a decir que “nadie puede articular lo que significa” [Solove06]. Esta dificultad se puede atribuir, entre otros, a su carácter: subjetivo, ya que depende del ámbito que cada persona considere como privado; multi-dimensional, ya que hay diferentes enfoques y tipos de privacidad; y contextual, ya que depende del contexto en el cual se encuentre la persona. Una de las dimensiones de la privacidad que ha obtenido mucha relevancia, es la que corresponde a los datos que permiten identificar a un individuo: los datos personales que, al ser parte del ámbito privado de una persona, deben ser protegidos. En los siguientes párrafos desarrollamos los conceptos de privacidad, datos personales, y su protección.

Respecto a la subjetividad, partimos de la Declaración Universal de Derechos Humanos de 1948 que en su artículo 12 establece que “Nadie será objeto de injerencias arbitrarias en su vida privada, su familia, su domicilio o su correspondencia ni de ataques a su honra o a su reputación. Toda persona tiene derecho a la protección de la ley contra tales injerencias o ataques” [UnitedN48]. Asimismo, la Real Academia Española (RAE) define a la privacidad como “el ámbito de la vida privada que se tiene derecho a proteger de cualquier intromisión”. En ambos casos hay referencia explícita a lo privado, que según la RAE es algo “particular y personal de cada individuo”. Estas definiciones permiten hacerse la idea del alto componente subjetivo de la privacidad, ya que, depende de cada individuo el alcance del ámbito personal a proteger. Por ejemplo, alguien podría ser muy reticente a compartir su correo electrónico personal en la web, mientras otro podría entregarlo siempre para suscribirse a boletines de ofertas.

En 1999, Clarke [Clarke97] se refirió al carácter multidimensional de la privacidad. Él define la privacidad como “el interés que tienen los individuos para mantener el espacio personal, libre de interferencias de otras personas y organizaciones”. Además, menciona explícitamente que el término ‘espacio’ no tiene un ámbito singular, sino, uno multidimensional y amplio.

Clarke [Clarke97] identificó cuatro tipos de privacidad: *de la persona* (el cuerpo), *del comportamiento personal*, *de las comunicaciones personales*, y *de los datos personales* (Figura 2.1). Posteriormente, en 2013, identificó una nueva dimensión a la que se refiere como la *privacidad de la experiencia personal*. La *privacidad de la persona* se refiere a la integridad del cuerpo de una persona; por ejemplo, una transfusión de sangre sin consentimiento sería una forma de violación de esta privacidad. La *privacidad de comportamiento* se enfoca en las maneras de comportarse en aspectos sensibles; por ejemplo, prácticas sexuales, políticas, etc. La *privacidad de las comunicaciones* se refiere a la capacidad de comunicarnos entre nosotros, a través de diferentes medios, sin ser monitoreados. La *privacidad de los datos personales* se refiere a la capacidad que deberíamos tener para controlar la recolección y el procesamiento de nuestros datos, aunque estos estén en posesión de otras personas u organizaciones. Finalmente, la *privacidad de la experiencia* se refiere a la privacidad de las acciones que realizamos y a través de las cuales se obtiene experiencia. Por ejemplo, al comprar cosas, al leer artículos, etc. Aunque Clarke [Clarke97] señala que la privacidad de datos personales suele también ser referida como *privacidad de la información*, considera que es más útil el usar este último para englobar la privacidad de las comunicaciones y de los datos.

En la misma línea, Finn et al. [Finn13] expandió el conjunto de tipos de privacidad de Clarke [Clarke97], agregando la *privacidad de pensamientos y sentimientos*, *de localización y espacio*, y la *de asociación*. La *privacidad de pensamientos y*

sentimientos se enfoca en el derecho de las personas a compartir o no sus pensamientos y emociones. La *privacidad de localización y espacio* se refiere al derecho de movernos en ambientes públicos o semipúblicos sin ser monitoreados. Finalmente, la *privacidad de asociación* se refiere al derecho de las personas de reunirse con quien quieran sin ser monitoreadas.

Figura 2.1: Tipos de privacidad acorde a Clarke

Solove [Solove06] adopta una perspectiva centrada en la información (los datos) y aborda la privacidad en términos de daños a la misma. En el contexto de la sociedad norteamericana y su legislación, Solove propone una taxonomía de actividades dañinas a la privacidad. Estas son agrupadas en cuatro categorías: *recolección de información*, *procesamiento de información*, *diseminación de información* e *invasión*. Dentro de la *recolección de información* incluye la vigilancia continua y la interrogación como forma coercitiva de obtener información. Respecto al *procesamiento de información* incluye la agregación de información de la persona, la identificación de la persona real, la inseguridad de cara a posibles futuros daños, el uso secundario de la información (es decir, el uso diferente a lo inicialmente aceptado) y la exclusión de la persona, al no ser correctamente informado sobre el uso de sus datos y no permitirle hacer algo al respecto. La *diseminación de información* incluye la violación de la confidencialidad de la persona, la divulgación de su información, la exposición de atributos físicos o emocionales primordiales de la persona, el incremento en la accesibilidad, el chantaje, la apropiación de la identidad o personalidad de un individuo, y la distorsión de cómo una persona es apreciada. Finalmente, la *invasión* incluye actividades dañinas como la intrusión o invasión en la vida (p. ej. estado y actividades) de la persona, y la interferencia en sus decisiones por parte del gobierno.

Por otro lado, Nissenbaum [Nissenbaum10] se refiere al carácter contextual de la privacidad; ella define a la privacidad como el “derecho a un flujo apropiado de la información personal”. Además, menciona que dicho flujo varía dependiendo del contexto, a lo que se refiere como integridad contextual. Esta tiene como elementos los contextos en los cuales se produce el flujo y las normas que establecen dicho flujo. La idea es que las personas (adoptando determinados roles) se comportan acorde a un contexto social, por ejemplo, de salud, de educación, etc. Las normas dependen del tipo de información, del rol de la persona, del remitente, del destinatario, y de los principios

bajo los cuales la información es transferida. De este modo se establecen restricciones sobre la recolección, uso y diseminación de la información personal.

Aunque los investigadores hayan abordado la privacidad desde diferentes perspectivas, y resulte complejo el llegar a una sola definición; parece haber un acuerdo de que es un elemento subjetivo, con múltiples dimensiones [Finn13], y contextual. Estas características se han tomado como punto de partida, a través del cual se han podido elaborar discusiones y realizar nuevas contribuciones. Por ejemplo, se ha analizado la importancia de la privacidad [Clarke97] desde diferentes enfoques:

- Psicológicamente. Debido a que la gente necesita un espacio privado para pensar y comportarse, independiente de que esté en un contexto público o no.
- Sociológicamente. Los seres humanos somos criaturas sociales, y necesitamos tener la confianza de poder comportarnos y asociarnos con otros sin temor de que estemos siendo monitoreados. La gente está cada vez más preocupada en su privacidad; como muestra, siete de cada diez ciudadanos europeos consideran su privacidad como muy importante [TNSPS16].
- Económicamente. Por un lado, las empresas están basando sus modelos de negocio en el uso de la información, que es necesario de cara a seguir brindando servicios. Por otro lado, deben proveer medios a través de los cuales garanticen la privacidad de las personas, minimizando posibles amenazas a su privacidad.
- Políticamente. Ya que las limitaciones a la privacidad de pensamiento y comportamiento, se pueden traducir en limitaciones a manifestarse, lo cual, puede amenazar la democracia.

Los tipos de privacidad y sus afectaciones no están aisladas o relacionadas unívocamente. De hecho, abordar un tipo de privacidad puede impactar varios aspectos de la persona. Por ejemplo, afectar la *privacidad de comportamiento* tendría implicaciones psicológicas, sociales y políticas. Además, los tipos de privacidad pueden estar relacionados entre sí. Dada la digitalización actual, la *privacidad de los datos* personales puede afectar el resto de tipos. Por ejemplo, hay datos personales que son generados de características corporales como el iris del ojo o una huella dactilar, o de las ubicaciones donde ha estado una persona y con quién se ha reunido. De este modo una afectación a los datos personales puede afectar la *privacidad de la persona* (el cuerpo) y la *privacidad de comportamiento*.

2.1.1. Datos personales

Los datos personales cobran gran relevancia en la sociedad digitalizada actual. Muchas de nuestras acciones se convierten en datos que nos pueden identificar. En efecto, los datos personales se definen como “cualquier información relacionada a una persona física identificada o identificable” [EuroPC16]. Es decir, información que, directa o indirectamente, e individual o colectivamente, lleva a la identificación de una persona. Por ejemplo, un nombre, el número de la seguridad social, una coordenada de localización, un patrón de comportamiento, etc.

Hay diversas formas de categorizar los datos personales: según la sensibilidad de los datos, el sujeto a quien pertenece los datos, los propósitos de uso de los datos, el contexto en el cual los datos son procesados, el grado de identificación, entre otros [OECD14]. Vale la pena resaltar la categorización basada en el origen de los datos reportada en [OECD14], como un esfuerzo para contribuir con una taxonomía que

ayude de manera práctica a la implementación de los principios de privacidad. Las categorías de datos propuestas incluyen:

- Aquellos datos que son *provistos* de manera consciente por acciones directas de la persona. Por ejemplo, datos como su dirección de correo electrónico, su fecha de nacimiento, provistos a través de un formulario.
- Aquellos datos que son *observados* y almacenados cuando un usuario usa o ha usado un servicio o dispositivo. Por ejemplo, su historial de búsqueda en la web, el registro de uso de un dispositivo.
- Los datos *derivados*, usando un razonamiento simple o matemáticas básicas, a partir de otros datos. Por ejemplo, la frecuencia de compra, o, la edad de una persona basada en su fecha de nacimiento.
- Los datos *inferidos* después de un proceso de análisis basados en la probabilidad. Por ejemplo, el perfil de riesgo crediticio de una persona.

Hoy en día, los datos personales tienen un alto valor; incluso, en 2009, la Comisaria Europea del Consumidor ha llegado a definirlos como “el nuevo petróleo de Internet y la nueva moneda del mundo digital” [Kuneva09]. Por un lado, las organizaciones están muy interesadas en explotar los datos personales, para entender a los usuarios, mejorar su experiencia de uso, y ofrecerles nuevos productos y servicios. Por otro lado, los usuarios son cada vez más conscientes de sus datos personales, y de las implicaciones en su privacidad [TNSPS16]. Por ende, tanto las organizaciones como las personas están altamente interesadas en proteger su privacidad y sus datos personales.

2.1.2. Protección de la privacidad y de los datos personales

La privacidad y los datos personales están relacionados, pero no son sinónimos. Asimismo, no son equivalentes la protección de la privacidad y la de los datos personales. Para aclarar la diferencia, es necesario profundizar en la definición de protección de la privacidad hasta llegar al nivel de los datos, y establecer entonces sus diferencias.

La importancia de la privacidad no tiene únicamente un carácter individual, sino también, uno grupal, y uno de sociedad. Clarke [Clarke97] llega a definir la *protección de la privacidad* como “un proceso de encontrar el balance adecuado entre la privacidad y múltiples intereses”. Estos intereses incluyen aquellos relativos o no a la privacidad, de la misma persona (p. ej. la privacidad y el obtener una mejor tasa de crédito), de un grupo (p.ej. diferentes niveles de privacidad requeridos por varias personas en un ambiente colaborativo) [Such18] o de la sociedad (p. ej. privacidad y uso de registros de salud ante una emergencia de salud pública).

Enfocándose en la privacidad de la información (privacidad de datos y comunicaciones), Clarke [Clarke97] la define como “el interés que la persona tiene en controlar, o al menos influenciar significativamente, el tratamiento de sus datos”. Por consiguiente, la *protección de la privacidad de los datos* implicaría encontrar el balance en el tratamiento de los datos de la persona y otros intereses. Este enfoque, manteniendo la subjetividad del término privacidad, vela por el interés y protección de la persona.

Por otro lado, la llamada *protección de datos*, dictada en algunas legislaciones, es abordada desde una perspectiva más pragmática que se concentra en proteger los datos acerca de las personas y no a las personas [Clarke97] (Figura 2.2).

Figura 2.2: Protección de la privacidad y de los datos personales

2.2. Legislación

La privacidad y los datos personales tienen un marco legislativo que vela por su protección. La privacidad ya se reconoce desde la Declaración Universal de Derechos Humanos de 1948 [UnitedN48]. Asimismo, tomando un enfoque más pragmático, distintas legislaciones han abordado los datos personales; y su auge se ha dado a raíz de la proliferación del uso de los computadores. Aunque los primeros principios se dieron en los Estados Unidos de América, pronto fueron acogidos por la Organización para la Cooperación y el Desarrollo Económico (OECD –Organisation for Economic Cooperation and Development), y posteriormente, han sido adoptados por la Unión Europea y sus Estados miembros. A continuación, se brinda una breve cronología.

Ya desde que los computadores comenzaron a proliferar en los años 70, se comenzó a pensar en las implicaciones que la recolección y tratamiento masivo de datos personales podría tener en la privacidad. Es por ello que en los Estados Unidos de América se estableció un conjunto de principios sobre la protección de información personal (*FIPPS –Fair Information Practice Principles*) en 1973 [USDHEW73], que luego fueron parte fundamental de su Acta de Privacidad –*Privacy Act*– de 1974. Los principios básicos fueron cinco:

- *transparencia*, para informar a la persona sobre la recolección de sus datos;
- *elección*, que brinda a la persona la posibilidad de decidir, o denegar, cómo se usarán sus datos;
- *revisión y corrección de la información*, que permite a la persona acceder y rectificar sus datos;
- *protección de la información*, que se refiere a salvaguardar la calidad e integridad de los datos de la persona;
- *rendición de cuentas*, que responsabiliza a las organizaciones del apego y cumplimiento a estos principios.

Estos principios fueron tomados como base dentro de las *directrices de protección de la privacidad y flujos transfronterizos de datos personales por la OECD* en 1980 [OECD80]. Dentro de sus directrices, se establecieron un conjunto de ocho principios:

- *limitación de recogida* de datos personales, estos deben ser obtenidos de manera lícita y justa, y con el conocimiento y consentimiento de la persona.
- *calidad de los datos* personales, estos deben ser pertinentes para los fines de uso, y deben ser exactos, completos y actualizados.

- *especificación de propósito* para los cuales los datos serán recogidos, esto debe realizarse antes de la recolección.
- *limitación de uso* de los datos personales, incluyendo su divulgación, acorde a lo especificado en el propósito. Sin embargo, esto podría cambiarse con el consentimiento de la persona o en virtud de la ley o de las autoridades.
- *salvaguardia de la seguridad* a través del uso de medidas razonables de seguridad para proteger los datos personales contra riesgos como la pérdida, acceso no autorizado, uso, modificación o divulgación.
- *transparencia* en cuanto a la evolución, prácticas y políticas relacionadas a los datos personales.
- *participación individual*, que se refiere al derecho de la persona a que se le confirme sobre la tenencia de sus datos; a que se le comuniquen sus datos (en un tiempo y forma razonables, a un precio no excesivo en caso de haberlo, y de manera inteligible); a que, en caso de denegación, se le explique y que pueda cuestionarla; a que pueda expresar dudas sobre sus datos, y que, en caso de una petición exitosa, sus datos se puedan eliminar, rectificar, completar o corregir.
- *responsabilidad* del controlador de los datos para hacer efectivos estos principios.

Específicamente, en el marco de la Unión Europea (UE), se adopta en 1995 la *Directiva de Protección de Datos (Directiva 95/46/CE)* [EuroPC95]. Esta estableció un marco regulador, que incorporaba los principios de la OECD, para la protección y la libre circulación de datos personales dentro de la Unión Europea. Al ser una directiva, ésta debía ser traspuesta al ordenamiento legal de cada país miembro para que defina sus propios medios de cara a conseguir los objetivos planteados en la directiva. En España, esto se llevó a cabo con la entrada en vigor de la *Ley Orgánica de Protección de Datos de Carácter Personal* de 1999 [JefEstEsp99].

Para evitar fragmentaciones y para tener una normativa vinculante para todos los Estados de la Unión Europea, entró en vigor el 25 de mayo de 2016 el *Reglamento General de Protección de Datos* [EuroPC16], y es de obligado cumplimiento desde el 25 de mayo de 2018. Este reglamento establece los requisitos que deben cumplir las empresas y organizaciones que recolectan y procesan datos personales. Uno de los aspectos más relevantes es que se aplica a todas las organizaciones, independientemente de su ubicación, que traten datos de residentes europeos. Además, incrementa el nivel de responsabilidad de los responsables del tratamiento de los datos, e incluso, impone multas que pueden alcanzar el 4% de la facturación mundial anual de la organización. En línea con los principios FIPPs y de la OECD, el RGPD incluye los siguientes:

- la *licitud y transparencia* en el procesamiento de datos, y la *lealtad* hacia la persona,
- la *limitación de finalidad* de tratamiento de los datos previo a su recolección,
- la *minimización de datos* a únicamente aquellos necesarios para cumplir la finalidad,
- la *exactitud de los datos* y que estos estén actualizados, o se permita su corrección, para cumplir la finalidad,
- la *limitación del plazo de conservación* de los datos personales solo hasta cumplir con la finalidad establecida,
- la *integridad y confidencialidad* de los datos personales a través de mecanismos técnicos y organizacionales que garanticen su seguridad.

Con el fin de adaptar el ordenamiento jurídico español al RGPD, se encuentra en vigor desde el 5 de diciembre de 2018 la *Ley Orgánica de Protección de Datos Personales y garantía de los derechos digitales*.

Finalmente, vale la pena mencionar la actual propuesta (aún en discusión) de protección de los datos personales en el sector de las comunicaciones electrónicas [EuroC17] de 2017, comúnmente denominada *Regulación ePrivacy*. Esta regulación es complementaria al RGPD, centrándose principalmente en las comunicaciones electrónicas. La regulación ePrivacy tiene un carácter más amplio que la Directiva en Privacidad y Comunicaciones Electrónicas [EuroC02] del 2002, a la que reemplazaría. Por ejemplo, incluye, además de los servicios de telecomunicación tradicionales, a los servicios de comunicación tales como WhatsApp, Facebook Messenger, Skype, etc.

El enfoque para abordar la protección de datos en los sistemas ha cambiado con el tiempo. Antes, se dejaba la responsabilidad del cumplimiento normativo en manos de abogados, quienes por lo general redactaban complejas políticas de privacidad para llevar la protección de datos a cabo. Sin embargo, las legislaciones actuales ponen de manifiesto dos aspectos clave: el llamado al uso de medios técnicos y su implementación desde el inicio de la concepción de los sistemas. Por ejemplo, el Artículo 25 (Protección de datos desde el diseño y por defecto) del RGPD indica que el responsable del tratamiento de datos personales aplicará “... *medidas técnicas y organizativas apropiadas [...] a fin de cumplir los requisitos del presente Reglamento y proteger los derechos de los interesados*.”. Además, estas medidas de protección deben establecerse desde la concepción de los sistemas (que tratan datos personales) y durante su operación, lo cual está alineado con el concepto de *Privacidad desde el Diseño* (PbD –Privacy by Design) [Cavoukian09].

2.3. Privacidad desde el diseño

La necesidad de tomar medidas técnicas, para velar por la privacidad y protección de datos de las personas, ha sido reconocida por los comisionados de privacidad y protección de datos alrededor del mundo [DPPC10]. Ellos reconocen a la *Privacidad desde el Diseño* como un componente clave en la protección de la privacidad; además, han acordado promover la adopción de sus principios fundamentales en las legislaciones dentro de su jurisdicción.

El concepto de *Privacidad desde el Diseño* promueve la introducción de las consideraciones de privacidad desde el inicio de los proyectos que involucren información personal, y no al final de los mismos [Cavoukian09]. Este concepto fue enunciado inicialmente por Ann Cavoukian, en ese entonces responsable de la Autoridad de Protección de Datos de Ontario en Canadá [Cavoukian09], y ha sido acogido luego en regulaciones como el RGPD (Art. 25) [EuroPC16]. Este enfoque lista siete principios:

- Proactivo en lugar de reactivo; prevenir en lugar de remediar. Hay que prever y prevenir los daños para la privacidad antes que estos sucedan. Es decir, se debe trabajar con anterioridad al daño, y no después de que este suceda.
- Privacidad por defecto. Los datos personales deben estar protegidos por defecto, antes incluso de las acciones del usuario.
- Privacidad integrada en el diseño. Los aspectos de privacidad deben ser considerados en el diseño de los sistemas de información, en lugar de aparecer como añadidos o complementos una vez que el sistema está construido.

- Funcionalidad completa; suma positiva no suma cero. Se busca acomodar los intereses y objetivos legítimos de las distintas partes, en lugar de sacrificar unos en nombre de otros (p.ej. privacidad frente a seguridad).
- Protección durante el ciclo de vida completo. Al integrar la privacidad en el diseño se puede proteger la información personal incluso antes de recogerla. Se busca brindar esta protección durante el ciclo completo de la información, incluyendo su destrucción.
- Visibilidad y transparencia. El sistema de información debe funcionar acorde a los objetivos que declara y debe permitir la verificación de ello.
- Respeto a la privacidad del usuario. Sobre todo, los desarrolladores e ingenieros deben considerar los intereses de los usuarios como lo más importante, puesto que protegerles es la finalidad de las medidas adoptadas.

A pesar de la adopción del PbD en la regulación, los ingenieros tienen problemas implementándolo [Hadar18][Senarath18]. Entre los problemas que se reportan están: la necesidad de tener requisitos operacionales en lugar de aquellos abstractos provenientes de la legislación, la falta de conocimiento y entendimiento en el área de privacidad de la información y protección de datos, y la falta de mecanismos que les guíen en el proceso de desarrollo de sistemas respetuosos con la privacidad.

Por un lado, se hace necesario el establecer mecanismos de educación que mejoren sus conocimientos y habilidades en el área de privacidad y protección de datos, o fomentar el uso de mecanismos que les permitan establecer soluciones a pesar de que su conocimiento en privacidad sea limitado. Por otro lado, se requiere desarrollar métodos sistemáticos y herramientas de soporte que permitan a los ingenieros lidiar con los problemas de privacidad y protección de datos en las diferentes actividades del proceso de desarrollo. Estas sugerencias constan en las recomendaciones establecidas por la ENISA [Danezis14].

2.4. Privacidad en la ingeniería de sistemas

La privacidad desde el diseño contempla la implementación de medidas (técnicas y organizacionales) que garanticen la privacidad de la persona desde el diseño del sistema y durante todo el ciclo de vida de los datos involucrados. Así, el punto inicial para tomar medidas son las actividades iniciales del ciclo de vida de los sistemas, y el punto final estará marcado por el final del ciclo de vida de los datos personales. Las actividades del ciclo de vida de un sistema incluyen las de desarrollo más otras como el despliegue, mantenimiento, soporte, evolución y retiro; y actividades que deben realizarse a lo largo de todo el ciclo como la administración de configuración y el aseguramiento de la calidad [IEEE14]. Dado el alcance de esta tesis, a continuación, nos enfocamos en dos de las primeras actividades del ciclo de vida del desarrollo: el análisis y el diseño. Inmediatamente después, se presenta el estado actual en las dos actividades desde la perspectiva de la ingeniería de privacidad.

2.4.1. Análisis y diseño de sistemas

El desarrollo de los sistemas comprende aquellas actividades desde la especificación requerimientos hasta la entrega de un producto acorde a ellos [IEEE14]. Estas actividades incluyen la especificación de requisitos, el diseño, la construcción, las pruebas, y la entrega del producto.

La actividad de especificación de requisitos (abordada por la ingeniería de requisitos) se encarga de las tareas relacionadas con el descubrimiento, análisis, documentación y

verificación de los requisitos en un sistema desarrollado [Sommerville11]. Los requisitos pueden ser considerados como características del sistema, y pueden ser categorizados en dos grandes grupos: funcionales y no funcionales.

Los requisitos funcionales son aquellos que definen lo que el sistema debería hacer. Por ejemplo, un editor de texto. Los requisitos no funcionales o requisitos de calidad [IEEE14], por otro lado, se expresan como limitaciones sobre las funcionalidades ofrecidas. Por ejemplo, requisitos relacionados con el rendimiento o a la seguridad del sistema. Los requisitos relacionados con la privacidad son considerados como requisitos no funcionales. Las restricciones de privacidad podrían venir del entorno legal, de la organización, de la sociedad, de la persona, etc. como se vio en la Sección 1.1.

La actividad de diseño se enfoca en analizar los requerimientos, definir la estructura del sistema en función de componentes, y describir su comportamiento [IEEE14]. Para esto, se define la arquitectura, los componentes, las interfaces, y las características necesarias que ayuden a representar el sistema a construir. En el proceso de diseño se pueden usar diferentes técnicas de cara a facilitar su realización. Así, para obtener una mejor definición del sistema, se pueden usar varios tipos de modelos que presentan diferentes perspectivas. Por ejemplo, modelos de estructura, modelos de comportamiento o modelos de información. Para tomar ventajas de las experiencias previas y acelerar su proceso de definición, se pueden usar patrones. Por ejemplo, el patrón de arquitectura cliente-servidor.

2.4.2. Ingeniería de privacidad

La *Ingeniería de Privacidad* se encarga de la integración de la privacidad en el proceso de ingeniería de los sistemas [ISO/IEC27550]. La investigación en el área se encarga del estudio de los enfoques sistemáticos que permitan llevar a cabo dicha integración [Gürses16]. Los elementos que fundamentan la *Ingeniería de Privacidad* son las teorías, métodos, técnicas y herramientas.

Las *teorías* de ingeniería de privacidad ayudan a conceptualizar la privacidad; estas han expuesto su carácter subjetivo, multi-dimensional y contextual (Sección 1.1). Las distintas legislaciones se han enfocado en la protección de datos personales como una vía más pragmática de la privacidad. Estas legislaciones (Sección 1.2), e incluso estándares, han llegado a establecer un conjunto de principios que deberían ser cumplidos por los sistemas que recolecten y procesen datos personales.

Los *métodos* de ingeniería de privacidad definen los enfoques sistemáticos a seguir, para abordar cuestiones de privacidad que surjan en cada etapa del desarrollo o ciclo de vida de un sistema. Acorde a [ISO/IEC24744], un método debería definir procesos a seguir, los productos a obtener, y las personas a cargo.

Las *técnicas* de ingeniería de privacidad son procedimientos, que podrían estar soportados por lenguajes o notaciones, y que permiten llevar a cabo una tarea para cumplir un objetivo.

Las *herramientas* de ingeniería de privacidad son instrumentos (automatizados) que brindan soporte a los ingenieros para llevar a cabo sus responsabilidades dentro de un método seguido.

Aunque la *Ingeniería de Privacidad* debe encargarse del estudio de los elementos sistemáticos a lo largo de todas las actividades en el desarrollo de sistemas; las tendencias actuales muestran que la mayoría de esfuerzos se han enfocado en las actividades de análisis de requisitos y de diseño [Morales-Trujillo19].

Análisis

La actividad de análisis cobra mucha relevancia debido a que está íntimamente ligada al establecimiento de requisitos, sobre los cuales debe diseñarse y construirse un sistema. Los requisitos de privacidad y protección de datos pueden provenir de diferentes fuentes. Así, el RGPD impone ciertas obligaciones que los responsables del tratamiento deben implementar en los sistemas. Estándares [ISO/IEC29100] y buenas prácticas internacionales [OECD14] también se traducen en requisitos que podrían ser implementados. Otras fuentes incluyen las políticas internas de las organizaciones [Young10], heurísticos derivados del estudio de usuarios [Gürses08], y los propios requisitos de los usuarios.

El estado del arte reporta dos enfoques complementarios [Shostack14][Notario15] para la elicitación de requisitos de privacidad: el basado en riesgos y el orientado a metas. Los dos consideran un conjunto de principios de privacidad obtenidos de marcos legales, éticos o filosóficos y pasan a transformarlos a requisitos operacionales siguiendo caminos distintos (la Figura 2.3 muestra el enfoque de [Notario15]).

Figura 2.3: Enfoques de elicitación de requisitos de privacidad

El enfoque basado en riesgos consiste en identificar un conjunto de problemas que deben ser solucionados (o contrarrestados). Acorde a [Notario15], inicialmente, se identifican los activos a proteger en el sistema en desarrollo y las amenazas que podrían comprometer la privacidad y la protección de datos. Por ejemplo, considerando el principio de confidencialidad de los datos personales, una amenaza sería que alguien no autorizado acceda a ellos. Las amenazas deben ser analizadas para determinar la relevancia de su riesgo asociado; esto se podría realizar considerando su probabilidad e impacto. Luego, se define un tratamiento, que sería el requisito operacional, para abordar el riesgo asociado a la amenaza. El tratamiento puede ser el no hacer nada

(aceptando el riesgo) o incluir medidas, en diferentes niveles, para evitar o reducir el riesgo (modificando su impacto o su probabilidad).

Por otro lado, en un enfoque orientado a metas, la intención es definir las metas a conseguir en un sistema que se está desarrollando. En el caso de considerar legislaciones o estándares, se describe cada principio como una meta de los responsables de datos. Estas metas deberían ser alcanzadas de cara a cumplir las obligaciones que impone la legislación. Asimismo, las organizaciones o el usuario, podrían definir directamente sus metas. Aquellas metas de alto nivel (principios) puede ser deconstruidas en un conjunto de directrices requeridas para poder alcanzar la meta, y cada directriz podría ser dividida en un conjunto de requisitos. Estos requisitos serán atendidos dependiendo de su nivel de prioridad.

Algunos de los métodos para el análisis de aspectos de privacidad reportados hasta el momento son: LINDDUN [Deng11], NFR [Cysneiros04], PriS [Kalloniatis08], Tropos y GSRM [Islam10], PRIPARE [Notario15]. Además, recientemente, en el año 2017, se ha publicado un estándar ISO con guías para realizar evaluaciones de impacto en la privacidad [ISO29134].

Diseño

Desde la perspectiva del análisis de requisitos de privacidad, en el diseño se tienen que establecer controles de privacidad como contramedidas (enfoque basado en riesgos) o medidas (enfoque orientado a metas) para satisfacer los requisitos especificados. Sin embargo, el término control de privacidad podría quedar muy limitado, dado que en la actividad de diseño se deben definir los elementos que describen la estructura y comportamiento del sistema, tanto a nivel de arquitectura como de componentes.

Diferentes métodos han sido propuestos, algunos de ellos se han enfocado exclusivamente en la actividad de diseño [Iachello05][Moncrieff09]; otros han abordado de manera conjunta el análisis y el diseño [Kalloniatis08] [Aljohani16]; y algunos han abordado el diseño dentro del proceso completo de desarrollo [Fhom11], [Notario15]. Hasta donde sabemos, entre las propuestas, no hay enfoques bien marcados; en su lugar, hay mucha variabilidad. Por ejemplo, en el elemento que se aborda. Mientras algunos se esfuerzan en definir los productos (p. ej. modelos, arquitecturas de referencia, patrones de diseño), otros se enfocan en los procesos y personas (p. ej. fases, roles).

En la misma línea, los elementos de métodos que han sido propuestos en la literatura también presentan diversidad. Así, hay propuestas de modelos de referencia [Karjoth02][Gope16], estrategias de diseño [Hoepman14], patrones (de privacidad) [Chung04][Romanosky06], que contribuyen a conformar la base de conocimiento de privacidad que se tiene actualmente.

Estos son problemas asociados al hecho de tener un área de investigación nueva e incipiente. Por un lado, aunque hay la impresión que la actividad de diseño va ganando atención por parte de los investigadores, no se conoce a ciencia cierta cómo va su crecimiento. Por otro lado, las propuestas en la actividad de diseño parecen tener mucha dispersión en torno a los elementos que abordan. Entonces, es necesario ordenar el conocimiento para determinar sus tendencias de crecimiento, enfoques adoptados; en general, tener conciencia de esta área de investigación, similar a lo que ha sucedido en la actividad de análisis de requerimientos [Kalloniatis09][Anthonysamy17]. Esto sería útil para los investigadores para identificar oportunidades de investigación, y para los ingenieros de cara a conocer la madurez de los elementos que pueden guiarles en el diseño sistemático de sistemas respetuosos con la privacidad.

A pesar de esta dispersión en las propuestas, vale la pena resaltar que varias de ellas han dado mucha relevancia a los patrones. Los patrones (de privacidad) han sido tenidos en cuenta en varios proyectos (p. ej. PRIS [Kalloniatis08], PrimeLife [Fischer-Hübner10], PRIPARE [Notario15]), siendo considerados elementos reusables útiles en el proceso de desarrollo de sistemas respetuosos con la privacidad.

2.5. Patrones de privacidad en el diseño de sistemas

El origen de los patrones se remonta al dominio de la arquitectura [Alexander77], luego de lo cual, se han movido a muchos otros como la ingeniería de software [Gamma95], o más recientemente, al de la seguridad [Yoder98] o al de la privacidad. Los patrones no permanecen aislados, sino que, evolucionan llegando a formar conjuntos cohesivos que tratan de abarcar completamente la creación de una solución. Actualmente, en el área de ingeniería privacidad, varias propuestas han considerado los patrones dentro de sus procesos y métodos. Las siguientes secciones desarrollan estos temas.

2.5.1. Patrones

Christopher Alexander [Alexander77] menciona que un patrón “describe un problema que ocurre una y otra vez en nuestro entorno, y luego describe el núcleo de la solución a ese problema, de tal forma que puedes usar esta solución un millón de veces, sin hacerlo de la misma manera dos veces”. Además, indica que, son una forma de sistematizar conocimiento en base a la experiencia previa colectiva de expertos.

Aunque el concepto surgió en el dominio de la arquitectura, este se ha movido rápidamente hacia otros; entre ellos, la ingeniería de software [Gamma95], y la ingeniería de seguridad [Yoder98]. La definición ha sido adaptada dependiendo del dominio, pero manteniendo su esencia. Por ejemplo, un patrón de diseño orientado a objetos “es una descripción de objetos y clases comunicantes que están configurados para resolver un problema de diseño general en un contexto particular” [Gamma95]. Asimismo, “un patrón de seguridad describe un problema particular y recurrente de seguridad que surge en un contexto de seguridad específico y presenta un esquema genérico bien probado para una solución de seguridad” [Yoder98].

En el contexto de esta tesis, nos referimos a un patrón de privacidad como un *elemento reusable que describe la solución a un problema recurrente en un contexto de privacidad específico*. La Figura 2.4 muestra un ejemplo bastante didáctico de un patrón de privacidad tomado del catálogo de patrones de privacidad reportado en [Fischer-Hübner10]. Asimismo, la Figura 2.5 muestra un ejemplo de implementación de este patrón en la red social Facebook.

Vale la pena notar que los patrones son descritos a través de un conjunto de campos que conforman su estructura. Cada uno de los campos tiene su respectivo contenido en aras de ser útiles a los desarrolladores. Aunque los campos que se utilicen en los patrones dependerán de los dominios específicos a los cuales estos pertenezcan, sí se puede discriminar entre dos tipos: aquellos obligatorios y los que son opcionales. Según [Meszaros97] la presencia de los campos obligatorios ayudará a que los patrones sean entendidos indistintamente de variaciones en el estilo adoptado por los autores, y los campos opcionales ayudarán a que se maximice el entendimiento. En el caso del ejemplo de la Figura 2.4, sus autores [Fischer-Hübner10] utilizan el campo *patrones relacionados* como opcional, mientras el resto son obligatorios.

Nombre	OPCIONES DE PRIVACIDAD EN REDES SOCIALES
Problema	En las redes sociales los usuarios tienen la posibilidad de proporcionar datos sobre sí mismos, pero no todos los datos deben ser visibles para todos los usuarios. Diferenciar qué usuarios están autorizados a ver qué datos es muy importante para la privacidad, así los usuarios podría controlar la visibilidad de su información.
Solución	La solución a este problema es un control de acceso selectivo para las redes sociales.
Contexto	Esta solución debería utilizarse en la implementación de redes sociales para garantizar la privacidad.
Patrones relacionados	- Panel de concienciación de la privacidad en espacios colaborativos - Control de acceso selectivo en foros de software - Programación de grupos respetuosa con la privacidad

Figura 2.4: Ejemplo de patrón de privacidad

Figura 2.5: Ejemplo de implementación de un patrón en una red social

Evolución de los patrones

Los patrones no están aislados [Buschmann96][Schumacher03], se agrupan en catálogos, se interrelacionan formando sistemas, y finalmente, evolucionan a lenguajes de patrones. Un catálogo de patrones tiene como principales objetivos el agrupar y clasificar un conjunto de patrones en función de ciertas categorías [Schumacher03].

Los sistemas, además de tener los patrones organizados, tratan de normalizar e interrelacionar los patrones previamente agrupados. Acorde a [Buschmann96] un sistema debería tener una suficiente cantidad de patrones uniformemente descritos, presentar las diferentes relaciones entre sus patrones, organizar sus patrones para una rápida búsqueda, mostrar cómo aplicar e implementar los patrones, y, poder evolucionar (p. ej. poder soportar la agregación de patrones, permitir la mejora de las descripciones de los patrones, etc.).

Finalmente, un lenguaje de patrones, además de cumplir con las características previas, debería cubrir cada aspecto de un problema que se esté resolviendo (p. ej. un sistema o módulo de software que se esté construyendo). Por tal motivo, el dominio de un lenguaje de patrones usualmente es más específico que el de un sistema de patrones. Es necesario tener en cuenta que el cubrir cada aspecto de un problema que se esté resolviendo hace que el llegar a proponer un lenguaje no sea una tarea en absoluto trivial.

Esta perspectiva de evolución se enfoca en la normalización, clasificación e interrelación entre los patrones, y trata de llegar a un elemento idealizado y complejo: el lenguaje de patrones. Este lenguaje estaría compuesto por un conjunto de elementos auto-contenidos y auto-suficientes, listos para su aplicación, que resolverían un problema; por ejemplo, la construcción de un sistema. Sin embargo, esta perspectiva no muestra algunas actividades necesarias como la identificación de un problema y solución previo a la definición de un patrón, o la revisión de la calidad de un patrón previo a su publicación. Este tipo de actividades pueden ser vistas a través de una perspectiva de ciclo de vida de los patrones.

Ciclo de vida de los patrones

Los patrones atraviesan un conjunto de etapas sistemáticamente organizadas desde su concepción hasta la evaluación de su aplicación (Figura 2.4). Según Yoshioka et al. [Yoshioka08] se pueden organizar las diferentes etapas en un proceso de extracción y uno de aplicación.

Figura 2.6: Etapas del ciclo de vida de un patrón

El proceso de extracción es llevado a cabo por un autor e inicia cuando se identifica, en nuestro conocimiento o en el reportado en un dominio, un problema recurrente en un contexto determinado junto a su solución. Con dicha información se escribe un patrón siguiendo una estructura determinada. Luego, el patrón se organiza acorde a una clasificación y se establecen las relaciones con otros patrones². Entonces el patrón debe ser revisado y corregido para garantizar que sea entendible y que tenga la calidad adecuada. Finalmente, el patrón debe ser publicado en un repositorio.

El proceso de aplicación se lleva a cabo por un usuario del patrón; por ejemplo, un ingeniero, un diseñador o un desarrollador de software. El proceso comienza cuando el practicante, dentro de su proceso normal, reconoce el contexto y problema de un patrón. A través de los medios necesarios, busca y selecciona el patrón más apropiado para

² Consideramos que esta actividad debe ser una etapa per se debido a su importancia en la evolución y aplicación de los patrones.

resolver el problema y procede a instanciarlo. Finalmente, se procede a evaluar los efectos de la instanciación del patrón; por ejemplo, en el elemento que se haya estado construyendo.

Conferencias sobre Patrones

Vale la pena mencionar las Conferencias PLOP –Pattern Languages of Programs– sobre lenguajes de patrones, debido a que estos eventos tienen por objetivo reunir investigadores, autores y usuarios de patrones para tratar los últimos avances en el área. Hay conferencias distribuidas alrededor del mundo; así se tienen eventos en Asia, en Estados Unidos de América, en Latinoamérica, y en el caso de Europa se tiene EuroPLOP –European Conference on Pattern Languages of Programs– [EuroPLOP04]. Las propuestas enviadas ahí son revisadas por pares, y además mejoradas a través de un proceso iterativo conocido como pastoreo –shepherding. Este proceso es un medio para garantizar la calidad de los patrones [Inventado16] y de las propuestas basadas en ellos.

2.5.2. Ingeniería con patrones de privacidad

Los patrones han sido agrupados dentro de distintos tipos de colecciones: catálogos, sistemas y lenguajes. Aunque ya en estas formas se pueden constituir como herramientas para ayudar en la actividad de diseño de sistemas respetuosos con la privacidad, otros elementos de ingeniería como los métodos también los han incluido dentro de sus procesos.

Colecciones existentes

La literatura muestra varios trabajos reportando patrones de privacidad, estos incluyen propuestas de patrones individuales o dispersos y otras que los agrupan dentro de diferentes tipos de colecciones: catálogos, sistemas y lenguajes.

Algunos patrones que han sido reportados de manera individual o dispersa incluyen dos reportados en [Schumacher03], uno sobre el uso de cookies, y el otro sobre el uso de correos electrónicos pseudónimos. En [Sadico05] se presenta un patrón para mostrar al usuario las políticas de privacidad usadas en un sitio. En [Romanosky06] se presentan tres patrones de privacidad útiles en interacciones sobre la web; es decir, en situaciones en las que se solicita un servicio. Por ejemplo, solicitar una página web o enviar un correo. Ya acercándose, a lo que se conocería como un catálogo de patrones, en [Lobato09] se presentan siete patrones para la definición de políticas de privacidad en sitios web.

Schümmer [Schümmer04] muestra un conjunto de 18 patrones de privacidad, de los cuales describe 6, dentro de un lenguaje sobre la cantidad de información compartida en sistemas de interacción entre usuarios. Por ejemplo, cuanta información se muestra sobre las acciones de un usuario en un sistema colaborativo. Los patrones son principalmente narrativos y permitirían el diseño de los componentes de un sistema.

Chung et al. [Chung2004] presentó 15 pre-patrones de privacidad dentro de conjunto de 45 en el dominio de computación ubicua. Los patrones son muy variados en tanto abordan cuestiones organizacionales como la mejora de la confianza de los usuarios, y técnicas relacionadas a la arquitectura del sistema y a las interfaces de usuario por ejemplo. Entre sus ventajas están el uso de una estructura común, y el mostrar los patrones relacionados. Se podrían considerar como un catálogo de pre-patrones de privacidad, entendiéndose que se trata de una propuesta inicial y que aún necesitan mecanismos que garanticen su validez.

Hafiz [Hafiz13] presentó un conjunto de 12 patrones de privacidad que ha alcanzado el nivel de lenguaje (una parte de ellos fueron presentados previamente en [Hafiz06]). El lenguaje se orienta a proveer anonimidad a varios tipos de sistemas, principalmente de comunicación. Por ejemplo, en redes de mezcla –redes que usan servidores proxy para mezclar datos antes de reenviarlos–. Los patrones son arquitecturales y narran soluciones de alto nivel. Cumple con todas características que implican el pertenecer a un lenguaje de patrones, y su calidad ha sido garantizada en un proceso de revisión por pares en un evento PLOP.

El proyecto PrimeLife [Fischer-Hübner10] presentó un conjunto de 15 patrones de privacidad en el dominio de interacción humano-máquina. Los patrones están orientados a ser utilizados en el proceso de diseño de interfaces de usuario usables para sistemas respetuosos con la privacidad. Los patrones presentados en esta colección se orientan a cómo mostrar políticas de privacidad, hacia el uso de íconos para mostrar información relacionada al manejo de la privacidad en un sistema, hacia la interacción con el usuario, y también muestran interfaces de configuración básica de privacidad. Una de las mayores ventajas de este conjunto de patrones, asociado directamente a su dominio, es el uso de bosquejos para la representación de sus soluciones. Esto hace que muchos de los patrones sean fáciles de entender. Otras de las características que tiene este conjunto de patrones es su estructura común, y el establecimiento de relaciones entre sus patrones. En el reporte encontrado, no se podría determinar el soporte a su evolución. Por lo que, se podría considerar como un catálogo bastante desarrollado y maduro, y casi llegando al nivel de sistema.

El proyecto PRIPARE [Notario15] reunió un conjunto de 26 patrones dentro de un repositorio web [PRIPARE15]. Los patrones no se enfocan en un dominio en particular, pero sí están clasificados acorde a estrategias de privacidad [Hoepman14] y etiquetas de los dominios de aplicación (p. ej. correo electrónico, redes de mezcla) en los cuales serían útiles. Cuenta con otras características como el tener una estructura común, el mostrar la existencia de relaciones con otros patrones. Sin embargo, no termina de explicar la implementación de los patrones o el soporte a la evolución de los patrones. Por lo tanto, podría ser considerado como un catálogo de patrones.

Drozd en [Drozd16a][Drozd16b] recolectó un conjunto de 40 patrones de privacidad pertenecientes a varios dominios. Estos incluyen patrones provenientes de la literatura en ese tiempo [Hafiz06][Fischer-Hübner10][Doty11a][PRIPARE15], por lo que, comprende varios dominios de aplicación, y de expertos consultados durante su estudio. Actualmente, esta colección de patrones es accesible a través de una herramienta web, que presenta a los patrones agrupados de manera jerárquica. Así, desde un enfoque de abajo hacia arriba, cada patrón está agrupado dentro de una instrucción, que a su vez está explícitamente listada en los principios de privacidad del estándar ISO/IEC 29100 [ISO/IEC 29100]. Este conjunto de patrones no muestra una estructura común, por lo que no todos reflejan las relaciones existentes; además, presenta únicamente resúmenes de muchos de los patrones. En su estado actual, podría ser considerado como un catálogo de patrones.

Luego que Schümmer [Schümmer04] y Hafiz [Hafiz13] propusieron sus lenguajes en 2004 y 2010 respectivamente, el resto de contribuciones han llegado únicamente al nivel de catálogo. Aunque el desarrollo de nuevos patrones y su agrupamiento es importante de cara a tener una buena base de patrones para trabajar o usar, es indispensable que ellos evolucionen hacia sistemas o lenguajes para lograr un mayor impacto y adopción de los mismos. Además, debido a que los dominios de aplicación en los sistemas o lenguajes son muy específicos, será necesario que haya muchas

propuestas en ese nivel para cubrir la amplia variedad de sistemas que se requieren construir. El evolucionar los catálogos existentes no es una tarea fácil; sin embargo, en este momento parecerían haber ciertas condiciones que favorecerían dicha evolución. Por ejemplo, hay una buena cantidad de patrones sobre los que se podría trabajar, se ha definido un esquema común de representación, entre otras. Muchas de estas se han definido en un esfuerzo común de un grupo de investigadores en el área colaborando en *privacypatterns.org* [Doty11b].

Colaboración hacia un repositorio común

Actualmente, hay un repositorio común en *privacypatterns.org* [Doty11b] que alberga un total de 72 patrones de privacidad y que es resultado del trabajo colaborativo de un conjunto de investigadores en el área. Este repositorio inicialmente incluía 9 patrones, y fue concebido por Nicky Doty y Mohit Gupta en la Escuela de Información de la Universidad de California en Berkeley. Debido a que es una iniciativa abierta a otros autores, actualmente tiene colaboradores pertenecientes a 8 instituciones diferentes (Clever, Escuela Politécnica Nacional, Inria, Radboud University, UC Berkeley School of Information, Ulm University, Universidad Politécnica de Madrid, Vienna University of Economics and Business). El repositorio contiene la mayoría de los patrones que se pueden encontrar en las otras colecciones y se espera que se convierta en un lugar común de acceso a los patrones de privacidad. Este repositorio ha integrado muchos avances de cara a apoyar la investigación y uso de los patrones de privacidad. Estos avances comprenden: el establecimiento de una estructura común y plantillas para nuevos patrones, el uso de un repositorio en GitHub [Doty11a] para soportar nuevas contribuciones, el uso de estrategias y tácticas como mecanismos de organización.

El esfuerzo para integrar los diferentes catálogos de patrones ha puesto en evidencia otros problemas que impiden su evolución. Así, varios patrones no definen sus relaciones, y aquellos que si las definen tienen esquemas diferentes según el catálogo original al que pertenecen; asimismo, algunos esquemas usados solo llegan a definir la existencia de una relación sin brindar características adicionales. Por lo tanto, es menester el establecimiento de un esquema común para definir las relaciones entre los patrones de privacidad, de cara a poder soportar su evolución.

Métodos usando patrones de privacidad

Actualmente, varios métodos utilizan patrones como elementos clave dentro de sus procesos; por ejemplo, los métodos PRIS [Kalloniatis08], Pears [Kung14], [Hoepman14], PRIPARE [Notario15], entre otros. Solos o formando parte de otros elementos, los patrones de privacidad han sido recomendados en los reportes emitidos por ENISA [Danezis14] y por la AEPD [AEPD19] de cara a construir sistemas respetuosos con la privacidad y en consonancia con la legislación aplicable.

Esta apertura hacia los patrones puede deberse a que son usualmente vistos como elementos que están en un nivel operacional, y que por lo tanto le serían muy útiles a un desarrollador. Los patrones de privacidad permitirían definir soluciones de diseño muy cerca de la implementación, o incluso podrían llegar a apuntar hacia PETs útiles [Hoepman14]. Ejemplos de este tipo de enfoques son los propuestos por Hoepman [Hoepman14] y por Drozd [Drozd16a]. Ambos casos utilizan una perspectiva en niveles, en el que partiendo de un nivel de abstracción bastante alto (esto es, principios de privacidad), terminan recomendando un conjunto de patrones de privacidad (más operacionales).

El enfoque de Hoepman [Hoepman14] ha sido ampliamente difundido y aceptado en la comunidad. Además de ser recomendado por ENISA y por la AEPD, ha sido recientemente incluido en la normativa ISO/IEC sobre Ingeniería de privacidad para los procesos del ciclo de vida de los sistemas [ISO/IEC27550]. Este enfoque propone un esquema de tres capas para ayudar a los ingenieros a proveer soluciones de diseño que satisfagan los principios de privacidad altamente abstractos. La capa superior está compuesta por un conjunto de ocho *estrategias de diseño de privacidad* que permiten alcanzar cierto objetivo de protección acorde a los principios de privacidad [Hoepman14]. La capa intermedia ofrece 25 *tácticas* diferentes que actúan como medio para llegar a las estrategias [Hoepman18]. Finalmente, los patrones de privacidad, categorizados dentro de estrategias y tácticas, ayudarían a tomar decisiones de diseño en los sistemas a construir.

Este enfoque y los adoptados en otros métodos comparten un problema común: considerar los patrones de privacidad como cajas negras que están listas (o lo suficientemente maduras) para ser usadas, que son fáciles de aplicar, y que su aplicación tiene beneficios. Si bien es cierto que los patrones de diseño orientados a objetos [Gamma95] han tenido una amplia adopción y éxito, esto aún no ha sucedido en otros dominios. Además, se debe tener en cuenta que las características y beneficios de los patrones son muy dependientes del dominio en el cual son aplicados [Zhang12]. Estos temas no han sido explorados en el caso de los patrones de privacidad, y es necesario hacerlo, de cara a proveer sustento para fomentar su adopción.

Estudios empíricos con patrones de privacidad

La relevancia que los patrones de privacidad han ido adquiriendo en la ingeniería de privacidad hace necesario que se demuestre cuán fáciles de aplicar y útiles son. Lamentablemente, el estado del arte [Lenhard17][Caiza19][Morales-Trujillo19] en la investigación en patrones de privacidad para el diseño de sistemas no muestra estudios empíricos sobre la aplicación y los beneficios de los patrones de privacidad; en su lugar, las propuestas se apoyan principalmente en una buena argumentación.

Otros dominios más maduros como el de los patrones orientados a objetos [Chung04] [Prechelt02][Ng06][Mangalaraj14], y el de los patrones de seguridad [Abramov12] [Yskout12][Yskout15] sí reportan estudios empíricos en el contexto de la aplicación de los patrones y de los beneficios a obtener de la misma. Sin embargo, los únicos consensos a los que se han podido llegar es que los patrones bien documentados pueden mejorar la comunicación entre los desarrolladores y los encargados del mantenimiento del software, y que los beneficios a obtenerse dependen de los patrones analizados y del contexto en el cual se aplique [Zhang12][Riaz15]. Acorde a [Riaz15] deberían realizarse más estudios empíricos considerando particularmente el contexto de cara a establecer generalizaciones.

Finalmente, vale la pena tener en cuenta los estudios empíricos en el ámbito de diseño usando patrones de seguridad [Yskout12][Yskout15] debido a su cercanía a nuestro dominio. En [Yskout12] se enfocan en determinar si anotaciones extra (p.ej. objetivos de seguridad como integridad o confidencialidad) en los patrones producen impactos en la productividad del desarrollador. Mientras que en [Yskout15], se enfocan en determinar si el uso de los patrones impacta el tiempo y el producto de diseño. Aunque los resultados no sean generalizables, su experiencia es valiosa de cara a llevar a cabo estudios empíricos sobre los patrones de privacidad.

2.6. Resumen

Este capítulo ha presentado los antecedentes necesarios para entender los temas abordados en esta tesis. Fundamentalmente se divide en dos bloques: uno que explica el marco teórico de la privacidad y otro que aborda su proceso de ingeniería.

En el primer bloque se describe la privacidad en su concepción subjetiva, multi-dimensional, y contextual; para posteriormente enfocarse en las legislaciones que la abordan y que, de forma más pragmática, se centran en la protección de datos. Entonces, se presenta el RGPD que, siendo una de las legislaciones más avanzadas en el tema de la protección de datos, establece la protección de datos (y por lo tanto la privacidad) desde el diseño como uno de los mecanismos para proteger a las personas.

En el segundo bloque se aborda la ingeniería de privacidad que investiga diferentes medidas (teorías, métodos, técnicas y herramientas) que se pueden aplicar para incluir cuestiones de privacidad a lo largo del proceso de ingeniería de los sistemas. Se resalta que la actividad de diseño tiene una cantidad creciente de contribuciones que incrementan el cuerpo del conocimiento en el dominio, pero que estas están dispersas y desordenadas.

Los patrones de privacidad engloban muchas de estas contribuciones, e incluso han empezado a formar parte de reportes y recomendaciones por agencias de ciberseguridad europeas. Sin embargo, sus beneficios no han sido probados y su proceso de aplicación no ha sido estudiado.

Tanto la dispersión de los elementos propuestos para el diseño de sistemas respetuosos con la privacidad como el proceso de aplicación de los patrones de privacidad (elemento más reportado para el diseño) son abordados en los siguientes capítulos.

Capítulo 3

Mapa de elementos de diseño en la ingeniería de privacidad

Actualmente, hay una demanda urgente de herramientas, técnicas y métodos para que los ingenieros puedan implementar aspectos de privacidad a lo largo del proceso de desarrollo de los sistemas de información. La urgencia queda manifiesta por la necesidad de implementar medidas que protejan los datos personales para cumplir con la legislación vigente (RGPD [EuroPC16] en el caso europeo), y así, proteger la privacidad de las personas. La urgencia se acrecienta, ya que los ingenieros aún tienen problemas integrando las cuestiones de privacidad en las distintas tareas de desarrollo.

Los problemas que actualmente tienen los ingenieros incluyen la falta de conocimiento y entendimiento del concepto de privacidad, la dificultad de pasar de un alto nivel de abstracción de los principios de privacidad a decisiones técnicas de diseño, y el uso de su propia opinión y experiencia para definir los requisitos de privacidad que un sistema debería cumplir [Ayalon17][Bednar18][Hadar18][Senarath18]. Es decir, no hay indicios de que se estén usando elementos como herramientas, técnicas o métodos [Gürses16]. Sin embargo, si existe un interés a usarlos siempre y cuando sean útiles y se acoplen al esquema de trabajo al cual están acostumbrados [Senarath19].

Por otro lado, la cantidad de contribuciones en el ámbito de la ingeniería de privacidad ha venido incrementándose [Gürses16], y más aún aquellas dentro de la actividad de diseño de sistemas respetuosos con la privacidad [Lenhard17] [Morales-Trujillo19].

En este contexto (número creciente de propuestas aún no adoptadas por los ingenieros) se hace indispensable organizar el conocimiento existente para luego establecer soluciones que en realidad puedan ser llevadas a la práctica. Estudios con este objetivo se han realizado en la ingeniería de requisitos [Kalloniatis09][Anthonysamy17] [Beckers12]; sin embargo, no existe contribución similar para la disciplina de diseño.

Para solucionar esta necesidad, este capítulo presenta **un mapa de elementos reusables para el diseño sistemático de sistemas de información respetuosos con la privacidad**. Los elementos reusables incluyen los métodos y sus componentes que pueden utilizarse al realizar tareas de diseño. Se muestra también la madurez de estos elementos y las tendencias más relevantes. Profundizando más en el área, y al ser los patrones de privacidad los elementos más reportado, se analizan los temas investigados para estos elementos, su madurez, y se plantean las cuestiones de investigación aún no resueltas. Estos resultados son obtenidos a través de un mapeo sistemático de literatura, usando las guías provistas en [Petersen08] y actualizadas en [Petersen15].

La estructura de este capítulo se ha organizado de modo que en la Sección 1 se definen los elementos reusables de diseño considerados en este estudio. A continuación, la Sección 2 muestra la metodología seguida, concentrándose principalmente en la planificación y ejecución. Posteriormente, la Sección 3 presenta el mapa de elementos reusables encontrados, su madurez y sus tendencias. En la Sección 4 se hace un análisis más en profundidad de los patrones de privacidad al ser el tema más investigado en el ámbito de este estudio. Finalmente, se presenta el resumen del capítulo y las contribuciones originales realizadas.

3.1. Definición de los elementos reusables de diseño

Dado que el objetivo es identificar los elementos reusables para la actividad de diseño, es indispensable acotar primero dicha actividad. La actividad de diseño es entendida como el “proceso de definición de la arquitectura de software, componentes, módulos, interfaces y datos para que un sistema de software satisfaga los requisitos especificados” y “los resultados de este proceso” [ISO/IEC/IEEE24765].

Respecto a los elementos reusables de diseño, sus definiciones provienen del esquema de elementos provisto en el Metamodelo de Ingeniería de Software para Metodologías de Desarrollo (SEMDM –Software Engineering Metamodel for Development Methodologies) expuesto en el estándar ISO/IEC 24744 [ISO/IEC24744] y de la extensión para métodos de ingeniería de privacidad propuesta en [Martín17].

La Figura 3.1 muestra una organización de los elementos reusables de diseño, diferenciando aquellos que pertenecen al estándar SEMDM (sin fondo) y a la extensión para ingeniería de privacidad (sombreados). Además, están organizados dependiendo del aspecto que abordan: los productos a utilizar/generar, el proceso a seguir, o las personas o herramientas involucradas (productores). Algunos de ellos, como las restricciones y directrices, pertenecen a más de un aspecto. La Tabla 3.1 muestra las definiciones de los elementos reusables.

Aspectos de un método							
Elementos Reusables	Producto			Proceso	Productor		
	Producto						
	Producto de trabajo	Modelo	Documento				
	Und. Mod.	Unidad de modelo					
	Recurso	Modelo conceptual de privacidad					
		Marco normativo de privacidad					
		Código de ingeniería de privacidad					
		Base de conocimiento en privacidad					
		Lenguaje					
		Notación					
		Restricción					
	Etapa	Directriz					
		Ciclo					
		Fase					
		Construcción					
		Hito					
	Und. de trabajo	Actividad					
		Tarea					
		Técnica					
	Productor	Herramienta					
		Equipo					
		Rol de ingeniería de privacidad					
Método							

Figura 3.1: Elementos reusables de diseño

Tabla 3.1: Definiciones de los elementos reusables de diseño

Elemento reusable	Descripción
Método	Un método es un conjunto de elementos que especifican el proceso a seguir, los productos de trabajo a obtener o utilizar, y las personas o herramientas encargadas de hacer que el proceso se lleve a cabo en el contexto del desarrollo de un sistema. Por ejemplo, el método de programación extrema (XP – eXtreme Programming).
Producto de trabajo	Representa un artefacto de valor que puede ser creado, modificado, utilizado o destruido dentro de una o más unidades de trabajo.
Modelo	Representa una abstracción de una cosa, captando sus características (para algún propósito específico), según algún lenguaje abstracto. Por ejemplo, modelo de datos, modelo de objetos.
Documento	La representación de modelos u otras entidades a través de medios escritos o gráficos. Por ejemplo, un glosario, un diagrama.
Unidad de modelo	Un componente atómico de un modelo, que representa un fragmento cohesivo de información en la entidad siendo modelada. Las unidades del modelo no son productos de trabajo en sí mismos, sino una parte de ellos. Por ejemplo, una clase, un objeto, un atributo, una operación.
Recurso	Una representación abstracta de elementos reutilizables utilizados “tal cual”.
Modelo conceptual de privacidad	“Proporciona una descripción conceptual de lo que es la ‘privacidad’ en el contexto de la teoría de la privacidad, en la que se basa un método específico de ingeniería de la privacidad”. Por ejemplo, la ISO 29100 especifica la privacidad como un conjunto de principios.
Marco normativo de privacidad	En este “se establecen los requisitos normativos que deben aplicar todos los métodos que afirmen cumplirlo, y puede incluir reglamentos vinculantes, así como prácticas óptimas recomendadas no vinculantes”. Por ejemplo, el Reglamento general de protección de datos [EuroPC16].
Código de Ingeniería de Privacidad	“Clarifica la aplicación del marco normativo de privacidad en situaciones específicas”. El artículo 40 del RGPD elabora códigos de conducta destinados a “contribuir a la correcta aplicación del presente Reglamento, teniendo en cuenta las características específicas de los distintos sectores de transformación y las necesidades concretas de las microempresas y las pequeñas y medianas empresas” [EuroPC16]. Un ejemplo de estos códigos es el código de conducta en la Cloud Security Alliance (CSA) para el cumplimiento del RGPD [CCCSA].
Base de conocimiento en privacidad	“Es una pieza de conocimiento generalmente reconocida que puede ser reutilizada ‘tal cual’ en los esfuerzos de ingeniería de privacidad, y cuyo valor y utilidad son aceptados colectivamente”. Por ejemplo, los patrones de privacidad a medida que recogen y sistematizan los conocimientos de los expertos.
Lenguaje	Consiste en conceptos estructurados en un vocabulario combinado con un conjunto de reglas gramaticales para modelar desde una perspectiva determinada. P.ej. el Lenguaje de Modelado Unificado (Unified Modeling Language –UML).
Notación	Una convención para el uso de símbolos concretos para representar un modelo creado con un lenguaje. Por ejemplo, una clase UML está representada por un rectángulo con 3 secciones (nombre, atributos y operaciones).
Restricción	Una condición que debe cumplirse en un tiempo determinado. Por ejemplo, permitir el acceso a algún recurso después de que se cumpla alguna condición.
Directriz	Una indicación de cómo se puede utilizar un elemento del método durante su representación.

	Por ejemplo, una directriz que explique la función que un proceso debe desempeñar.
Etapas	Representa un marco de tiempo durante el cual se ejecutan un conjunto de unidades de trabajo cohesionadas temporalmente.
Ciclo	El período de tiempo en el cual el objetivo es entregar un producto o servicio final. Por ejemplo, un sprint en el método SCRUM.
Fase	El período de tiempo que conforma una partición lógica grande de un ciclo. Por ejemplo, la fase de diseño (desde una perspectiva de tiempo).
Construcción	Un período de tiempo durante el cual se completa una iteración del desarrollo de una nueva versión de un producto de trabajo. Por ejemplo, la primera iteración dentro de la fase de construcción de un sistema.
Hito	El instante en el que se logra un conjunto cohesivo de objetivos significativos (tareas, productos de trabajo). Por ejemplo, fase de diseño completada.
Unidad de trabajo	Representa un trabajo realizado como parte de un conjunto de procesos en un esfuerzo común.
Actividad	También conocido como un proceso en SEMDM. Es una unidad de trabajo de gran tamaño que modela una colección cohesiva de una o más tareas, operando en un área de especialización. Por ejemplo, la implementación de un módulo dentro de un sistema.
Tarea	Una unidad de trabajo de pequeño tamaño, que expresa una acción específica para lograr un propósito determinado. Por ejemplo, la selección de patrones de arquitectura o la documentación de un diseño.
Técnica	Una unidad de trabajo de pequeño tamaño que se centra en cómo se puede lograr un propósito determinado. Por ejemplo, categorización, encuesta.
Productor	Representa a alguien o algo que lleva a cabo unidades de trabajo y produce (o puede ayudar en la producción de) productos de trabajo.
Herramienta	Una aplicación informática que es utilizada por otros productores para automatizar la ejecución de sus responsabilidades. Por ejemplo, herramientas de documentación: procesador de textos.
Equipo	Consiste en un conjunto de productores que se centran en unidades de trabajo comunes. Por ejemplo, el equipo de proyecto, el equipo de arquitectura.
Rol de ingeniería de privacidad	Representa a alguien que entiende el marco de la privacidad y diseña sistemas favorables a la privacidad (considerando los elementos del método de privacidad) o participa dentro del método para diseñar uno. Por ejemplo, el ingeniero de privacidad.

3.2. Metodología

Un mapeo sistemático de literatura permite llegar a tener una visión general de un área de investigación en particular [Petersen08]. Esto permite tener evidencia cuantitativa para identificar tendencias, áreas dónde hay un mayor o menor número de contribuciones, para así fijar la atención en áreas específicas. A continuación, se describen las etapas de este tipo de estudios, para después, desarrollarlas en el contexto de este estudio.

Las guías establecidas en [Petersen08] y actualizadas en [Petersen15] han sido ampliamente aceptadas por los investigadores, y son las que se usan en este estudio. Acorde a estas guías, un mapeo sistemático de literatura tiene tres etapas (Figura 3.2): una para planificar la forma en la que se va a llevar a cabo el mapeo (planificación), una

en la que se lleva a cabo el plan (ejecución) y, por último, una para analizar los resultados, mostrarlos y concluir (reporte).

Figura 3.2: Etapas de un mapeo sistemático

En la etapa de planificación se llevan a cabo las tareas descritas en la Figura 3.3. Estas están orientadas principalmente a definir el ámbito del estudio, el proceso a seguir, y los materiales a usarse en la etapa de ejecución.

En la etapa de ejecución se materializan las decisiones tomadas en la planificación. Se inicia ejecutando la cadena de búsqueda, para luego, sobre los artículos obtenidos, llevar a cabo la inclusión y exclusión, y así obtener un conjunto de artículos para codificar y extraer datos. Esta codificación y extracción de datos se realiza utilizando los esquemas de clasificación definidos durante la planificación.

En la etapa de reporte, los datos obtenidos en la etapa anterior son analizados y representados apropiadamente. Finalmente, todo el procedimiento se describe en un reporte, incluyendo los principales elementos de todas las etapas: objetivo, preguntas de la investigación, los materiales, actividades desarrolladas, los resultados, las amenazas a la validez del estudio, el análisis y la discusión de los principales hallazgos del estudio. Esto se describe a continuación. El material completo utilizado se puede encontrar en el paquete para replicación (<http://dx.doi.org/10.17632/y5x695378r.1>).

Figura 3.3: Tareas en la etapa de planificación

3.2.1. Planificación

Definición del ámbito del estudio

El estudio ha sido delimitado por las siguientes preguntas de investigación (PI):

PI1. *¿Qué elementos reusables han sido propuestos para el diseño de sistemas basados en software respetuosos con la privacidad?*

El conocer los elementos reportados es el objetivo fundamental de este estudio.

PI2. *¿Cuál es el nivel de madurez de los estudios que han reportado estos elementos?*

La madurez de los estudios puede ayudar a determinar cuán cerca se está de una aplicación práctica. La madurez, acorde a los tipos de investigación propuestos en [Wieringa06], permite diferenciar las contribuciones desde aquellas que son únicamente opiniones hasta llegar a aquellas que han sido evaluadas empíricamente en la práctica.

Identificación del estudio

La estrategia que se usó en el desarrollo del estudio es la de búsqueda en una base de datos. La base de datos usada fue Scopus debido a su buen desempeño para identificar trabajos en el área de la computación [Cavacini15] y también, según consta en su sitio web [Elsevier20], por ser el repositorio más grande de resúmenes y referencias de trabajos revisados por pares. Scopus incluye referencias a bases de datos muy relevantes en computación: ACM (Association for Computing Machinery), IEEE (Institute of Electrical and Electronics Engineers), Springer, entre varias más.

Cadena de búsqueda

La cadena de búsqueda (Tabla 3.2) ha sido definida considerando tres secciones: la primera aborda la privacidad, la segunda a los sistemas de información basados en software y la tercera a los elementos reutilizables. Los términos de cada sección se han definido usando las siguientes fuentes, la mayoría de ellas estándares:

- los metamodelos de donde se obtienen los elementos reusables considerados en este estudio (ISO/IEC 24744 [ISO/IEC24744] y su extensión [Martín17]),
- consultas con investigadores experimentados en ingeniería de privacidad, y
- taxonomías y vocabularios (2017 IEEE Thesaurus Versión 1.0 [IEEE17], ACM Computing Classification System 2012 Revision [ACM12], y Systems and Software Engineering-Vocabulary ISO/IEC/IEEE 24765 [ISO/IEC/IEEE24765]).

Tabla 3.2: Cadena de búsqueda

("privacy" OR "data protection")		AND		("software" OR "system" OR "information" OR "computing" OR "computation" OR "computer" OR "informatics")		AND		("design" AND ("package" OR "component" OR "module" OR "framework" OR "design principle" OR "design stage" OR "design product" OR "methodology" OR "method" OR "language" OR "notation" OR "design constraint" OR "design guideline" OR "design cycle" OR "design phase" OR "design activity" OR "design task" OR "design technique" OR "workflow" OR "design artifact" OR "architecture" OR "model" OR "diagram" OR "design tool" OR "design rule" OR "strategy" OR "pattern" OR "design process" OR "heuristic" OR "goal" OR "standard" OR "convention" OR "design plan" OR "design policy" OR "design specification")	

La cadena de búsqueda ha sido desarrollada y evaluada en un proceso de dos iteraciones, en el que se agregaron (p. ej. “system”, “goal”) o se eliminaron (p. ej. “communication”) términos. En cada iteración, los resultados obtenidos al ejecutar la cadena en Scopus fueron comparados con un conjunto de artículos de referencia identificados por un investigador experimentado en ingeniería de la privacidad. La cadena de búsqueda final incluyó nueve de los 10 artículos definidos por el investigador.

Criterios de inclusión y exclusión

Para definir específicamente el dominio del estudio, para evitar el ruido en los resultados y para gestionar los recursos limitados como el tiempo, se han definido un conjunto de criterios para incluir o excluir artículos. Estos criterios fueron aplicados de manera automatizada (Tabla 3.3), usando la versatilidad de Scopus y a través de una hoja de cálculo, y de manera manual (Tabla 3.4), siendo leídos y evaluados por los investigadores trabajando en este estudio.

Los criterios aplicados con las herramientas proporcionadas por Scopus son auto-explicativos y ayudan a acotar el dominio, el intervalo temporal del estudio, y el tipo de recurso sobre el que se está trabajando. La hoja de cálculo ha servido para encontrar y eliminar duplicados, y para limitar la cantidad masiva de resultados. Esto se ha hecho en función de una valoración aproximada de la calidad de un artículo basado en la cantidad de citas recibidas. Para ello hemos definido un mínimo de citas: para los artículos publicados en el año que se realizó el estudio (2017) el mínimo es cero, para aquellos publicados en los seis años anteriores (entre el 2011 y 2016) el mínimo es aquel definido para estar en el 50% de artículos más citados acorde a los indicadores de Thomson Reuters [ThomsonR18] en la categoría de Ciencias de la Computación – Computer Science-, y para aquellos con más de seis años de antigüedad (anteriores a 2011) el límite fue cinco.

Tabla 3.3: Criterios de inclusión y exclusión automatizados

Filtros	Valores
Campo de investigación	Ciencia de la computación Ingeniería Ciencias de la decisión
Fecha de publicación	<= 2017
Tipo de documento	Artículo de congreso Artículo
Idioma	Inglés
Palabras clave	“privacy” o “data protection”
Citas mínimas	Número
	Año
	5
	Antes de 2011
	4
	2011 o 2012
	3
	2013 o 2014
	2
	2015
	1
	2016
	cualquiera
	2017

La aplicación manual de criterios de inclusión y exclusión fue llevada a cabo por 2 investigadores y revisada por un supervisor. Este proceso se realizó en iteraciones y los artículos se categorizaron como:

- Incluido, cuando el artículo cumple con todos los criterios de inclusión y ninguno de exclusión.
- Excluido, cuando el artículo no cumple con ningún criterio de inclusión o cumple con uno de exclusión.
- No claro, cuando el investigador tiene dudas respecto a incluirlo o excluirlo.

Para normalizar los criterios de los investigadores, se llevó a cabo una fase piloto previo a aplicar los criterios de manera masiva en todos los artículos resultado de la búsqueda. En este proceso de normalización, los investigadores leyeron los títulos y resúmenes de un mismo conjunto de artículos para categorizarlos adecuadamente. Las diferencias en la categorización fueron discutidas entre los investigadores y el supervisor al final de cada iteración. Para garantizar la homogeneidad entre los criterios de los investigadores, se utilizó el coeficiente ‘alfa de Krippendorff’ [Krippendorff09], que mide el nivel de acuerdo entre distintos codificadores, con un valor mínimo de 0,8 que es un valor aceptado en la mayoría de estudios [Lombard10]. Esto se alcanzó luego de la segunda iteración.

Tabla 3.4: Criterios de inclusión y exclusión manuales

Criterios	
Inclusión (todos deben cumplirse)	El artículo trata sobre privacidad o protección de datos.
	El artículo trata sobre sistemas de información basados en software respetuosos con la privacidad.
	El artículo es una contribución primaria.
	Hay por lo menos un elemento reutilizable (como contribución principal o secundaria en este documento).
	El elemento o elementos reusables propuestos pueden utilizarse en la etapa de diseño de los sistemas de información respetuosos con la privacidad.
Exclusión (ninguno se debe cumplir)	El artículo reporta un estudio secundario o terciario.
	El artículo reporta un metamodelo en el contexto de este estudio.

La fase principal de inclusión y exclusión se realizó en tres iteraciones. En la primera, cada investigador trabajó de manera independiente sobre grupos diferentes de artículos. Se leyó el título y el resumen previo a la categorización respectiva como incluido, excluido o no claro. Durante esta iteración se definieron cuatro puntos de control para verificar el coeficiente de acuerdo. En estos puntos de control, los investigadores y el supervisor categorizaron el mismo grupo de artículos y discutieron las divergencias hasta llegar a un acuerdo.

En la segunda iteración cada investigador revisó sus artículos etiquetados como ‘no claros’, esta vez, añadiendo la sección de conclusiones a la lectura. Los artículos fueron categorizados únicamente como incluidos o excluidos.

En la tercera y última iteración, los dos investigadores leyeron el título, resumen y conclusiones del mismo conjunto de artículos para tomar la decisión de incluir o excluir definitivamente. Esta decisión fue consensuada por los dos investigadores.

Clasificación y extracción de datos

Para poder analizar los artículos de cara a responder las preguntas de investigación, fue necesario categorizar (codificar) los artículos. Para ello, se obtuvieron los siguientes datos de los artículos finalmente incluidos:

- Los datos de identificación que incluyó un identificador del artículo dentro del estudio, su URL (Localizador de Recurso Uniforme –Uniform Resource Locator) en Scopus, e información de los autores.
- El contenido textual de las secciones a leer, esto es, el título, resumen, las palabras clave de autor.
- Datos útiles como el año, el tipo de documento (artículo de revista o conferencia), y el nombre de la revista o conferencia.

Además, también se guardó una copia de los artículos para los casos en los que fue necesario leer otras secciones adicionales a los datos ya extraídos.

Para categorizar los artículos, los investigadores leyeron el título, resumen, palabras clave de autor y las conclusiones; y posteriormente clasificaron considerando el tipo de elemento reusable de diseño (Tabla 3.1) y el nivel de madurez (usando la propuesta de [Wieringa 06] Tabla3.5). El conjunto de artículos incluidos, separado en lotes, fue leído por los dos investigadores. Después de leer cada lote, las codificaciones fueron discutidas y se estableció una codificación consensuada.

Tabla 3.5: Tipos de investigación usados como niveles de madurez

Tipo de investigación	Descripción
Investigación de evaluación	“Las técnicas se aplican en la práctica y se realiza una evaluación de la técnica. Esto significa que se muestra cómo se aplica la técnica en la práctica (aplicación de la solución) y cuáles son las consecuencias de la aplicación en términos de ventajas e inconvenientes (evaluación de la aplicación). Esto también incluye la identificación de problemas en la industria.”
Investigación de validación	“Las técnicas investigadas son novedosas o una actualización importante y aún no se han puesto en práctica. Las técnicas utilizadas son, por ejemplo, experimentos, es decir, trabajos realizados en el laboratorio.”
Propuesta de solución	“Se propone una solución para un problema, la solución puede ser novedosa o una extensión significativa de una técnica existente. Los posibles beneficios y la aplicabilidad de la solución se demuestran con un pequeño ejemplo o una buena línea de argumentación.”
Artículos filosóficos	“Estos artículos esbozan una nueva forma de ver las cosas existentes mediante la estructuración del campo en forma de una taxonomía o marco conceptual.”
Artículos de opinión	“Estos artículos expresan la opinión personal de alguien, si una cierta técnica es buena o mala, o cómo se deben hacer las cosas. No se fundamentan en trabajos relacionados y métodos de investigación.”
Artículos de experiencia	“Los artículos de experiencia explican qué y cómo se ha hecho algo en la práctica. Tiene que ser la experiencia personal del autor.”

3.2.2. Ejecución

En esta etapa se llevaron a cabo las decisiones tomadas en la etapa de planificación. La Figura 3.4 muestra las diferentes tareas realizadas y la cantidad de artículos que se han

obtenido al término de cada una de ellas. Estas tareas han sido llevadas a cabo por dos investigadores y vigiladas por un supervisor.

Figura 3.4: Ejecución del mapeo sistemático

La cadena de búsqueda fue ejecutada el 5 de febrero de 2018 para buscar en el título, resumen y palabras clave de los artículos existentes en la base de datos Scopus a esa fecha. Dado que la cadena de búsqueda es bastante inclusiva, se tuvo un conjunto de 7847 artículos devueltos.

Fue necesario entonces realizar un proceso de filtrado, el cual fue dividido en dos partes: automatizado y manual. La primera para aprovechar las ventajas de filtrado de Scopus y reducir así drásticamente el tiempo de procesamiento de los resultados. Aquí fueron aplicados los criterios definidos en la Tabla 3.3 y además se eliminaron duplicados. La segunda sirvió para identificar de manera más exacta los artículos que permitirían contestar a las preguntas de investigación. Aquí se aplicaron los criterios de la Tabla 3.4 a manera de árbol de decisión.

Dada la alta cantidad de artículos (2122) en el filtrado manual, se los dividió en lotes de 21 o 22, y se optó por que cada investigador trabaje en grupos diferentes de artículos. Para ello se llevó a cabo una fase piloto para unificar criterios. En la fase piloto se realizaron 2 iteraciones, al final de las cuales se procesaron 44 artículos en total y se superó el umbral definido en la etapa de planificación (Alfa de Krippendorff igual a 0,8). En la fase principal se procesaron los 2078 artículos restantes.

La fase principal del filtrado manual se llevó a cabo en tres iteraciones bajo las condiciones mencionadas en la planificación. Al término de la primera iteración, fundamentalmente con cada investigador codificando de manera individual, se obtuvieron 131 artículos incluidos, 117 no claros y 1874 excluidos. En la segunda iteración, luego que los autores revisaran sus propios artículos no claros, se obtuvieron 144 artículos incluidos y 1978 excluidos. Finalmente, en la tercera iteración, al revisar los dos investigadores el mismo conjunto de artículos y consensuar sobre los mismos, se obtuvieron 66 artículos (Anexos: Tabla A.1) que fueron posteriormente codificados para determinar los elementos reusables reportados y su madurez. Los resultados de la codificación se muestran y se analizan en la siguiente sección.

3.3. Mapa de elementos reusables de diseño

Los 66 artículos finalmente analizados en este estudio permiten construir el mapa de elementos reusables de diseño de sistemas respetuosos con la privacidad, el cual es analizado en los párrafos siguientes. Luego se analiza la madurez y tendencias de los elementos reusables como conjunto y de los elementos individuales más reportados.

Posteriormente, se realiza un análisis más en profundidad de los elementos de la *base de conocimiento en privacidad* y de los *métodos*.

El mapa de elementos reusables de diseño es presentado en la Figura 3.5, y contiene 73 apariciones de 12 tipos de elementos. Se puede notar que hay una concentración en tres tipos de elementos: la *base de conocimiento en privacidad*, los *métodos* y las *herramientas*, sumando entre los tres el 81% (59 de 73) de apariciones. Para mejorar la comprensión, los elementos han sido organizados acorde a su tipo y según el aspecto que abordan. Los tres aspectos denotan “el proceso a seguir, los productos a usar y generar, y la gente y herramientas involucradas (productores)” [ISO/IEC24744].

Figura 3.5: Mapa de elementos reusables de diseño

El aspecto de producto tiene claramente la mayoría de instancias (46), de las cuales 37 aportan a la *base de conocimiento en privacidad*, que serían las que más atención han captado por parte de los investigadores. En la Sección 3.3.3 se analiza más en detalle este tipo de elemento, que se desagrega en otros como los patrones de diseño (17 de los 37), las estrategias de diseño (3) y los modelos de referencia (17). El resto de elementos dentro del aspecto producto tienen pocas o ninguna contribución. En el proceso de filtrado varios elementos, como *modelos* y *documentos*, fueron descartados al ser muy específicos al sistema que se está diseñando y, por lo tanto, no reusables. Así también, hubo *modelos* que no eran específicos a un sistema, más bien, lo eran a un dominio, y además eran reutilizables, ajustándose más a la definición de un modelo de referencia dentro de la *base de conocimiento de privacidad*.

El segundo aspecto con más cantidad de apariciones encontradas es el de productor (7), muy distante de la cantidad en el aspecto producto. Las contribuciones se concentran únicamente en las *herramientas*, cuatro de los cuales proporcionan a los desarrolladores el apoyo para seleccionar patrones de privacidad. Las otras herramientas proporcionan información adicional a los *modelos* a través de anotaciones para ayudar al ingeniero en el proceso de diseño. El artículo 1935 permite anotar entidades y atributos, y el artículo 1023 anotar procesos. Por otro lado, el artículo 2266 propone un sistema para comprobar si una arquitectura cumple un conjunto de requisitos de privacidad.

El aspecto de proceso ha recibido escasas contribuciones reportadas de manera individual. Sin embargo, se ha notado que este tipo de elementos se han reportado normalmente dentro de métodos completos, como se ve en la Tabla 3.6.

Los *métodos* tienen especial relevancia debido a que son el segundo elemento más reportado (11 métodos en 15 artículos seleccionados) y ya que dentro de ellos hay más elementos reusables (mostrados en la Tabla 3.6). Se puede notar que los métodos reportados principalmente presentan elementos del aspecto de proceso, es decir, se enfocan en cómo llevar a cabo el diseño. Nueve de los 11 métodos describen elementos metodológicos pertenecientes al aspecto proceso, como *fases*, *actividades*, *tareas* y *técnicas* de diseño. Estos resultados permiten complementar los mostrados en la Figura 3.5, que muestran mayor concentración en los otros aspectos.

Aunque ocho de los 11 métodos presentan elementos dentro del aspecto de producto, los casos en los que se proponen *modelos*, estos solo son mencionados, pero no descritos. No obstante, los elementos de la *base de conocimiento en privacidad* vuelven a ser relevantes, ya que cuatro métodos incluyen por lo menos uno de ellos.

Finalmente, hay dos métodos en los que se han mencionado, sin ser descritos, *roles* genéricos que corresponden al aspecto de productor.

3.3.1. Madurez

Los 66 artículos estudiados fueron categorizados como investigación de evaluación, investigación de validación, propuestas de soluciones y artículos filosóficos (Figura 3.6), y no como artículos de opinión o de experiencia. Los artículos estudiados tienden en su mayoría (38%) a ser propuestas concretas a nivel de solución, luego están las varias propuestas en investigación de validación (30%), llegando a una escasa investigación a nivel de evaluación (9%). Los estudios a nivel de validación y evaluación pueden ser considerados como empíricos ya que proveen evidencia en un entorno de laboratorio y de producción respectivamente. En conjunto representan un porcentaje (39%) comparable al de las soluciones. Finalmente, las propuestas de nivel filosófico tienen un porcentaje (23%) razonable de contribuciones.

Al enfocarnos en los elementos reusables más reportados se notan algunas diferencias dependiendo del elemento. Mientras los elementos de la *base de conocimiento de privacidad (BCP)* tienen mayor porcentaje de contribuciones a nivel empírico (44%); los *métodos* llegan a un nivel de solución en su mayoría (46%). Las *herramientas*, aunque tienen una baja cantidad de contribuciones (7), también presentan más estudios empíricos (57%). Vale la pena notar que, para estos elementos más reportados, las contribuciones de nivel evaluación son escasas, e inclusive son nulas en los métodos. Respecto a las contribuciones a nivel filosófico, estas mantienen un porcentaje sostenido (~25%) en los diferentes elementos, a excepción de las herramientas.

Tabla 3.6: Elementos reusables dentro de los métodos

Año	Artículo	Metodología	Producto	Proceso	Productor
2005	12	Iachello and Abowd's		- Fase - Técnica	
2005-2008	384, 221, 577, 219	PRIS	- BCP (patrones de privacidad, modelo de referencia) - Técnica	- Actividad	
2009	162	Moncrieff et al.'s	- BCP (modelo de referencia)		
2011	182	Fhom and Bayarou's	- Modelo (mencionado)	- Fase - Actividad	
2011	100	Kost et al.'s	- Modelo (mencionado)	- Tarea	
2014	2198	PEARs	- BCP (estrategias, tácticas, patrones de privacidad, PETs)		
2014	2309	SALT	- Documento (especificación de diseño de un sistema)	- Tarea	
2014-2015	1726, 1825	PRIPARE	- BCP (modelo de referencia)	- Actividad	
2017	1882	Aljohani et al.'s		- Fase	- Rol (mencionado)
2017	1740	Feth et al.'s	- Modelo (mencionado) - Documento - Diagrama	- Fase - Actividad - Tarea	
2017	1751	SeeME		- Tarea	- Rol (mencionado)

3.3.2. Tendencias

El primer artículo considerado en este estudio es del 2002 y desde ahí el número de artículos reportando elementos reusables de diseño muestra una tendencia creciente (Figura 3.7), especialmente marcada desde 2014. De hecho, la mitad de los artículos han sido publicados desde ese año, y 24 de los 66 (36%) artículos corresponden a los dos últimos años analizados, 2016 y 2017.

Figura 3.6: Madurez de los elementos reusables de diseño

Figura 3.7: Cantidad de artículos por año

Esa concentración de artículos desde 2014 también es visible en todos los niveles de madurez (Figura 3.8 (a)), exceptuando las validaciones. En este período, se han publicado el 60% de artículos filosóficos (9 de 15), 56% de solución (14 de 25), 100% de evaluación (6 de 6). Sin embargo, solo las propuestas a nivel de solución muestran una tendencia creciente y sostenida.

Al dividir los tipos de contribuciones en empíricas (validaciones y evaluaciones) y no empíricas (soluciones y filosóficas), ambos grupos crecen conforme crece la cantidad de contribuciones en el tiempo. Sin embargo, se debe notar que las contribuciones empíricas siguen siendo menores a las no empíricas en instantes puntuales de tiempo.

Se puede notar en la Figura 3.8 que las tendencias en los elementos reusables de diseño se deben mayoritariamente a las que ocurren en las contribuciones a la *base de conocimiento en privacidad*. Por ende, este tipo de elemento será analizado más en detalle en la Sección 3.3.3.

Los *métodos* presentan un comportamiento diferente en torno a sus niveles de madurez. Por un lado, sus contribuciones empíricas solo llegan a nivel de validación y están concentradas previamente al 2009. Por otro lado, las contribuciones recientes, aunque considerables en cantidad (7 desde el 2014), se han quedado a nivel filosófico y de

solución. Finalmente, las *herramientas* se han mantenido con bajas contribuciones de diferente nivel de madurez a lo largo del tiempo.

Se nota entonces que los elementos *base de conocimiento de privacidad* y *método* son muy importantes de cara a analizar los elementos reusables y sus tendencias. El primero juega un rol importante al marcar la pauta en las tendencias de los elementos reusables y de su madurez. El segundo es un elemento compuesto, cuyos elementos internos (mayoritariamente orientados al proceso) complementan a los elementos reusables que han sido reportados de manera individual (mayoritariamente orientados al producto). Ambos tipos de elementos serán analizados más en detalle en las siguientes secciones.

Figura 3.8: Tendencias en madurez de elementos reusables de diseño

3.3.3. Base de conocimiento en privacidad

Las apariciones (37) de este tipo de elementos corresponden a *patrones de privacidad* (17), *estrategias de diseño* (3), y *modelos de referencia* (17). Los patrones de privacidad son elementos que no han sido estudiados de manera aislada, sino que han sido abordados de manera conjunta con estrategias de diseño, han sido considerados parte

fundamental en algunos métodos, o han sido el foco de algunas herramientas. Por lo que, se han convertido en un tema de investigación transversal a varios de los elementos reusables estudiados; por ello, son analizados en profundidad en la Sección 3.4.

Respecto a los *modelos de referencia*, estos permiten el establecimiento de marcos comunes y reutilizables para que los ingenieros que trabajan en un dominio específico se entiendan de una manera normalizada y clara. En línea con la definición provista en [Mackenzie06], un modelo de referencia describe el conjunto mínimo de características (incluidas, por ejemplo, las entidades y sus relaciones), comunes a todos los sistemas de un tipo determinado en un dominio específico, de manera neutral desde el punto de vista tecnológico e independiente de la aplicación.

La relevancia de los *modelos de referencia* queda plasmada en su tendencia de crecimiento y en la madurez que van alcanzando. Son elementos que han cobrado mayor relevancia en los últimos años, más de la mitad de estos (nueve de los 17) se han publicado entre 2015 y 2017 (Tabla3.7). En lo que respecta a su madurez, el 41% se ha comprobado empíricamente (en consonancia con las cifras globales de la *base de conocimiento en privacidad*); pero sólo dos contribuciones han alcanzado un nivel de evaluación, que se publicaron entre 2015 y 2017. Además de conocer estas tendencias, y acorde a su naturaleza, es importante analizar el dominio en el cuál se aplicarían (si lo hay), y el objeto del modelo de referencia (es decir, el artefacto de diseño).

Tabla 3.7: Investigación en modelos de referencia

Año	Contribución	Madurez	Dominio	Objeto
2002	138 (Arq. IBM)	Solución	Sistemas informáticos de empresa	Arquitectura
2004	372 (CONFAB)	Validación	Sistemas espaciotemporales	Arquitectura
2009	313	Filosófico	Sistemas espaciotemporales	Arquitectura
	725	Validación	Sistemas de procesamiento de información móvil	Arquitectura
2010	9	Validación	Plataformas de computación	Arquitectura
	958	Filosófico	Sistemas eléctricos	Arquitectura
2011	701	Filosófico	Sistemas eléctricos	Arquitectura
2013	2537	Validación	Sistema colaborativo y social	Arquitectura
2015	1837	Solución	Electrónica de consumo	Arquitectura
	2298	Evaluación	Sistemas informáticos de empresas	Arquitectura
	3014	Validación	Sanidad	Arquitectura
2016	1873	Filosófico	Sanidad	Arquitectura
	2709 (ENCASE)	Filosófico	Sistema colaborativo y social	Arquitectura
	3235	Solución	Sanidad	Seguridad
2017	1869	Solución	Sistemas espacio-temporales	Arquitectura
	2566	Solución	Sanidad	Datos
	2899	Evaluación	Infraestructuras inteligentes	Arquitectura

Los dominios que tienen más contribuciones (Tabla3.7) son la asistencia a la salud (4) y la de sistemas espacio-temporales (3) (es decir, servicios basados en la localización, redes de sensores, sistemas de información geográfica, etc.). Luego de estos, los

dominios de sistemas de información empresarial, sistemas de energía y los sistemas de computación colaborativa y social tienen 2 contribuciones cada uno. Hay que resaltar que hay modelos bastante completos y desarrollados (artículos 1837, 2298, 2899), e incluso incluyen técnicas que podrían utilizarse en la etapa de implementación del proceso de desarrollo de software (artículos 9, 372). Las contribuciones en cada dominio aparecen dispersas a lo largo del período estudiado. Únicamente las contribuciones relacionadas con la asistencia a la salud podrían representar un primer paso de una tendencia futura a tener más publicaciones, ya que todas ellas se concentran entre 2015 y 2017.

En el caso del objeto del modelo, hay una tendencia clara de los investigadores a proponer modelos arquitectónicos de referencia (15 de los 17). Esto resulta ser relevante considerando la importancia y urgencia que tiene este elemento en el proceso de diseño, la arquitectura de un sistema es la que se esboza al inicio, describe fundamentalmente a un sistema, y, además, juega un papel determinante en los requisitos no funcionales como la privacidad. Las otras dos contribuciones se centran en otros tipos de modelos: el artículo 2566 presenta un modelo de datos de referencia y el 3235 presenta un modelo de referencia de seguridad para manejar eficientemente los datos de los registros electrónicos de salud de los pacientes.

3.3.4. Métodos

Los métodos son el segundo elemento reusable más reportado en este estudio (15 apariciones reportando 11 métodos), y la mayoría de ellos (seis de los 11) se concentran entre 2014 y 2017, lo que demuestra la relevancia que ha adquirido este tema en los últimos años. Sin embargo, solo 2 métodos de los estudiados (PRIS y el reportado en el artículo 12) habrían llegado a un nivel de validación, ambos reportados en artículos previo a 2008. Aparentemente, una vez la investigación asociada finaliza, la evolución de los métodos también. A continuación, se analizan los métodos (Tabla 3.8) desde tres perspectivas: su enfoque de creación, esto es, si se definen desde cero o amplían métodos existentes; el dominio (si lo hay) de los sistemas a construir; y las etapas del ciclo de vida del desarrollo de software que abordan.

La mayoría de los métodos (7 de los 11) amplían procesos conocidos de desarrollo de sistemas para abordar cuestiones de privacidad. Este enfoque podría ayudar a superar los problemas reportados por los ingenieros sobre la adopción de prácticas de ingeniería de la privacidad [Hadar18][Senarath18][Senarath19], ya que, utilizar procesos de desarrollo conocidos modificados les resultaría más fácil que aprender un nuevo método desde cero. Sin embargo, según se observa, los esfuerzos en los últimos años por alcanzar nuevos niveles de madurez parecen limitados. Es decir, luego de sus propuestas como soluciones, no hay artículos que informen sobre estudios empíricos, o sobre mejoras incrementales.

Respecto al dominio que abordan, hay casi igual número de métodos que se centran en un dominio específico (5 de 11) y los que no (6 de 11), por lo que no hay una clara inclinación hacia uno de estos dominios; y tampoco se puede detectar ninguna tendencia. Podría ser que la definición de métodos enfocados en dominios específicos pueda facilitar la adopción de prácticas de privacidad, debido a la orientación concreta y detallada que podrían proporcionar. Esto cobra especial relevancia en dominios donde los contextos son muy particulares, por ejemplo, al manejar datos extremadamente sensibles en sistemas relacionados a la salud.

Aunque se esperaban métodos enfocados en la etapa de diseño, aparecieron también métodos que se ocupaban de las actividades de requisitos y de diseño, otros llegando

incluso a abarcar el desarrollo completo de un sistema. Hay una distribución aproximadamente equilibrada entre los tres enfoques a lo largo del tiempo.

Finalmente, varios de estos métodos reportan usar patrones de privacidad, lo cual parece ser algo común con otros elementos reusables encontrados.

Tabla 3.8: Investigación en métodos de diseño

Año	Método (Artículos)	Madurez	Enfoque de creación	Dominio	Etapas de desarrollo
2005	Iachello and Abowd's (12)	Validación	Ampliación (basada en modelos de procesos de diseño centrados en usuario)	Sistemas espacio-temporales	Diseño
2005 – 2008	PRIS (384, 221, 577, 219)	Validación	Ampliación (modifica un enfoque orientado al proceso desde la definición de los objetivos)	N/A (artículo no provee información)	Requisitos y diseño
2009	Moncrieff et al.'s (162)	Solución	N/A (artículo no provee info.)	Sistemas espacio-temporales, Sanidad	Diseño
2011	Fhom and Bayarou's (182)	Filosófico	Nueva metodología	Sistemas eléctricos	Todas
2011	Kost et al.'s (100)	Solución	Ampliación (amplía los métodos de diseño actuales)	N/A (artículo no provee info.)	Requisitos y diseño
2014	PEARs (2198)	Solución	Ampliación (complemento para una metodología de PbD)	N/A (artículo no provee info.)	Diseño
2014	SALT (2309)	Filosófico	Nueva metodología	Sistemas de información multimedia	Todas
2014 – 2015	PRIPARE (1726, 1825)	Solución	Ampliación (nuevas medidas para metodologías de desarrollo de sistemas bien adoptadas)	N/A (artículo no provee info.)	Todas
2017	Aljohani et al.'s (1882)	Filosófico	Ampliación (al concepto de participación del usuario final)	Sanidad	Requisitos y diseño
2017	Feth et al.'s (1740)	Solución	Nueva metodología	N/A (artículo no provee info.)	Diseño
2017	SeeME (1751)	Solución	Ampliación (amplía los métodos existentes de diseño socio-técnico)	N/A (artículo no provee info.)	Requisitos y diseño

3.4. Patrones de privacidad

Los patrones de privacidad son también elementos que integran la *base de conocimiento de privacidad*. Los resultados muestran que estos elementos tienen interacciones con otros elementos como las estrategias, los métodos y las herramientas. Por lo que, para poder captar esta transversalidad, lo abordamos como un área de investigación en la ingeniería de privacidad.

En el contexto de este estudio, los patrones de privacidad son el área de investigación que más atención ha captado por parte de los investigadores en la ingeniería de privacidad. Hay en total 32 apariciones (del total de 73 en el estudio) de elementos en el contexto de los patrones de privacidad reportados en 28 artículos, cuyo número ha aumentado constantemente desde 2014 (Tabla 3.9).

Esta tendencia a investigar sobre patrones de privacidad podría deberse al hecho de que encajan naturalmente en el contexto de la ingeniería de privacidad. Por un lado, organizan el conocimiento existente (generalmente de expertos) ofreciendo soluciones generales y sistemáticas para el diseño de sistemas respetuosos con la privacidad [Gürses16]. Por otro lado, los patrones fueron concebidos para ser utilizados por profesionales no expertos, lo que facilita su uso por ingenieros que no tienen un conocimiento adecuado sobre los aspectos de la privacidad y no saben cómo incorporarlos en sus diseños [Hadar18][Senarath18]. De cualquier modo, dada la atención puesta sobre los patrones de privacidad, es necesario realizar un análisis más en profundidad sobre la investigación en esta área.

Para ello, se han mapeado los temas de investigación en patrones de privacidad encontrados en los artículos estudiados a las categorías de investigación en patrones de diseño expuestas en [Bafandeh17] (Tabla 3.9). Por ejemplo, las propuestas de nuevos patrones corresponden a la categoría de creación de patrones. Asimismo, los métodos que incluyen a los patrones de privacidad dentro sus procesos, se pueden mapear a la categoría de utilización de patrones, y las herramientas encontradas a la categoría de búsqueda y selección de patrones. Al final, los artículos de este estudio se enmarcan en dos categorías principales: el desarrollo de patrones y el uso de patrones.

En la categoría de desarrollo de patrones, se ha encontrado artículos contribuyendo a la creación de patrones (14), de lenguajes (3), y a la categorización (3).

Las contribuciones se concentran mayoritariamente en la creación de patrones, sumando 17 artículos entre patrones y lenguajes. En este grupo de patrones se puede observar el inicio de la evolución de los mismos, en línea con lo indicado en los antecedentes de esta tesis (Capítulo 2). Es así que, aunque se tienen 14 artículos reportando la creación de nuevos patrones, también han aparecido contribuciones como un catálogo en línea (artículo 1807), una taxonomía de relaciones (artículo 2272) para conectar los patrones de privacidad, y un lenguaje de patrones (artículo 2031) para el anonimato en las tecnologías de soporte a la privacidad PETs. Soportando esta tendencia evolutiva también se debe tener en cuenta los demás catálogos descritos en la Sección 2.5.2.

Respecto a la categorización de patrones de privacidad, se ha encontrado que las estrategias de diseño reportadas en los artículos estudiados han sido utilizadas en la práctica como formas de categorización de patrones. Sin embargo, sus capacidades van más allá, ya que definen diferentes niveles de abstracción para facilitar el entendimiento y operacionalización de los requisitos de privacidad. El enfoque del artículo 1847 [Hoepman14] expone los fundamentos, y define un conjunto de estrategias, que facilitan la identificación de patrones para implementarlas. Este enfoque ha sido ampliamente aceptado en este dominio de investigación (41 citaciones en Scopus a 5 de febrero de 2018 y 110 a 23 de mayo de 2020). Además, ha sido utilizado como base para los artículos 1781 y 1728. El primero propone dos nuevas estrategias de diseño enfocadas en la rendición de cuentas (accountability), y el segundo expone un conjunto de 25 tácticas, que añaden un nuevo nivel de abstracción entre las estrategias propuestas y los patrones de privacidad.

Los artículos que tratan sobre el uso de patrones se dividen en dos subcategorías: la utilización de patrones (8) y la búsqueda y selección de patrones (4).

La categoría de utilización de patrones incluye, por el momento, únicamente métodos que utilizan patrones de privacidad. Los ocho artículos encontrados reportan cuatro métodos diferentes. Por lo general, estos métodos consideran los patrones de privacidad como elementos de abstracción de bajo nivel o más operacionales. El problema principal identificado es que los patrones son asumidos como elementos (cajas negras) listos para ser usados, fáciles de aplicar, y que su aplicación traerá beneficios. Para que el método pueda ser completamente validado o evaluado necesitará que los patrones de privacidad demuestren sus características y beneficios asumidos.

Tabla 3.9: Investigación en patrones de privacidad

Categorías de investigación en patrones de diseño	Desarrollo del patrón			Uso del patrón	
	Creación de patrones	Creación de un lenguaje de patrones	Categorización de patrones	Utilización del patrón	Búsqueda y selección de patrones
Elemento reusable	Base de conocimiento en privacidad				
Temas de investigación encontrados	Propuestas de patrones	Colecciones de patrones	Estrategias de diseño	Métodos	Herramientas
2005				PRIS: 384 (Sol.)	
2006	167 (Sol.) 746 (Sol.)				
2007	221 (Val.)			PRIS: 221 (Val.) 577 (Val.)	
2008	219 (Val.)			PRIS: 219 (Val.)	
2009					
2010					289 (Sol.)
2011					418 (Val.)
2012	383 (Sol.)				
2013	2731 (Sol.)	ID2031 (Sol.)			
2014	1798 (Fil.)		ID1847 (Fil.)	PRIPARE: 1726 (Fil.) PEAR: 2198 (Sol.)	1798 (Fil.)
2015				PRIPARE 1825 (Sol.)	
2016	1794 (Val.) 1926 (Sol.) 1996 (Val.) 2932 (Val.)	1807 (Sol.)	1728 (Fil.)		1807 (Sol.)
2017	1804 (Val.) 1808 (Eval.) 2126 (Eval.)	2272 (Val.)	1781 (Fil.)	Aljohani et al. 1882 (Phil.)	

Finalmente, en la subcategoría de búsqueda y selección de patrones se encuentran aquellas herramientas para encontrar los patrones más apropiados para problemas específicos. Los artículos encontrados se orientan a listar las ventajas de disponer de un repositorio en línea para la búsqueda de patrones de privacidad (artículo 1798), a describir un catálogo de patrones de privacidad en línea (artículo 1807), o a presentar herramientas para ayudar a los ingenieros en la selección de los patrones (artículos 289 y 418) proveyendo de recomendaciones en función del contexto.

Además de conocer el qué se está investigando, es necesario saber el grado de madurez del área de patrones de privacidad, de cara a usarlos en la práctica como parte de métodos, procesos, estrategias, etc.

3.4.1. Madurez

Hemos analizado la madurez desde dos enfoques, uno centrándonos en la madurez del área de investigación considerando las categorías de investigación cubiertas, y otro considerando el tipo de investigación reportado en cada uno de los artículos.

Considerando las categorías de investigación del área, se tiene que, aunque los patrones de privacidad hayan sido los elementos reusables más reportados en el marco de este estudio, su nivel de cobertura respecto a las categorías de investigación posibles en los patrones de diseño [Bafandeh17] es aún bajo. La Figura 3.9 resulta reveladora en tanto muestra que la cobertura de categorías es limitada: solo se han encontrado artículos dentro de 2 de las 5 categorías principales. Y dentro de estas, las únicas subcategorías que tienen un número considerable de artículos son las de los nuevos patrones y las de utilización de patrones. El resto de categorías con contribuciones, tienen una cantidad escasa de ellas. Hay categorías que requieren atención urgente; por ejemplo, la subcategoría de aplicación, ya que naturalmente luego de que los patrones son desarrollados, estos deberían ser aplicados por los ingenieros, siguiendo o no un método en particular. Asimismo, la categoría de evaluación de la calidad tiene especial relevancia de cara a garantizar la calidad de los patrones creados, y a evaluar las implicaciones (entre estas, los beneficios) de la aplicación de los mismos.

Figura 3.9: Madurez del área de investigación de los patrones de privacidad

Considerando específicamente el tipo de estudio reportado en cada artículo relacionado a los patrones de privacidad, menos de la mitad de los artículos (11 de 28) han sido probados empíricamente (Figura 3.10); y de ellos, la mayoría (9) solo han llegado a un nivel de validación. Considerando las categorías que más artículos tienen, la creación de patrones resulta tener más artículos empíricos (a nivel de validación o evaluación) (8 de 14), y los artículos proponiendo métodos solo reportan 3 validaciones (hay que tener en cuenta que los 3 artículos son sobre el método PRIS). La única categoría que reporta evaluaciones es la creación de patrones. En la categoría de creación de lenguajes, solo se reporta una validación en un artículo sobre relaciones (Tabla 3.9). Las estrategias de diseño permanecen a un nivel filosófico. En cuanto a las herramientas, se tiene solo una validación en el 2011.

Figura 3.10: Madurez de categorías de investigación en patrones de privacidad

3.4.2. Tendencias

Considerando la cantidad de artículos relacionados a los patrones de privacidad (Tabla 3.9), hay una concentración de artículos (17 de 28) desde el 2014 hasta el 2017, y se esperaría que haya más en los años posteriores. Además, se nota que los desarrollos de patrones de privacidad individuales son los que están tendiendo a proveer cierta evidencia empírica. Por el contrario, los artículos relacionados con *métodos*, luego del método PRIS en 2008, no han llegado a reportar estudios empíricos. La cantidad de contribuciones en el resto de categorías permanece baja, pero se han mantenido en los años recientes, por lo que podrían ser el origen de nuevas tendencias.

Teniendo en cuenta la perspectiva de categorías, la investigación dentro del área de patrones de privacidad, aunque ahora concentrada en el desarrollo de patrones, se está extendiendo al resto. Es así que actualmente hay varios catálogos, y los medios para definir relaciones, lo que hace pensar que pronto aparecerán más sistemas de patrones. La propuesta de nuevos lenguajes de patrones de privacidad podría llevar algo más de tiempo, debido a la complejidad asociada para abarcar el diseño completo de un sistema.

En cuanto a las estrategias de diseño propuestas por Hoepman (artículos 1847, 1728), aunque sea una de las contribuciones más aceptadas en el ámbito de la ingeniería de privacidad; en lo que compete a los patrones de privacidad, sólo se han utilizado como esquema de categorización (por ejemplo, en el catálogo [Doty11b]). En esta tesis consideramos que el potencial de las estrategias de diseño todavía requiere ser explotado, de cara a ofrecer directrices más concretas en procesos o métodos. De hecho, mientras se ha desarrollado esta tesis, su autor ha propuesto recientemente un conjunto de guías para implementar la privacidad desde el diseño [Hoepman18]. Aunque las definiciones de las estrategias están bastante maduras y las guías resultan ser

recomendaciones bastante útiles, todavía dependen mucho de la experiencia del ingeniero que las implementa.

En cuanto a los métodos, desde el 2014 han aparecido nuevas propuestas. Sin embargo, la publicación de artículos relacionados con un método particular suele estar ligada a la duración del proyecto de investigación asociado. Sería interesante el evaluar más en detalle los métodos propuestos, tal vez a través de estudios sistemáticos de literatura más profundos, de cara a definir sus ventajas, a mejorarlos, y a tratar de llevarlos a la práctica. Finalmente, no encontramos tendencias en las herramientas, y su cantidad sigue siendo baja a pesar de su valor para ayudar a los ingenieros en el proceso de desarrollo de los sistemas.

En general, se nota que en todas las categorías hay un interés en seguir proponiendo contribuciones en torno a los patrones de privacidad. Son destacables los esfuerzos que se están haciendo para evolucionar los patrones de privacidad hacia posibles sistemas o lenguajes y hacia su uso dentro de métodos o conjuntamente con las estrategias de diseño. Sin embargo, de cara a demostrar su validez en la práctica, es urgente investigar la aplicación de los mismos, dado que aún hay cuestiones relevantes sin resolver:

- ¿Cuán fácil es aplicar los patrones de privacidad existentes?
- ¿Cuán efectiva es la aplicación de patrones de privacidad en el diseño de sistemas respetuosos con la privacidad?
- ¿Qué beneficios aportan los patrones de privacidad a los ingenieros?
- Si no se es posible aún el aplicarlos, ¿qué problemas están obstaculizando su aplicación? ¿Cómo podrían resolverse esos problemas?

Contestar estas preguntas implica directamente la realización de estudios de carácter empírico, bien sea a nivel de validaciones o de evaluaciones.

3.5. Resumen del capítulo y contribuciones originales

Este capítulo ha identificado los elementos reusables reportados en el estado del arte para el diseño sistemático de sistemas de software respetuosos con la privacidad, junto con su nivel de madurez y las tendencias en este dominio de investigación. Para ello, se han presentado las definiciones de los distintos elementos y se ha descrito la metodología seguida: un mapeo sistemático de la literatura, detallando su planificación y ejecución. Finalmente, se han presentado los resultados obtenidos y su análisis respectivo.

El análisis de los resultados permitió elaborar las siguientes contribuciones originales:

- **Un mapa de elementos reusables para el diseño de sistemas de información respetuosos con la privacidad.** En él se destacan los elementos reusables más reportados, su madurez y tendencias.
- El estado actual de la investigación en el área de los patrones de privacidad, siendo el área más investigada dentro de los elementos reusables de diseño en este dominio. También se establecen los pasos próximos en los cuales debería enfocarse la investigación en el área de los patrones de privacidad.

El área de investigación en patrones de privacidad ha cobrado mucha relevancia en la ingeniería de privacidad. Muestra de ello es el creciente número de contribuciones, no solo en el desarrollo de nuevos patrones, sino también en las propuestas de estrategias de diseño, métodos y herramientas que usan patrones de privacidad. Sin embargo, el área aún es incipiente y siguiendo su crecimiento natural, requiere de contribuciones en

el resto de categorías (p.ej. la aplicación y la evaluación) acompañadas de la respectiva validación empírica. Este tipo de contribuciones podrían ayudar a determinar las características y beneficios usualmente asumidos de los patrones de privacidad; los cuales, de ser probados, podrían fomentar su adopción y la de los elementos desarrollados en base a ellos. Este tema es explorado en el siguiente capítulo.

Los resultados de este estudio se han detallado en el artículo “Reusable Elements for the Systematic Design of Privacy-Friendly Information Systems: A Mapping Study” [Caiza19] que ha sido publicado en la revista “IEEE Access” (JCR³: Q1).

Este estudio se ha llevado a cabo en colaboración con investigadores del proyecto “PDP4E - Métodos y herramientas para el cumplimiento del RGPD a través de la Ingeniería en Protección de Datos y Privacidad”⁴ dentro del programa de investigación e innovación Horizonte 2020 de la Unión Europea.

Además, se ha llevado a cabo un evento de difusión con el título “Necesito hacer mi tesis, ¿cómo hago mi estudio del estado del arte?”⁵ directamente relacionado a la metodología aplicada en este estudio. Esta conferencia versó sobre el proceso a seguir para llevar a cabo estudios de literatura tanto en estudios de pregrado como de posgrado.

---

³ Factor de impacto Web of Science Journal Citation Reports <https://jcr.clarivate.com/>

⁴ PDP4E <https://cordis.europa.eu/project/id/787034/es>

⁵ Conferencia en enero de 2020 <https://detri.epn.edu.ec/necesito-hacer-mi-tesis-como-hago-mi-estudio-del-estado-del-arte/>

Capítulo 4

Estudio exploratorio de la aplicación de los patrones de privacidad

Como se vio en el Capítulo 3, los patrones de privacidad son un área de investigación que ha captado mucha atención por parte de los investigadores en ingeniería de privacidad. Por su naturaleza, se ajustan al contexto actual de la ingeniería de privacidad donde se requieren elementos reusables que sistematicen la inclusión de cuestiones de privacidad en el proceso de desarrollo de sistemas, y que brinden un soporte adecuado a ingenieros no expertos en privacidad.

La investigación en el área de los patrones de privacidad se ha enfocado principalmente en proponer nuevos patrones, en hacer esfuerzos hacia la construcción de lenguajes y en proponer métodos de ingeniería de privacidad que los usen (Sección 3.4.1). A pesar de que esta área está en una etapa temprana de madurez, varios de los enfoques que usan patrones de privacidad ya constan en guías de ingeniería de privacidad provistas por agencias europeas de ciberseguridad (ENISA [Danezis14], AEPD [AEPD19]).

Sin embargo, en camino a investigar el proceso de aplicación de los patrones de privacidad, solo ha habido propuestas [Pearson10][Pearson11][Drozdz16a] enfocadas en la búsqueda y selección de patrones. Peor aún, no hay contribuciones empíricas orientadas a evaluar la facilidad de aplicación de los patrones o sus beneficios, quedando únicamente como asunciones extrapoladas de otros dominios como el de la orientación a objetos [Gamma95], donde han tenido un notable éxito. Avanzar en la aplicación de patrones de privacidad está alineado al flujo natural de investigación en dominios más maduros [Bafandeh17]. Más aún, la investigación en estos temas es necesaria ya que ayudaría a demostrar la utilidad de los patrones de privacidad, dando soporte a los elementos basados en ella y fomentando su adopción.

Específicamente, preguntas como ¿están los patrones de privacidad listos para ser usados?, ¿son fáciles de aplicar? y ¿proveen beneficios? han quedado planteadas y deberían ser respondidas de cara a avanzar en la investigación en patrones de privacidad y dar pruebas de su utilidad.

Para empezar a proveer respuestas a estas preguntas, en este capítulo se presenta un estudio empírico exploratorio sobre la aplicación de los patrones de privacidad. Esto se lleva a cabo a través de un experimento que sigue las guías de experimentación en ingeniería del software propuestas en [Wohlin12]. La realización del experimento ha permitido realizar un **análisis en profundidad sobre los instrumentos y oportunidades de experimentación en la aplicación de los patrones de privacidad**.

La estructura de este capítulo se ha organizado de modo que en la Sección 1 se describe el diseño y operación del experimento. Posteriormente, los resultados son mostrados en la Sección 2. A continuación, la Sección 3 elabora un análisis detallado sobre las posibilidades de experimentación actuales en torno a la aplicación de los patrones de privacidad. En la Sección 4 se hace una reflexión sobre las afectaciones del proceso de aplicación de los patrones de privacidad. Finalmente, en la Sección 5 se presenta el resumen del capítulo y las contribuciones realizadas.

4.1. Metodología

De cara a evaluar y mejorar los elementos como procesos, métodos, o herramientas usados en los procesos de ingeniería de software, es necesario llevar a cabo estudios empíricos [Sjøberg05]. Dentro de los diferentes tipos de estudios empíricos, la experimentación se orienta a confirmar teorías, concepciones, a explorar relaciones, entre otras [Hannay07]; por lo que resulta ser el tipo apropiado de cara a demostrar las concepciones que se tienen sobre los patrones de privacidad.

El proceso de experimentación, según [Wohlin12], tiene 5 etapas (Figura 4.1). La primera etapa corresponde a la definición del ámbito del estudio, en la que se define de manera clara el objetivo del experimento. A continuación, en la planificación se determina el contexto del experimento en detalle; así se incluye la definición de variables, de los participantes, del diseño del experimento, de la instrumentación, del objeto, de los procesos de medición y de la validez de los resultados. A continuación, en la operación se prepara y se lleva cabo el experimento. Posteriormente, se analizan e interpretan los datos; para finalmente, presentar los resultados y preparar un paquete de replicación que contiene todo el material necesario para reproducir el experimento.

Figura 4.1: Proceso de experimentación en ingeniería de software

4.1.1. Ámbito

Para construir evidencia empírica respecto a la aplicación y beneficios de los patrones de privacidad, este estudio tiene por objetivo, evaluar su efectividad para resolver problemas de diseño por desarrolladores no expertos en privacidad. Específicamente, se evalúa si al aplicar patrones de privacidad se mejora la calidad de la solución a un problema de diseño. Además, se analizan las percepciones que los desarrolladores tienen al aplicar los patrones de privacidad.

4.1.2. Planificación

Contexto

Este estudio se lleva a cabo en un área de investigación emergente y en crecimiento, en la que no hay estudios empíricos o experiencia previa específicamente en torno a la aplicación de los patrones para resolver tareas de diseño. Por lo que, en línea con las recomendaciones provistas en [Yskout18], se ha decidido orientarlo como estudio exploratorio. En la misma línea, dado que se quiere dar los primeros pasos en la investigación en la aplicación de patrones de privacidad, el uso de estudiantes es justificado. Acorde a [Carver10] se puede trabajar con estudiantes en estudios empíricos cuando se busca probar nuevas ideas o dar perspectivas iniciales.

Debido al carácter exploratorio del experimento y al uso de estudiantes de pregrado, se definieron problemas de diseño autónomos donde se trató de mantener la complejidad en un nivel medio.

Finalmente, se usa específicamente el grupo de patrones propuesto en [Hafiz13]. Este grupo de patrones es uno de los dos únicos lenguajes de patrones [Schümmer04][Hafiz13] de privacidad, siendo además el único que ha sido evaluado por pares expertos en patrones. Lo cual asegura que este grupo de patrones tiene criterios de homogeneidad y de calidad. Usar otro de los catálogos disponibles e incluso

repositorios que recolectan patrones de varias fuentes presentaría varios problemas de heterogeneidad.

Preguntas de investigación e hipótesis

En línea con el objetivo planteado en el ámbito del estudio, se definieron las siguientes preguntas de investigación (PIX):

- PIX1: ¿Cuál es el efecto en las soluciones de diseño al usar patrones de privacidad?
Esta pregunta guarda relación directa con la determinación de los beneficios asumidos al aplicar patrones de privacidad. Se evalúa la completitud –cualidad de estar completa– de las soluciones de diseño usando una calificación basada en escala provista por un experto.
- PIX2: ¿Cómo son percibidos los patrones de privacidad por los desarrolladores?
En esta pregunta se determina específicamente si los patrones son percibidos como fáciles de comprender y de aplicar. Para el primer caso, se usa la percepción en términos de facilidad de lectura y entendimiento de los patrones de privacidad. Para el segundo caso se pregunta por la facilidad percibida al aplicar los patrones de privacidad.

De cara a responder la pregunta de investigación PIX1, se ha definido la hipótesis inicial:

- H1o: La completitud de la solución a un problema de diseño con requisitos de privacidad es la misma al usar o no patrones de privacidad.

La pregunta de investigación PIX2 se responde describiendo las percepciones de los desarrolladores.

Variables

La Tabla 4.1 muestra las variables utilizadas en el experimento. Se utilizan variables cualitativas tanto para evaluar la completitud de la solución al problema de diseño como para recoger la percepción de los participantes.

De cara a contestar la PIX1 a través de la H1o se ha identificado una variable independiente y una dependiente. La variable independiente es el uso de patrones de privacidad, que define los tratamientos aplicados: resolver un problema de diseño sin o con el uso de patrones de privacidad. La variable dependiente es la completitud de la solución, su valor es definido por un experto en privacidad (diferente al investigador que ejecuta el estudio) que evalúa la solución de los participantes.

Las variables definidas para contestar la PIX2 no se consideran dependientes o independientes dentro del experimento, pero permiten recoger la percepción de los participantes después de utilizar los patrones de privacidad.

Respecto a otras variables que podrían interferir en los resultados del experimento, estas han sido tratadas de ser controladas a lo largo del estudio. Así, se lleva a cabo una etapa de capacitación para nivelar el conocimiento previo en privacidad, en diseño de software y en diseño con requisitos de privacidad. Además, se controla el nivel de dificultad de los problemas de diseño usados.

Participantes

Se previó que los participantes del estudio serían estudiantes de pregrado de carreras asociadas a la tecnología o al desarrollo de software de dos universidades ecuatorianas (Escuela Politécnica Nacional y Universidad de las Américas).

Los requisitos para participar en el experimento eran tener al menos conocimiento básico sobre diseño de software y medio en el ámbito general del desarrollo de software. Asimismo, era un requisito deseable estar en los últimos semestres de la carrera. La participación en el experimento fue voluntaria y no influía en modo alguno en las asignaturas cursadas por los estudiantes.

Tabla 4.1: Variables utilizadas en el experimento

Variable	Rol	Naturaleza		Escala
- Uso de patrón de privacidad	Independiente	Cualitativa	Nominal	Sí No
- Completitud de la solución	Dependiente	Cualitativa	Ordinal	Buena: se cumple con los requisitos de privacidad Aceptable: se cumple medianamente con el requisito Deficiente: cumplimiento escaso del requisito Ninguna: no hay criterios o no ayudan a cumplir el requisito
- Facilidad de lectura percibida	N/A	Cualitativa	Ordinal	Muy de acuerdo
- Facilidad de comprensión percibida				De acuerdo
- Facilidad de aplicación percibida				En desacuerdo Muy en desacuerdo

Tipo de diseño

El tipo de diseño utilizado es el de un factor con dos tratamientos (tratamiento 1: sin usar patrones en el proceso de resolución de un problema de diseño, tratamiento 2: usando patrones en el proceso de resolución de un problema de diseño), uno de los tipos estándares según [Wohlin12]. Dado que se usan cuatro problemas de diseño de privacidad; dos están destinados a ser resueltos sin patrones de privacidad, antes que los participantes reciban una capacitación sobre patrones de privacidad; y tras recibir la capacitación se resuelven dos problemas de diseño más.

Además, la asignación de los problemas a los participantes se realiza de manera aleatoria para limitar amenazas a la validez y para equilibrar la distribución de los problemas.

Instrumentación

Los instrumentos utilizados incluyen aquellos para la resolución de los problemas de diseño propuestos y el material didáctico para la capacitación de los participantes. Dentro del primer tipo de instrumentos encontramos los problemas con requisitos de privacidad, el conjunto de patrones de privacidad a usar y los instrumentos de medición de las diferentes variables definidas. Dentro del segundo grupo se engloban diapositivas sobre la privacidad, los datos personales, el diseño de sistemas, el diseño respetuoso con la privacidad, los patrones de privacidad y ejemplos de resolución de problemas con

requisitos de privacidad. Los instrumentos son descritos a continuación y pueden ser accedidos en https://github.com/juliocaiza/exploratory_experiment_privacy_patterns.

Problemas de diseño con requisitos de privacidad

Los problemas de diseño fueron definidos de tal manera que sean autónomos y dentro de un contexto de un sistema de monitorización de la salud. Además, se controló la complejidad de los problemas según el tiempo disponible y al ser la primera vez que los participantes resolvían problemas con requerimientos de privacidad.

Como enfoque para resolver los problemas (Figura 4.2), se usó aquel basado en estrategias propuesto en [Hoepman14]. En este enfoque se debe identificar la estrategia a seguir, las tácticas más específicas y luego se establecen las decisiones de diseño para el sistema a construir [Hoepman18]. Este proceso sin modificaciones corresponde al tratamiento uno (sin usar patrones de privacidad en el proceso de resolución). Para el tratamiento dos (usando patrones de privacidad en el proceso de resolución) luego de la definición de las tácticas, se selecciona y aplica los patrones adecuados.

Cada problema incluyó un contexto, una sección de requisitos funcionales y una sección con requisitos de privacidad (Figura 4.3). Además, considerando que todos los patrones pertenecen a un solo conjunto orientado a proveer anonimato en los sistemas, se pre-estableció una única estrategia, la de *Ocultar*, para todos los problemas. Así, los participantes debían identificar las tácticas a utilizar, los patrones de privacidad (únicamente en el tratamiento apropiado) y finalmente debían establecer las decisiones de diseño a tomar. Las decisiones de diseño fueron evaluadas en última instancia.

Figura 4.2: Proceso de resolución de problema con requisitos de privacidad

Patrones de privacidad utilizados

El grupo de patrones de privacidad utilizado incluyó los 12 patrones de privacidad reportados en el lenguaje de patrones provisto en [Hafiz13]. Este lenguaje tiene como objetivo ayudar a tomar decisiones de diseño para proporcionar anonimato en las comunicaciones. Cabe destacar que este lenguaje y cada uno de sus patrones han sido evaluados por pares expertos en patrones. Este hecho garantiza la calidad de la descripción desde la perspectiva de los autores de los patrones.

Cada patrón se describe usando una plantilla que tiene las secciones: nombre, resumen, contexto, problema, solución, ejemplos, consecuencias, usos conocidos y relaciones (la Figura 4.4 muestra un ejemplo tomado de [Hafiz13]). Además, cada patrón de privacidad fue categorizado de acuerdo a una táctica determinada dentro de la estrategia *Ocultar*. Para evitar problemas de comprensión debido al idioma se decidió traducir todos los patrones, originalmente en inglés, al español que es el idioma nativo de los participantes.

Contexto	Se está pensando en implementar un servicio de monitoreo de salud remoto a pacientes (uno o más miembros de la familia) desde sus hogares. El servicio de monitoreo requiere recolectar y procesar un conjunto de datos de signos vitales para su correcto funcionamiento; sin ellos, no sería viable el servicio. El servicio, además de ofrecer la visualización de los datos, también ofrece alertas ante posibles escenarios que requieran de atención médica urgente. Para la implementación del servicio se usarán los siguientes dispositivos: una banda a usarse en la muñeca, un dispositivo HUB a ubicarse en el hogar, y un servidor centralizado para recibir las mediciones. La banda tiene distintos sensores, limitada capacidad de procesamiento y WIFI como único medio de comunicación; el HUB tiene capacidades suficientes para procesar datos, almacenarlos, y para recibir y transmitir; y el servidor guarda las mediciones para que puedan ser analizadas por médicos en caso de que se requiera atención médica. Se usará Internet para transportar los datos.
Requisitos funcionales	Establecer los mecanismos adecuados de comunicación de la banda al HUB, y de éste al servidor.
Requisitos de privacidad	Proteger los datos sensibles en los dos enlaces de comunicación para evitar que puedan ser leídos y conocer a quién le pertenece.
Estrategia a usarse	Ocultar (Hide) [Predefinido]
Tácticas a usar	Restringir el acceso a los datos Ofuscar los datos de identificación Disociar los datos y la identificación del usuario
Patrones	Identidad seudónima (Pseudonymous identity) Representación cambiada (Morphed representation)
Decisiones de diseño	- Usar un mecanismo de pseudonimización de identidad del usuario para los datos a enviarse de la banda al HUB. - La limitada capacidad de procesamiento de la banda no le podría permitir usar un mecanismo de cifrado. Además, la recolección y transmisión de datos debe ser rápida. - Este mecanismo ofusca los datos de identificación y los disocia de los datos. - Usar un mecanismo de cifrado de datos para la comunicación entre el HUB y el servidor. - Esto restringirá accesos no deseados mientras los datos atraviesan Internet.

Figura 4.3: Ejemplo de problema de diseño con requisitos de privacidad

Nombre	IDENTIDAD SEUDÓNIMA
Estrategia	Ocultar
Táctica	Ofuscar
Resumen	IDENTIDAD SEUDÓNIMA oculta los objetivos a anonimizar bajo un seudónimo. Es similar a usar un alias en una organización encubierta; solo unos pocos elegidos conocen la identidad real, un extraño no.
Contexto	Los objetos a anonimizar pueden ser personales o impersonales, pero todo tiene alguna forma de identidad. Por ejemplo, la identidad de una persona es un nombre, la identidad de un paquete es la información en su encabezado, la identidad de una entrada de datos en una base de datos es su clave primaria, etc.
Problema	Si las identidades están expuestas, las entidades también lo estarán. ¿Cómo se pueden proteger las entidades de una exposición?
Solución	Ocultar la identidad adoptando un seudónimo aleatorio que no se relaciona con el original. Almacena la asignación del seudónimo porque también se necesitará la asignación inversa. Protege la asignación de seudónimo, porque el éxito de un seudónimo depende de mantenerlo en secreto. Hay dos problemas críticos para implementar este patrón: encontrar objetivos a anonimizar y seudónimos apropiados y la proteger el mapeo del seudónimo. Los objetivos a anonimizar y los seudónimos varían basados en el dominio del problema tratado. La garantía de privacidad de este patrón depende en gran medida del segundo problema: proteger el mapeo del seudónimo...
Ejemplos	Para las interacciones web anónimas, un encabezado de paquete es un objetivo a anonimizar porque contiene metadata acerca del emisor de los paquetes. Se puede usar un proxy entre el emisor de la solicitud y un destinatario para quitar la información del encabezado (ej. HTTP USER AGENT) y enviarlo en nombre del emisor. Este esquema se usa en herramientas de anonimización conocidas como Anonymizer e iProxy. ...
Consecuencia	N/A
Usos conocidos	La parte que almacena el mapeo puede convertirse en un único punto de falla. Además, quien mantiene un proxy de anonimización o un remailer puede ser forzado por las autoridades legales a divulgar esta información, quebrantando el anonimato de los usuarios. Esta es la razón por la cual el remailer Penet fue cerrado. En 1995, la Iglesia de la Cienciología obligó al dueño de Penet, Johan Helsingius a revelar la identidad de un usuario con el seudónimo "-AB-" y la identificación anónima an144108@anon.penet.fi al ganar una orden judicial. Luego, en 1996, ...
Relaciones	Los seudónimos deben cambiarse dentro de una mix network siguiendo una REPRESENTACIÓN CAMBIADA.

Figura 4.4: Ejemplo de patrón de privacidad: Identidad seudónima

Herramientas de medición

En cuanto a las herramientas de medición, se utiliza una herramienta de soporte al experimento y encuestas en formularios de Google. La herramienta [Guzman20] (Figura 4.5) ha sido utilizada para mostrar los problemas de privacidad y para recolectar las soluciones enviadas por los participantes. La herramienta fue construida específicamente para dar soporte a este experimento, pero ha sido diseñada para apoyar la experimentación con cualquier tipo de patrones de diseño. Esta sigue mejorándose para futuras oportunidades de experimentación. El código fuente y los manuales de la herramienta están incluidos en el repositorio de datos y materiales.

Las encuestas se han utilizado para recolectar la percepción de los participantes de cara a responder la PIX2 y para obtener percepciones en cuanto a instrumentos que posiblemente influyen en el experimento (por ejemplo, la influencia de la herramienta en el proceso de resolución del problema de diseño). Se utilizaron cuatro encuestas dentro del estudio: la primera para evaluar los conocimientos previos; la segunda para evaluar las impresiones de los participantes sobre la capacitación recibida; y la tercera y cuarta para evaluar las percepciones sobre la resolución de los problemas de diseño sin y con el uso de patrones de privacidad.

The image displays two screenshots of a web-based experimental interface. The top screenshot, titled "New template", shows a configuration screen for a design template. It includes fields for "Name" (Template patrones con clasificacion) and "Description" (Template para patrones de diseño cuyas secciones incluyen dos clasificaciones). Below these are two tables: "Available sections" and "Selected sections".

N	Name	Data Type
4	Resumen	Text
5	Contexto	Text
8	Ejemplo (fig)	File

N	Name	Data type	Mandatory	Main
1	Nombre	Text	✓	✓
2	Estrategia	Classification	✓	✓
3	Tactica	Classification	✓	✓
4	Problema	Text	✓	
5	Solucion	Text	✓	

The bottom screenshot, titled "Experiment: Escenario 2 con patrones", shows the problem scenario presented to participants. It includes instructions in Spanish about solving a problem related to a health monitoring system. The problem is titled "Problem 1 of 1: Problema 1". The text describes a system for monitoring patient health in a hospital, where a HUB device is used to collect and process vital signs data. The system is designed to be secure, with data encrypted and access controlled. The participant's task is to solve the problem by indicating the strategy, tactic, and design criteria to be used.

Figura 4.5: Ejemplo de interfaces del experimentador (superior) y los participantes (inferior)

Amenazas a la validez

Respecto a las amenazas a la validez interna, se han utilizado diferentes medidas para reducirlas. Así, la asignación al azar de los problemas a los participantes evita favorecer alguno de los tratamientos aplicados y reduce los efectos del conocimiento previo de los participantes. El uso de un conjunto de patrones de privacidad con calidad confirmada permite controlar los posibles efectos sobre el proceso que ocasionarían patrones deficientes. Para reducir el sesgo en la elaboración de los instrumentos, especialmente en los problemas de privacidad y los formularios, se utilizó una doble validación con un experto en privacidad. Se decidió que la participación fuera voluntaria de cara a evitar un sobrecargo en el trabajo que cada participante tiene en sus actividades cotidianas.

A pesar de estas medidas preventivas, tras el estudio, hubo reportes de posibles afectaciones. Por ejemplo, aunque casi todos los participantes estuvieron de acuerdo o muy de acuerdo en que se familiarizaron con los temas examinados en la capacitación, algunos de ellos dijeron que todavía necesitaban más ejemplos de referencia. Asimismo, resulta que algunos participantes tuvieron dificultades para comprender las instrucciones sobre cómo describir sus soluciones, limitándose en ocasiones a únicamente enviar los patrones usados. La participación voluntaria resulta ser un arma de doble filo, ya que por un lado es un medio para evitar sobrecargar a los participantes con actividades impuestas adicionales, pero por otro ha actuado como amenaza debido al compromiso que los participantes tuvieron en el estudio (p.ej. 4 abandonaron el proceso). Se pueden explorar otras alternativas para aumentar el compromiso. Finalmente, vale la pena destacar la buena percepción respecto al uso de la herramienta; todos los participantes estuvieron de acuerdo en que era fácil de usar y que no los perturbaba en el proceso de diseño.

En cuanto a la validez externa, la principal amenaza es que los resultados de este estudio son aplicables al conjunto de los 12 patrones utilizados y a los atributos medidos, lo cual consta en la literatura sobre experimentación en patrones de diseño [Riaz15]. Los resultados podrían ser susceptibles de generalización si hubiera más patrones pertenecientes a la misma colección de patrones o a un grupo con características similares. Asimismo, la cantidad de participantes resulta ser muy reducida para generalizaciones. Contar con más participantes y con un mayor número de patrones de privacidad normalizados (o similares) y de calidad ayudarían a tener resultados más generalizables.

En cuanto a la validez del constructo, las percepciones de los participantes y la evaluación de la completitud son trianguladas de cara a saber si lo percibido finalmente genera un beneficio.

4.1.3. Operación

Preparación

En esta etapa se llevaron a cabo las actividades para preparar todo lo necesario previo a la ejecución del experimento. El experimento fue elaborado en el marco de un curso de capacitación de 10 horas de duración titulado “Diseño de sistemas con requisitos de privacidad” orientado a alumnos que cumplieran con los requisitos previamente descritos en la Sección 4.1.2 de planificación.

Tras el proceso de inscripción y confirmación de los participantes, se obtuvieron 15 inscripciones, que incluyeron estudiantes y profesores que también mostraron interés. De los 15 participantes, 11 atravesaron todas las etapas del estudio, ocho estudiantes y tres profesores. Los ocho estudiantes estaban en los últimos semestres de dos

titulaciones: Tecnologías de la Información e Ingeniería en Sistemas. Los profesores forman parte del personal de la carrera de Ingeniería en Sistemas.

Los participantes declararon tener un conocimiento medio o bueno en diseño de software. Sin embargo, declararon conocimientos medios y bajos en privacidad, y bajos o nulos en diseño de privacidad (Tabla 4.2). Estos resultados no fueron sorprendentes, ya que sus carreras incluyen temas de diseño de software, pero no incluyen aspectos de privacidad.

Tabla 4.2: Conocimiento de los participantes antes del curso

Escala	Privacidad (Estudiante Docente)	Diseño de software (Estudiante Docente)	Diseño de privacidad (Estudiante Docente)
Excelente	--	--	--
Bueno	--	5 2	--
Intermedio	4 1	2 1	1 1
Bajo	3 2	1 0	6 0
Ningún	1 0	--	1 2

Tras obtener los datos de los estudiantes, fue posible realizar la asignación aleatoria de los problemas, cumpliendo así el tipo de diseño definido. La Tabla 4.3 muestra la asignación aleatoria de los problemas a los 11 participantes.

Tabla 4.3: Número de participantes asignados a los problemas de diseño

	Problema 1	Problema 2	Problema 3	Problema 4	Total
Sin patrones	5	7	5	5	22
Patrones	6	4	6	6	22
Total	11	11	11	11	

Esta etapa incluyó también la creación de todo el material de capacitación y para la resolución de los problemas de diseño con requisitos de privacidad, descritos en la planificación. Respecto a la herramienta utilizada, en esta etapa se realizó la instalación y pruebas de la versión utilizada en este experimento.

Ejecución

El estudio se llevó a cabo a largo de cuatro etapas (Figura 4.6), similar al experimento con patrones de seguridad reportado en [Yskout15], dos dedicadas a capacitar a los participantes en cuestiones de diseño, privacidad y patrones de privacidad, y dos más para que aplicaran sus habilidades para resolver problemas de diseño con requisitos de privacidad. Las etapas de capacitación incluían la teoría respectiva y ejemplos sobre la solución de problemas de diseño con requisitos de privacidad. La primera etapa se llevó a cabo en cuatro horas mientras que las demás duraron dos horas cada una.

Las etapas número dos y cuatro son aquellas en las que se resuelven los problemas, dos de ellos sin usar patrones de privacidad, y luego dos más usando patrones. Como paso previo a finalizar cada una de estas etapas, se recolectaron las soluciones a evaluar. Tras haber enviado las soluciones en cada etapa, los participantes rellenaron una encuesta sobre su experiencia en la resolución de los problemas de diseño sin usar y usando patrones de privacidad. Adicionalmente, también se cumplimentaron dos encuestas sobre la capacitación, una antes de la capacitación para determinar los conocimientos previos, y otra después para evaluar la capacitación.

Figura 4.6: Etapas dentro del estudio

4.2. Resultados

Los resultados obtenidos no permiten llegar a determinar de manera concluyente las características (facilidad de lectura, entendimiento y aplicación) y beneficios (completitud) que se querían probar. Esto se debe a que, aunque el conjunto de patrones de privacidad evaluado fue percibido positivamente, los resultados sobre el beneficio en el diseño no demostraron que haya ventajas al usar patrones de privacidad. A continuación, se analizan los resultados y se responde a las preguntas de investigación.

4.2.1. Pregunta de investigación PIX1: ¿Cuál es el efecto en las soluciones de diseño al usar patrones de privacidad?

Esta pregunta se orientó a determinar si realmente el uso de patrones de privacidad provee beneficios sobre el diseño realizado. En este caso, se evaluaron las decisiones de diseño de cara a determinar su grado de completitud para solventar los requisitos de privacidad planteados en los problemas. La evaluación se basó en una escala de cuatro niveles (Tabla 4.1) y fue llevada a cabo por un experto en privacidad diferente al experimentador.

La Tabla 4.4 muestra la frecuencia de cada nivel de evaluación por problema trabajando sin o con patrones. Considerando que el número de muestras difiere entre cada problema, e incluso al usar o no patrones dentro de ellos, únicamente el problema 2 y el problema 4 muestran cierta mejora al usar patrones. Asumiendo que la complejidad planteada en los problemas fuese la misma, únicamente el número de soluciones buenas aumenta.

También se realizó una prueba U de Mann-Whitney de dos colas con $\alpha = 0,10$, que es utilizada para probar variables dependientes ordinales, con tamaños de muestra pequeños y diferentes [Marin03]. La tabla 4.5 muestra los resultados, y como cada valor U es mayor que el U crítico, no hay evidencia para rechazar H_0 : *la completitud de la solución a un problema de diseño con requisitos de privacidad es la misma al usar o no patrones de privacidad*. Esto quiere decir que, en el contexto preliminar de este estudio, no hay una diferencia significativa entre usar o no patrones de privacidad.

Tabla 4.4: Frecuencias de evaluaciones a los problemas de diseño

	Problema 1		Problema 2		Problema 3		Problema 4			
Evaluación	Sin	Con	Sin	Con	Sin	Con	Sin	Con	Sin	Con
Buena	2	2	2	1	0	0	1	4	5	7
Aceptable	1	0	2	3	3	0	2	1	8	4
Deficiente	0	2	1	0	0	1	1	0	2	3
Ninguna	2	2	2	0	2	5	1	1	7	8
Total	5	6	7	4	5	6	5	6	22	22

Tabla 4.5: Valores para la prueba U de Mann-Whitney para cada problema

	Problema 1	Problema 2	Problema 3	Problema 4
U	14	10	7	8
U crítica	5	4	5	3
Resultado	No rechazar H1o			

4.2.2. Pregunta de investigación PIX2: ¿Cómo son percibidos los patrones de privacidad por los desarrolladores?

Esta pregunta de investigación tenía como objetivo evaluar la percepción sobre la facilidad de lectura de las descripciones, de entendimiento (asociada a una posible complejidad) y de aplicación de los patrones de privacidad. La Figura 4.7 muestra los resultados obtenidos.

La mayoría de los participantes piensa que las descripciones son fáciles de leer (claras) (10 de los 11) y fáciles de aplicar/usar (9 de los 11). Sin embargo, hay cierto desacuerdo sobre la complejidad de los patrones de privacidad, ya que 3 de los 11 consideran que no son fáciles de entender.

Los resultados en torno a la facilidad de lectura están en consonancia con lo que se esperaba ya que todos los patrones utilizados han sido previamente revisados por pares en un evento PLOP. Asociado a la dificultad en el entendimiento, algunos de ellos reportaron tener dificultades con la terminología, tuvieron la impresión que algunos patrones se parecían y quisieron conocer de manera más clara las ventajas y desventajas. Ante esto, sugirieron más entrenamiento con problemas de referencia. Por último, la facilidad de aplicación/uso se percibe bien, puede ser que esto se deba a únicamente la petición de decisiones de diseño; por lo tanto, necesitaban leer los patrones para obtener ideas y no para aplicarlas estrictamente, lo que podría haber sido bastante más complejo.

Figura 4.7: Percepción de los participantes sobre los patrones de privacidad

4.3. Posibilidades de experimentación con patrones de privacidad

Más allá de los resultados no concluyentes, este estudio exploratorio ha permitido realizar reflexiones profundas sobre el experimento de cara a proponer estudios más exhaustivos a posteriori, en línea con lo dicho en [Yskout18]. Para empezar, se ha notado que los instrumentos, y sus capacidades actuales, han afectado el alcance de este estudio. Por ejemplo, el tipo de solución provista en los patrones de privacidad usados afecta directamente el alcance del problema de diseño planteado, en tanto que el tipo de solución pedida debe estar en consonancia con el tipo de la solución que proveen los patrones de privacidad. El problema radica en que aún no existe un consenso en el tipo de solución que un patrón de privacidad debería proveer.

La afectación de los instrumentos al alcance del estudio se incrementa al requerir otros que provienen de la misma área de investigación, que se sabe es incipiente (Sección 3.4.1). Por ejemplo, el experimento requiere un proceso de diseño de sistemas usando patrones de privacidad (los existentes aún requieren estudios empíricos) y de un mecanismo para evaluar la aplicación de los patrones de privacidad (aún no hay propuestas formales en el área).

De este modo, estamos en una situación en la cual los objetivos de los estudios empíricos se ven limitados por el estado actual de los instrumentos que provienen de la misma área de investigación. Sin embargo, sí se pueden llevar a cabo estudios de este estilo, indispensables de cara a probar las características y beneficios de los patrones de privacidad en su estado actual y cuando sean mejorados.

A continuación, nos enfocamos en determinar:

- el estado actual de los instrumentos necesarios para llevar a cabo experimentos en este ámbito;
- los enfoques de calidad (el efecto primario que se estudia en el experimento, p.ej. completitud –calidad de estar completo) que se pueden estudiar, ahora y una vez superadas las limitaciones de los instrumentos; y
- la validez del uso de estudiantes como participantes en experimentos en el área.

4.3.1. Estado actual de los instrumentos para experimentación

Un primer grupo de instrumentos (Figura 4.8) gira en torno al proceso de resolución de problemas de diseño con requisitos de privacidad. Su importancia radica en que son una pieza fundamental del estudio y deben ser considerados tanto para la capacitación (para ejemplificar el proceso a seguir), como para el experimento (para su uso definitivo en la resolución de los problemas). Estos instrumentos incluyen los problemas de diseño con requisitos de privacidad (en nuestro estudio, el objeto a analizar), el proceso para el diseño con requisitos de privacidad (en nuestro estudio, donde se incluyen los tratamientos) y un conjunto de patrones (en nuestro estudio, los elementos usados en uno de los tratamientos). Además de estos, también identificamos dos elementos implícitos muy importantes: un problema de referencia, en función del cual se podrían definir problemas de diseño, y el tipo de solución esperada en los problemas (en nuestro estudio, el tipo del resultado al aplicar los tratamientos), que está restringida por el tipo de solución que tienen los patrones de privacidad utilizados.

El resto de instrumentos incluye los mecanismos de evaluación para poder comprobar lo que se requiera en el experimento, dependiendo del enfoque de calidad adoptado y las herramientas para recolectar los valores de las mediciones, o para soportar todo el proceso de experimentación.

La definición y preparación de todos estos instrumentos no es una tarea trivial, de hecho, es complicada debido a las relaciones y dependencias que existen entre ellos, a la ausencia de algunos instrumentos en el estado del arte, así como a las limitaciones de madurez que presentan aquellos existentes.

Instrumentos faltantes en el estado del arte

Aún hacen falta *problemas (o casos) de referencia* que se aproximen a un escenario real con consideraciones de privacidad. Estos podrían ser reutilizados para definir problemas para la capacitación y para el experimento, contribuyendo así a reducir el tiempo de planificación y preparación de los experimentos. Además, un elemento de este estilo favorecería la reducción de posibles sesgos del experimentador al permitir la creación

de los problemas de diseño de manera sistemática. Este tipo de elementos está disponible en otros ámbitos, como los patrones de seguridad. Por ejemplo, la Fundación Nacional de Ciencias (National Science Foundation - NSF) de los Estados Unidos ha llevado a cabo un proyecto en esta línea [Yuan10]. Desafortunadamente, hasta donde sabemos, no hay problemas de referencia disponibles en el dominio del diseño con requisitos de privacidad.

Figura 4.8: Relaciones y dependencias entre los instrumentos

También se requieren *mecanismos o marcos de evaluación* para determinar el impacto de usar patrones de privacidad en la calidad de los productos de diseño y en el proceso de diseño mismo. El estado del arte en la investigación en patrones de privacidad [Lenhard17][Caiza19][Morales-Trujillo19] aún no reporta propuestas de este tipo; sin embargo, el estado del arte en el dominio general de los patrones de diseño [Bafandeh17] ha empezado a explorar el área con contribuciones que principalmente usan métricas basadas en código para los estudios cuantitativos y evaluaciones de expertos para los cualitativos. Un marco de evaluación para el ámbito de experimentación en patrones de privacidad debería usar, en principio, métricas aplicables a evaluar procesos y productos de diseño más que a evaluaciones a nivel de código. Un instrumento de este estilo podría proporcionar un enfoque sistemático para evaluar cualitativa y cuantitativamente las diferentes características y beneficios y contribuiría en gran medida a reducir la subjetividad en las mediciones y posteriores evaluaciones. Más adelante, se aborda más en detalle el tema de las métricas que podrían usarse en experimentos.

Instrumentos existentes en el estado del arte

Aunque se han desarrollado y reportado muchos *patrones de privacidad*, tienen aún varias limitaciones. Características como el dominio de la aplicación, el tipo de solución proporcionado dentro de los patrones y la cantidad de patrones de privacidad con calidad probada son fundamentales para la definición de los otros instrumentos, e incluso pueden llegar a limitarlos. Por ejemplo, el lenguaje de patrones de privacidad usado en este estudio [Hafiz13] tiene pocos patrones (12) y se enfocan en un dominio muy específico (anonimato en comunicaciones), lo que limitó el espacio de maniobra al

definir los problemas de privacidad. Asimismo, las soluciones provistas en los patrones se orientan a brindar recomendaciones que podrían ser tomadas en cuenta para tomar decisiones de diseño ante un problema planteado. Esto limita la definición del problema a usarse en el experimento, ya que el tipo de la solución esperada debe guardar consonancia con el tipo de la solución dentro del patrón. Por ello, en los problemas propuestos se solicitaron únicamente decisiones de diseño.

Es necesario tener en cuenta que este lenguaje de patrones fue para nosotros la mejor opción para el experimento, debido a que es el único conjunto de patrones que de manera colectiva e individual ha sido revisado por pares en un evento PLOP especializado en patrones. Esto resulta fundamental de cara a encontrar patrones con un buen nivel de calidad.

Aunque consideramos el uso de otros catálogos o repositorios de patrones de privacidad más grandes, por ejemplo, los presentados en [Drozd16b] [Doty11b], estos tenían mayores limitaciones. Una de las más importantes es la falta de garantías de calidad, ya que casi ninguno de los catálogos ha pasado por un proceso de revisión individual de sus patrones por parte de los expertos en patrones. Además, dado que los repositorios recolectan los patrones encontrados en la literatura, muchas veces han hecho resúmenes de las fuentes originales, por lo que se ha perdido información. Finalmente, el recolectar patrones de diferentes fuentes puede llevar a problemas de heterogeneidad entre las descripciones de los patrones.

Por lo tanto, se necesitan más conjuntos de patrones de privacidad validados de manera individual y colectiva. Esto contribuiría a ampliar los dominios en los cuales se pueden plantear problemas de privacidad. En la misma línea, y como enfoque complementario a los eventos PLOP, consideramos apropiado el proponer mecanismos para evaluar la calidad de los patrones, incluso antes de experimentar, y con más razón antes de su publicación en repositorios abiertos. Estudios posteriores a estos mecanismos, deberían estar orientados a evaluar la calidad de los patrones de privacidad, en particular del repositorio común que se está conformando en [Doty11b]. Finalmente, también consideramos importante que los patrones de privacidad proporcionen soluciones representadas a través de medios a los que los diseñadores están más acostumbrados (p. ej. modelos y diagramas). Esto permitiría realizar experimentos que vayan más allá de evaluaciones cualitativas basadas en escalas, eliminando más el grado de subjetividad del evaluador.

Para orientar a los participantes en el proceso de resolución de los problemas se requiere un *proceso de diseño con requisitos de privacidad* establecido, que defina los pasos a seguir para evitar comportamientos arbitrarios y posibles efectos en los resultados. De este modo, es posible enfocarnos en el efecto de la aplicación de los patrones de privacidad. La limitación que tienen la mayoría de propuestas de este estilo es que no han sido validados o evaluados.

En nuestro caso particular, usamos el enfoque basado en estrategias propuesto en [Hoepman14], ampliamente aceptado por la comunidad de investigación. En su actualización más reciente [Hoepman18] las estrategias y tácticas se describen a manera de instrucciones de alto nivel, acompañadas de ejemplos de sistemas en diferentes dominios brevemente descritos. En este enfoque no se tiene un conjunto exhaustivo y detallado de pasos que sean parte de un proceso sistemático; sino que, el desarrollador debería identificar las estrategias y tácticas más adecuadas en función de sus descripciones y ejemplos. Es por ello que la experiencia adquirida del desarrollador bien en su vida práctica o como parte de una capacitación será fundamental. Acorde a

[Colesky16] el enfoque aún debe ser validado. Como primer paso hacia esto, hicimos una pregunta a los participantes sobre su percepción al utilizar este enfoque, y todos ellos estuvieron de acuerdo en que las estrategias y tácticas son útiles para guiarlos. Sin embargo, consideramos que todavía faltan experimentos enfocados en evaluar el uso de estrategias y tácticas durante el proceso de diseño con requisitos de privacidad.

4.3.2. Enfoques de calidad

Usualmente los patrones de privacidad han sido asumidos como elementos que tienen ciertas características y que producen beneficios por el mero hecho de ser definidos por sus autores como ‘*patrones*’. Aunque no ha habido estudios previos (en este dominio) en torno a comprobar lo dicho, sí se han llevado a cabo estudios experimentales en el dominio general de los patrones de diseño. Estos se han enfocado en demostrar únicamente los beneficios que produce la aplicación de dichos patrones [Riaz15]. La Tabla 4.6 además de indicar los beneficios que se han tratado de demostrar, los clasifica considerando sobre qué o quién se produce el beneficio, y se incluyen las características que usualmente se han sobreentendido sobre los patrones.

Tabla 4.6: Características y beneficios atribuidos a los patrones de diseño

	tienen características como	y producen beneficios	por ejemplo,
Los patrones de diseño	facilidad lectura, facilidad de comprensión y facilidad de uso/aplicación;	a las personas (desarrolladores),	mejorando sus habilidades de diseño, promoviendo mejores prácticas y generando mejor comunicación (entre los desarrolladores actuales y con los que hagan mantenimiento),
		al proceso,	mejorando la productividad y reduciendo los costes de desarrollo y de refactorización y
		al producto;	mejorando la calidad.

Los estudios sobre experimentación con patrones de diseño reportan que los beneficios dependen en gran medida del conjunto de patrones que se evalúa y del contexto de la evaluación [Riaz15]. Por ello, hay que guardar cautela al generalizar los resultados de un conjunto de patrones a otros y aún más a otros dominios.

Dada la situación actual de los patrones de privacidad, los experimentadores deben hacer las siguientes consideraciones.

- Respecto a las características de los patrones de privacidad. Aunque todos los patrones deberían tenerlas, como usualmente se asume, no hay garantía de ello al no tener todos los patrones una calidad probada bien en una revisión por pares o a través de un marco de evaluación. Incluso, cuando existe la revisión por pares, aún se requiere una evaluación desde la perspectiva del desarrollador quien es su usuario final y no precisamente un experto en el desarrollo de patrones. Las características pueden evaluarse directamente mediante medidas cualitativas basadas en escalas. También pueden medirse indirectamente a través de los beneficios producidos en las personas, los procesos o los productos, es decir, si el conjunto de patrones demuestra beneficios, puede suponerse que tienen las características garantizadas. Este enfoque es el que se observa en la mayoría de estudios empíricos con patrones de diseño.

- En cuanto a los beneficios en los productos de diseño. Estos pueden ser evaluados cualitativamente por ingenieros experimentados en cuestiones de privacidad mediante, por ejemplo, evaluaciones basadas en escalas para evaluar las decisiones de diseño. De tener tipos de soluciones y representaciones más objetivas como los modelos, el tipo de mediciones también lo serían (p.ej. evaluación de una arquitectura). Las mediciones cuantitativas aún son mayormente un reto; por un lado, porque necesitan de requerimientos bien definidos en problemas más reales; por otro lado, porque todavía se carece de marcos de evaluación objetiva y sistemática de la privacidad [Wagner18] [Guaman20].
- Acerca de los beneficios en el proceso de diseño. Estos podrían ser determinados mediante mediciones cuantitativas y cualitativas. Las mediciones cuantitativas en su mayoría suelen estar relacionadas con el tiempo (p.ej. tiempo para resolver una tarea), y es necesario diferenciar el tiempo de lectura del problema y los patrones a usar del tiempo de resolución del problema. En problemas con un alcance reducido, el tiempo de lectura puede afectar considerablemente las mediciones. En cuanto a las mediciones cualitativas, se suelen valorar el proceso de selección de patrones o el de aplicación de los mismos. En ambos casos, para evaluar adecuadamente el proceso, se precisarían de problemas en escenarios más complejos para poder apreciar el comportamiento de los desarrolladores en un margen amplio de tiempo. Además, para poder hablar de selección, sería necesario un número alto de patrones de privacidad (con calidad garantizada) entre los cuales escoger. Asimismo, para poder hablar de un proceso de aplicación aún hacen falta elementos que indiquen un proceso sistemático a seguir.
- Sobre los beneficios para las personas. Estos pueden determinarse utilizando medidas cualitativas orientadas a evaluar el entendimiento, el conocimiento adquirido o la comunicación. No vemos mayores limitaciones para realizar este tipo de mediciones en las condiciones actuales de los patrones de privacidad. Hay que tener en cuenta que, si lo que se busca es una evaluación más profunda del comportamiento de los desarrolladores durante el proceso de diseño, por una parte, se requerirían problemas más complejos y acorde a la práctica real, y por otra, se podrían usar métodos de investigación empírica como los estudios de casos o los estudios etnográficos.

Como se puede notar, las limitaciones actuales de los instrumentos afectan también los enfoques de calidad de los estudios a llevarse a cabo. Sin embargo, si se puede y se debe seguir haciendo contribuciones experimentales. Una vez superadas las limitaciones, los investigadores podrían realizar evaluaciones usando más medidas cuantitativas con las que se podrían realizar estudios comparativos y más generalizables (dentro de contextos adecuados). Esto es, se podrían realizar evaluaciones más objetivas sobre la utilidad de los patrones de privacidad.

Cómo se ha visto en el análisis, parte fundamental de este tipo de estudios es la determinación de las métricas que se pueden utilizar para llevar a cabo las mediciones. La Tabla 4.7 presenta un conjunto amplio de métricas que podrían ser adecuadas para evaluar diferentes enfoques de calidad, que a su vez permitirían definir los beneficios de la aplicación o uso de los patrones de privacidad al diseñar sistemas de información. Varias de estas métricas han sido recogidas de otros estudios en el ámbito más amplio

de los patrones de diseño (Khomh y Guéhéneuc 2008; Zhang y Budgen 2012; Riaz y otros 2015; Khomh y Gueheneuc 2018). Estas métricas están clasificadas según su naturaleza cuantitativa o cualitativa, y acorde a los elementos que normalmente interactúan en los procesos de ingeniería de software: producto, proceso y personas. Este podría constituir un primer paso de cara a avanzar hacia un marco para la evaluación de la aplicación de los patrones, que actualmente es inexistente en el estado del arte.

Tabla 4.7: Enfoques de calidad y métricas para experimentos con patrones de privacidad

Beneficiario	Tipo de medida	Métrica	Enfoque de calidad
Producto	Cuantitativa	Número de las tareas requeridas realizadas	Complejidad
		Número de requisitos cumplidos	Complejidad
		Número de las decisiones arquitectónicas tomadas	Complejidad
		Número de errores existentes	Corrección (calidad de estar correcto)
		Número de casos de prueba fallidos	Corrección
	Cualitativa	Evaluación basada en escala de la adhesión a principios o requisitos de privacidad	Complejidad
		Precisión/utilidad de la solución	Corrección
		Nivel de comprensibilidad (del producto) basado en una escala	Complejidad
		Evaluación basada en escala de la arquitectura	Modularidad
		Evaluación basada en escala de la creatividad	Creatividad
Proceso	Cuantitativa	Tiempo para completar una tarea con éxito	Eficiencia
		Tasa de patrones seleccionados sobre los patrones considerados	Eficiencia
	Cualitativa	El nivel de logro en un conjunto de tareas durante un período de tiempo	Eficiencia
		Evaluación basada en escala del proceso de selección de patrones	Usabilidad
		Evaluación de la selección de patrones en comparación con una solución concebida	Usabilidad
		Evaluación basada en escala de la aplicación de los patrones	Usabilidad
		Evaluación de la aplicación de patrones en comparación con una solución concebida	Usabilidad
Personas	Cualitativa	Eficiencia de aprendizaje	Eficiencia
		Tipo de carga cognitiva alcanzada	Complejidad
		Usando las líneas de comunicación en las transcripciones de conversaciones	Comunicabilidad
		Evaluación a escala de decisiones de diseño documentadas	Comunicabilidad
		Evaluación a escala de la satisfacción durante la realización de una tarea	Usabilidad

4.3.3. Estudiantes como participantes en experimentos en el área

En cuanto a los participantes, los estudiantes encajan perfectamente en esta área de investigación emergente. Por un lado, nuestras condiciones de experimentación están alineadas con las expresadas en [Carver10] [Falessi18] en relación con la participación

de los estudiantes en estudios empíricos: los enfoques de ingeniería de privacidad basados en patrones apenas comienzan a ser explorados y las posibilidades actuales de experimentación están más alineadas con la medición de los beneficios a las personas.

A esto se suma el hecho de que los estudiantes de último año podrían reflejar en buen nivel el conocimiento de los ingenieros ya que ya poseen conocimiento en diseño, tienen cierto nivel de práctica y bajo nivel en temas de privacidad (estudios actuales en escenarios reales han mostrado que los ingenieros adolecen de conocimiento en temas de privacidad [Hadar18] [Senarath18]).

De cualquier forma, al menos por el momento, se requerirá realizar una capacitación previa en temas de privacidad y de diseño con requisitos de privacidad independientemente de si los participantes son estudiantes o ya ingenieros.

4.4. Consideraciones sobre la aplicación de patrones de privacidad

El análisis previo de Sección 4.3.1 desvela que las características de los patrones de privacidad pueden afectar el proceso de aplicación, lo que debería ser tenido en cuenta de cara a futuros estudios sobre la aplicación y adopción de los patrones de privacidad.

Por ejemplo, las soluciones de los patrones usados en este experimento son recomendaciones de alto nivel que hacen que su aplicación se traduzca a un ejercicio de reflexión. Esto es, después de leer estas recomendaciones se conciben nuevas ideas y se toman decisiones basadas en ellas. Este proceso de reflexión basado en las recomendaciones hace que se requiera cierta experiencia previa de parte de los desarrolladores para poder usar los patrones. Al contrario, que los patrones de privacidad describiesen sus soluciones de forma más concreta podría resultar favorable para el entendimiento de los desarrolladores no expertos.

En la misma línea, también consideramos que la representación de las soluciones de los patrones usados podría afectar el proceso de su aplicación. Por ejemplo, se podría considerar el uso de medios de representación, como UML o Diagramas de Flujo de Datos (Data Flow Diagram –DFD), que favorezcan el entendimiento de la solución por parte de los desarrolladores, o incluso, que permitan la automatización del proceso de aplicación.

Posteriormente a la definición de los patrones de privacidad, se hace necesario tener mecanismos para evaluar su calidad. En este estudio, la calidad de los patrones ha sido una limitante para establecer el conjunto de patrones disponible a los desarrolladores para que resuelvan los problemas. En este contexto, el uso de patrones sin garantías de calidad podría afectar de manera negativa el proceso de aplicación como tal e incluso el resultado del mismo. Al contrario, usar patrones con calidad garantizada podría favorecer el proceso de aplicación. Además, la existencia de medios para garantizar la calidad de los patrones de privacidad podría contribuir a homogeneizarlos para obtener resultados más exactos respecto a un grupo de patrones determinado, e incluso, para conseguir posibles generalizaciones.

Además, hay que considerar que el conjunto de patrones de privacidad utilizado es parte de un lenguaje de patrones, el mismo que tiene un conjunto de características propias (p.ej. uso de una plantilla común, la definición de relaciones) que no han sido evaluadas en el ámbito de este estudio, pero que se cree deberían facilitar el proceso de aplicación.

Finalmente, estas consideraciones dan cabida a pensar en que el nivel de abstracción de una solución, el tipo de representación usada, la calidad de los patrones, e incluso las características propias de un lenguaje podrían afectar la aplicación de los patrones. Ante

esto es necesario realizar un estudio, más allá de estas ideas iniciales, para establecer en detalle los elementos que podrían afectar al proceso de aplicación de un patrón. Esto sería necesario para futuros estudios empíricos.

4.5. Resumen del capítulo y contribuciones originales

Este capítulo describe un estudio experimental exploratorio en el ámbito de la aplicación de los patrones de privacidad, con el que se analiza la situación actual de los instrumentos requeridos para experimentar y se elaboran las posibilidades actuales y futuras de experimentación en este ámbito. Para esto, se ha descrito la metodología seguida: un experimento controlado (que evalúa el impacto de aplicar patrones de privacidad en la calidad de una solución diseño y la percepción de los desarrolladores al aplicarlos), detallando su planificación y operación. Posteriormente, se presentan los resultados y se elabora un análisis en profundidad sobre las posibilidades futuras de experimentación.

Este estudio provee las siguientes contribuciones:

- **Un análisis en profundidad de los instrumentos y las oportunidades para la experimentación en la aplicación de patrones de privacidad.**
- Un conjunto de enfoques de calidad (p.ej. completitud de las soluciones), y sus métricas, para experimentos más exhaustivos en el área. Se incluyen aquellos enfoques de calidad que se pueden estudiar ahora, con las limitaciones existentes, o más adelante, una vez superadas las limitaciones actuales.
- El desarrollo de instrumentos de experimentación, que podrían ser reutilizados en experimentos más exhaustivo; estos incluyen problemas de diseño con requisitos de privacidad y una herramienta de soporte a la experimentación con patrones.

Este estudio muestra que actualmente hay posibilidades de experimentar en el área, aunque están limitadas principalmente a realizar evaluaciones cualitativas en el producto de diseño y en las personas (desarrolladores) a cargo. Sin embargo, superadas las limitaciones de los instrumentos actuales, se podrían realizar estudios más exhaustivos agregando medidas cuantitativas. Además, el análisis de las relaciones entre los diferentes instrumentos utilizados hace pensar que el proceso de aplicación de los patrones de privacidad podría estar afectado desde la etapa de su desarrollo mismo, y en etapas intermedias antes de aplicarlos. Por ejemplo, al definir la representación a utilizar o al garantizar la calidad de un patrón. Por ello, en el siguiente capítulo se realiza un estudio más exhaustivo y se define un marco que contiene los elementos que afectan la aplicación de los patrones de privacidad.

Este estudio se ha llevado a cabo en colaboración con investigadores del proyecto “CLIPP - Consolidación del Liderazgo en Investigación en Ingeniería de Privacidad” en el Programa V-PRICIT de Apoyo a la realización de Proyectos de I+D para jóvenes investigadores de la Comunidad de Madrid y la Universidad Politécnica de Madrid.

El experimento se llevó a cabo durante una estancia de investigación en la Escuela Politécnica Nacional del Ecuador, en el marco del curso de

capacitación “Diseño de sistemas con requisitos de privacidad”⁶ organizado por la Escuela Politécnica Nacional y la Universidad de las Américas (Ecuador).

Para la creación de la herramienta de soporte a este estudio se ha asesorado un trabajo fin de carrera con el título “Desarrollo de un sistema de soporte en el proceso de experimentación con patrones de privacidad” [Guzman20].

Además, durante la estancia, se ha llevado a cabo (por dos ocasiones) la conferencia con el título “¡Mi teléfono me escucha! La importancia de la privacidad y los datos personales”^{7 8} para divulgar la investigación en el área de la privacidad y en la ingeniería del mismo.

---

⁶ Curso en diciembre de 2019 <https://detri.epn.edu.ec/proteccion-de-datos-privacidad-desde-el-diseno/>

⁷ Conferencia en diciembre de 2019 <https://detri.epn.edu.ec/mi-telefono-me-escucha/>

⁸ Conferencia en enero de 2020 <https://www.redineditaepn.com/wordpress/mi-telefono-me-escucha-la-importancia-de-la-privacidad-y-los-datos-personales/>

Capítulo 5

Marco de referencia para la aplicación de patrones de privacidad

Los patrones de privacidad han captado mucha atención por parte de investigadores y de organizaciones que necesitan elementos de ingeniería para incluir cuestiones de privacidad en el diseño de los sistemas. Aunque el área de investigación está en una etapa temprana de madurez (Capítulo 3), los patrones de privacidad y propuestas basadas en ellos (p.ej. métodos [Notario15] y estrategias [Hoepman14]) ya constan en guías de ingeniería de privacidad provistas por agencias europeas de ciberseguridad (ENISA [Danezis14], AEPD [AEPD19]).

Mientras los esfuerzos en el área se han enfocado principalmente en el desarrollo de nuevos patrones y en métodos que los usan, el resto de temas aún tienen contribuciones incipientes (p.ej. búsqueda y selección) o incluso nulas (p.ej. aplicación y evaluación de la aplicación). El flujo natural de la investigación en el área [Bafandeh17] y la necesidad de validar su efectividad (para dar soporte a los métodos, a las guías y demás elementos basados en patrones) nos colocan en un punto en el cual se necesita un análisis en profundidad sobre la aplicación de los patrones de privacidad.

En el Capítulo 4 se ha reportado un estudio que inicia el camino en el ámbito de la aplicación de los patrones de privacidad; este describe un experimento exploratorio y analiza las posibilidades de experimentación. El análisis ahí realizado ha puesto de manifiesto que la aplicación de los patrones de privacidad no se reduce únicamente a una mera tarea de instanciación; en su lugar, parece haber elementos que afectan dicha instanciación desde etapas en las cuales los patrones son desarrollados.

La aplicación de los patrones de privacidad no ha sido analizada de manera exhaustiva; se dispone únicamente de estudios aislados que han ido reportando defectos en los patrones [Chung04][Bier12][Hazeyama17][Lenhard17][Suphakul17]. Estos defectos van desde definiciones inconsistentes de lo que es un patrón de privacidad, pasando por descripciones deficientes y problemas de heterogeneidad, hasta la falta de medios para instanciar el patrón adecuado para una tarea en particular.

En este capítulo se elabora **un marco de referencia para la aplicación de patrones de privacidad** que comprende los elementos que pueden afectar su aplicación exitosa, y se lo usa para describir la situación actual. Esto se lleva a cabo adoptando un enfoque para evitar “reinventar la rueda”, aprendiendo del dominio cercano y más maduro de los patrones de seguridad donde la aplicación es un tema de investigación de gran interés [Ito15][Washizaki17]. De este modo, se lleva a cabo una búsqueda sistemática en ese dominio, para, usando la teoría fundamentada, elaborar el marco, el mismo que es usado para estudiar el contexto de los patrones de privacidad.

La estructura de este capítulo se ha organizado de modo que en la Sección 1 se describe la metodología, detallando la estrategia seguida para cada etapa del estudio. En la Sección 2 se presenta y explica el marco de referencia para la aplicación de patrones. En la Sección 3 se describe el estado actual de la investigación en los patrones de privacidad vista a través del marco propuesto. Esto es usado para presentar en la Sección 4 las mejoras priorizadas que podrían realizarse en este ámbito. Finalmente, en la Sección 5 se presenta el resumen del capítulo y las contribuciones realizadas.

5.1. Metodología

Este estudio está compuesto por diferentes etapas (Figura 5.1) que incluyen: una para explorar el dominio de los patrones de seguridad para buscar estudios sobre la aplicación de los patrones, otra para elaborar el marco mediante el análisis y organización de la información encontrada, y otra más para analizar el dominio de los patrones de privacidad.

En la primera etapa se usa una estrategia mixta de búsqueda sistemática de información, siguiendo fundamentalmente las guías propuestas en [Petersen15]. Inicialmente se realiza una búsqueda de información en una base de datos buscando estudios primarios y secundarios; y, luego, se realiza una búsqueda en referencias hacia delante y hacia atrás –*forward and backward snowballing*– en los estudios secundarios incluidos.

En la segunda etapa se adopta el enfoque de la teoría fundamentada –*grounded theory*– [Glaser17] para la construcción de nuevo conocimiento. Así un proceso de análisis de los datos recogidos en la anterior etapa hace que emerjan características relevantes, para luego ser organizadas dentro de un marco.

En la tercera etapa se toma una estrategia similar a la primera; y adicionalmente se incluye un mapeo de diferentes propuestas existentes al marco de referencia propuesto. Con ello se analiza el estado actual de la aplicación en los patrones de privacidad y se identifica contribuciones de impacto para el trabajo futuro.

Figura 5.1: Etapas en la construcción del marco de referencia para la aplicación de patrones

5.1.1. Exploración del dominio de los patrones de seguridad

La búsqueda se llevó a cabo en Scopus, que es considerado como el repositorio más grande de referencias a trabajos revisados por pares [Elsevier20] y con un gran desempeño para identificar trabajos en el área de la computación [Cavacini15]. Se ejecutaron dos cadenas de búsqueda (Tabla 5.1) en febrero de 2019. La primera tenía por objeto buscar aquellos estudios que explícitamente tratan sobre la aplicación de los patrones de seguridad o temas asociados; y por ende, se usan los prefijos de varios términos relacionados con la aplicación. La segunda buscaba estudios secundarios para identificar el tema de la aplicación de patrones dentro del análisis o categorizaciones que en ellos se llevan a cabo.

Tabla 5.1: Cadenas de búsqueda para el dominio de la seguridad

Descripción	Sintaxis para Scopus
Cadena de búsqueda 1 (Usando términos relacionados a la aplicación)	TITLE-ABS-KEY ((appl* OR us* OR implement* OR enforc* OR adopt*) W/3 ("security pattern" OR "security design pattern"))
Cadena de búsqueda 2 (Para estudios secundarios)	TITLE-ABS-KEY ("security pattern" OR "security design pattern") AND ("review" OR "mapping" OR "survey" OR "state of the art" OR "literature")

La cantidad de artículos entre las dos cadenas sumó 322 (230 de la primera cadena y 92 de la segunda) que luego fueron sometidos a un proceso de depuración.

El proceso de depuración incluyó la eliminación de 37 duplicados y un proceso de selección manual de dos iteraciones. En la primera iteración se leyeron los títulos y se marcó cada artículo como: incluido (si abordaba el tema de la aplicación de patrones), excluido (si no abordaba la aplicación de patrones) y no claro (si existían dudas). En la segunda iteración se leyó también el resumen, pero esta vez solo se marcaron los artículos como incluido o excluido. De este modo, se seleccionaron 27 artículos que trataban la aplicación de los patrones de seguridad. Estos incluían 21 estudios primarios tratando el tema de aplicación y 6 estudios secundarios que dentro de sus temas tratados incluían a la aplicación. Los estudios secundarios fueron usados para agregar más artículos a nuestro estudio. Al final, se agregaron dos artículos secundarios; obteniendo en total 29 artículos que trataban sobre la aplicación de los patrones de seguridad (Anexos: Tabla A.2).

5.1.2. Elaboración del marco de referencia para la aplicación de patrones

Esta etapa fue fundamental dentro del estudio y es donde se construyó el marco de referencia presentado. Para esto, se tomaron los estudios seleccionados en la etapa anterior y se analizó su contenido de manera iterativa. Se inició identificando y recolectando aquellos elementos reportados como posibles afectaciones a la aplicación de los patrones de seguridad (p.ej. la plantilla de un patrón y la presentación de un patrón.). Posteriormente se los agrupo en categorías similares, a las cuales llamamos aspectos (p.ej. estructura de un patrón). Por un lado, los aspectos fueron ordenados, desvelando entonces que podían acoplarse a las diferentes etapas del ciclo de vida de los patrones propuesto en [Yoshioka08] (p.ej. la estructura es definida en la etapa de redacción de un patrón). Por otro lado, los aspectos reportaban defectos (p.ej. el uso de diferentes secciones dentro de uno o más patrones para referirse a una misma característica), mejoras (p.ej. usar una plantilla común), o ambos. Todos estos componentes fueron ordenados, obteniendo así el marco presentado más adelante.

5.1.3. Exploración del dominio de los patrones de privacidad

En esta etapa se llevó a cabo una búsqueda en Scopus en febrero de 2019 y se utilizaron tres cadenas de búsqueda (Tabla 5.2). La primera y segunda tenían por objeto buscar aquellos estudios explícitamente asociados a la aplicación de los patrones de privacidad. En la primera se abarcan varias expresiones en torno a los patrones de privacidad. En la segunda se usan términos específicos para los patrones de privacidad y se busca incluir variaciones de términos relacionados con la aplicación. La tercera cadena buscaba estudios secundarios para identificar aquellos que categoricen o analicen el tema de la aplicación de patrones. Además, también se incluyeron artículos pertinentes y que provenían de dos estudios de literatura en patrones de privacidad [Lenhard17][Caiza19].

La búsqueda en Scopus devolvió 72 artículos en total, incluyendo 17 de la primera cadena, nueve de la segunda y 46 de la tercera. A estos se agregaron 22 artículos de los estudios de literatura adicionales. De este modo, se obtuvieron 94 artículos que pasaron a la etapa de depuración.

En la etapa de depuración se eliminaron 29 artículos duplicados. Al finalizar la etapa de selección de artículos basados en la lectura de título y resumen, no se encontró ningún artículo abordando el tema de la aplicación de los patrones de privacidad. Sin embargo, si se obtuvieron 13 artículos reportando defectos ([Chung04][Bier12][Lenhard17][Suphakul17]) y soluciones ([Pearson10][Pearson11][Hoepman14][VanRest14][Xia18]

[Drozd16a][Caiza17][Colesky18][Alshammari18]) individuales que podrían afectar el proceso de aplicación de los patrones. Estos 13 artículos fueron analizados para entender el estado actual en el dominio de los patrones de privacidad.

Tabla 5.2: Cadenas de búsqueda para el dominio de la privacidad

Descripción	Sintaxis de Scopus
Cadena de búsqueda 1 (Proximidad usando términos relacionados a la aplicación)	TITLE-ABS-KEY (("implementability" OR "applicability" OR "enforceability" OR "usability" OR "usage" OR "adoption") AND ("privacy" PRE/1 "pattern"))
Cadena de búsqueda 2 (Variación de términos y proximidad)	TITLE-ABS-KEY ((appl* OR us* OR implement* OR enforc* OR adopt*) W/3 ("privacy pattern" OR "privacy design pattern"))
Cadena de búsqueda 3 (Para estudios secundarios)	TITLE-ABS-KEY ("privacy pattern" OR "privacy design pattern") + búsqueda manual en el título

5.2. Marco de referencia para la aplicación de patrones

5.2.1. Elementos del marco

Acorde a [Yoshioka08], un patrón se mueve a través de un conjunto de etapas que pertenecen a dos procesos principales: la extracción y la aplicación. El proceso de extracción es llevado a cabo por un autor (creador) que tiene por objetivo encontrar patrones y, tras redactarlos, organizarlos y revisarlos, llegar a publicarlos en algún repositorio. El proceso de aplicación es llevado a cabo por un desarrollador que tiene como finalidad llegar a resolver una tarea de diseño. Para ello, identifica un contexto y problema (dentro de una tarea) que podría ser resuelto usando un patrón, luego busca y selecciona el patrón más apropiado, lo aplica, y evalúa los resultados de la aplicación.

Dentro del proceso de aplicación está la etapa de aplicación per se del patrón (a la cual llamaremos instanciación para evitar confusiones). La instanciación no es una tarea aislada; sino que, existen elementos que la afectan que vienen desde las etapas tempranas del ciclo de vida. Por ejemplo, una descripción deficiente de un patrón puede provocar que los desarrolladores no lo comprendan y, por ello, no puedan instanciarlo.

La **instanciación de un patrón** es una etapa específica en la cual el desarrollador usa la solución establecida dentro del patrón para resolver una tarea determinada.

La instanciación es una de las etapas dentro del **proceso de aplicación** que es llevado a cabo por un desarrollador; este proceso incluye: el reconocimiento de un problema y contexto donde se podría instanciar un patrón, la selección de los patrones, su instanciación y la evaluación de la instanciación.

Estos elementos que pueden influir en la instanciación exitosa de un patrón, pueden hacerlo de manera negativa, siendo defectos, (p.ej. a través de descripciones heterogéneas) o de manera positiva, siendo mejoras, (p.ej. usando términos normalizados en la descripción). Tanto los defectos como las mejoras se han agrupado dentro de categorías a las cuales hemos llamado aspectos, los cuales están distribuidos a lo largo de las diferentes etapas del ciclo de vida de los patrones.

Un **aspecto de aplicación** es una categoría que incluye elementos que afectan positivamente (**mejoras**) y elementos que afectan negativamente (**defectos**) a la instanciación exitosa de un patrón.

El **marco de referencia para la aplicación** de patrones incluye todos los aspectos de aplicación de los patrones, organizados acorde a las etapas del ciclo de vida de un patrón (Figura 5.2).

La Figura 5.2 muestra todos los elementos del marco, definidos hasta el momento, organizados y distribuidos tanto en el proceso de extracción como en el de aplicación. Por el momento, el marco incluye los aspectos de los cuales se han encontrado defectos o mejoras reportados en el dominio de la seguridad hasta la realización del estudio; por lo que, el marco podría ser extendido posteriormente.

Figura 5.2: Marco de referencia para la aplicación de patrones

5.2.2. Aspectos de aplicación

Los aspectos agrupan tanto defectos como mejoras que podrían afectar una instanciación exitosa. Idealmente, si existe un defecto, debería haber una mejora, ya propuesta o a proponerse, que lo corrija. Por ejemplo, dentro de la etapa de *redacción*, en el aspecto *estructura* se puede mencionar el defecto de tener diferentes plantillas de patrones para un mismo tipo (o catálogo), cuya mejora sería el tener una plantilla común para esos patrones. Pero también podría existir una mejora que, aunque no esté asociada a un defecto per se, facilite la tarea de instanciación de un patrón. Por ejemplo, dentro de la etapa de *selección*, en el aspecto *selección* una técnica para localizar el patrón más apropiado es una mejora que facilitará la instanciación.

A continuación, se definen los diferentes aspectos establecidos y se describen defectos y mejoras encontradas en los artículos estudiados del dominio de los patrones de seguridad.

Aspectos en la etapa de redacción

Ámbito. – Este aspecto trata la determinación del dominio al cual pertenece un patrón. Por ejemplo, un defecto que se señala frecuentemente es el uso indistinto del término genérico "patrón de seguridad" para diversas tareas (ámbitos) en el proceso de desarrollo de software: en la recolección de requisitos de seguridad, en tareas de análisis, en el diseño arquitectónico, en la administración, etc. Esta diversidad de ámbitos a los que se refiere el mismo término "patrón de seguridad" puede provocar

inconvenientes de heterogeneidad en las descripciones de los patrones, lo cual hace que se dificulte su comprensión.

La diversidad de ámbitos denota la existencia de diferentes tipos de patrones. Una taxonomía que defina claramente los diferentes tipos de patrones y sus respectivos ámbitos podría solventar el defecto.

Estructura. – Este aspecto se refiere a la organización del contenido de un patrón. Los defectos encontrados dentro de este aspecto incluyen el uso de diferentes plantillas para organizar el contenido de patrones dentro de un mismo ámbito. Esto puede provocar el uso de diferentes secciones y la definición de diferentes tipos de opcionalidad (una sección obligatoria u opcional) para un mismo tipo de contenido. Estos defectos pueden dificultar tareas siguientes como la clasificación de los patrones, la selección de los patrones, y la construcción de herramientas que automaticen esas tareas.

Una plantilla común es una mejora que normaliza las secciones que deberían tener los patrones del mismo tipo.

Comprensión. – Incluye las características que hacen que la descripción de un patrón sea entendible. Algunos defectos que se han encontrado en este aspecto incluyen descripciones abstractas, una especificación excesiva o insuficiente y la ausencia de uso de formas de representación que faciliten su comprensión (p.ej. gráficas). Las descripciones deficientes pueden dificultar la lectura, la comprensión y la instanciación de los patrones.

Posibles mejoras incluyen el uso de ejemplos para ayudar a entender un patrón mostrando cómo podría instanciarse, el uso de pistas de evaluación que indican cómo se va a probar un patrón y el uso de medios de representación útiles para hacer más comprensible o reutilizable la solución expresada en un patrón. Por ejemplo, una solución expresada en un lenguaje de modelado conocido, como UML, podría ser más fácil de entender, permitiría abordar problemas relacionados con la abstracción y con la heterogeneidad de las descripciones, y, además, podría facilitar el desarrollo de técnicas para la instanciación automatizada de las soluciones de los patrones.

Aspectos en la etapa de organización

Clasificación. – Se ocupa de cuestiones de categorización de los patrones. Entre los defectos que actualmente se han reportado están la existencia de demasiados esquemas de clasificación (p.ej. por dominio de aplicación o por principio de seguridad), varios de ellos sin lograr total plenitud (esto es, poder categorizar a cualquier patrón de seguridad) y exclusividad mutua (esto es, colocar un patrón de seguridad en una sola categoría dentro de una clasificación dada). Esta variedad de clasificaciones, o su deficiencia en la plenitud y exclusividad mutua, dificulta la organización de los patrones de seguridad tanto para los autores que buscan donde colocar sus patrones, como para los desarrolladores que quieren localizar y seleccionar un patrón.

Un esquema de clasificación común es usualmente discutido como mejora para hacer frente a los defectos relacionados a la variedad. En la misma línea, el esquema debería cumplir con ser pleno y mutuamente exclusivo.

Interconexión. – Este aspecto comprende cuestiones relacionadas a las formas de conectar patrones (p.ej. algunos patrones pueden complementarse para resolver un problema). Las conexiones pueden ayudar a los desarrolladores, y a las herramientas de soporte, a navegar a través de grandes repositorios de patrones hasta encontrar el más apropiado para un problema determinado.

Aquellas mejoras en este aspecto deben ir más allá de únicamente establecer la existencia de una conexión, y proponer diferentes tipos que provean a los autores de información específica que los ayuden a comprender más profundamente las conexiones. Otras soluciones más sofisticadas que requieren del establecimiento de las conexiones entre los patrones incluyen los sistemas y los lenguajes de patrones. Estos elementos contienen grupos de patrones relacionados y cohesivos en un dominio determinado.

Aspectos en la etapa de revisión

Calidad. – Este aspecto se refiere a las características asociadas a los patrones que permiten juzgar su valor (p.ej. llegar a definir si es bueno o malo). Una de las carencias actuales es la falta de un mecanismo de evaluación de los patrones previo a que estos se hagan públicos. Lo cual puede significar que patrones deficientes y sin garantías estén a disposición de los desarrolladores.

La mejora directa es la definición de criterios de calidad, y mecanismos basados en ellos, para determinar los niveles de calidad (cuantitativos, cualitativos o ambos) de un patrón previo a su publicación. Una propuesta de este tipo permitiría aumentar las posibilidades de una instanciación exitosa al garantizar niveles mínimos de calidad en el contenido de un patrón. Además, promovería un bucle de mejora continua, basado en una identificación y resolución de defectos, de los patrones previo a ser publicados.

Aspectos en la etapa de publicación

Almacenamiento. – Este aspecto comprende cuestiones relacionadas al almacenaje y acceso a los patrones. Los defectos reportados incluyen la falta de un esquema de codificación para facilitar su almacenamiento y acceso por medios automáticos, la existencia de repositorios incompletos y la carencia de medios que posibiliten modificarlos. Esto dificulta la creación de herramientas que automaticen procesos relacionados al acceso, a la búsqueda y a la modificación de los patrones; e incluso la interoperabilidad entre varias herramientas.

Una posible mejora es usar un sistema de codificación común, normado o bien conocido, para facilitar el almacenamiento y el uso de los patrones (p.ej., Extensible Markup Language –XML– o JavaScript Object Notation –JSON). Asimismo, definir un repositorio común proporciona un lugar único para que los desarrolladores puedan acceder a los patrones.

Aspectos en la etapa de selección

Selección. – En este aspecto se incluyen cuestiones relacionadas con la búsqueda y elección de los patrones más apropiados para su posterior instanciación. Como posibles mejoras se han encontrado técnicas de búsqueda y selección, que se traducen en formas sistemáticas de abordar el patrón más apropiado para un problema determinado. En la misma línea, las herramientas pueden implementar las técnicas para proveer soporte automatizado.

Aspectos en la etapa de instanciación

Instanciación. - Este aspecto agrupa los medios que guían la instanciación de un patrón. Todas las demás mejoras deberían contribuir a la instanciación exitosa de un patrón.

Los medios para llevar a cabo esta tarea incluyen: técnicas de instanciación, que actuarían como formas sistemáticas de aplicar un patrón (p.ej. técnicas usadas en la ingeniería dirigida por modelos [Schmidt06]), procedimientos que incluyan

instrucciones o recomendaciones para instanciar un patrón (p.ej. directrices que expliquen cómo instanciar un patrón), y herramientas de instanciación que implementen las técnicas o procedimientos para proveer apoyo automatizado al desarrollador.

Aspectos en la etapa de evaluación de la aplicación

Evaluación. – Este aspecto incluye los medios que podrían utilizarse para determinar los efectos que la instanciación de los patrones tiene en el proceso, en el producto de diseño, o en las personas a cargo. Esta evaluación podría contribuir a validar y perfeccionar las mejoras reportadas en el resto de aspectos. En el ámbito de los patrones de seguridad aún hay una carencia de mecanismos de evaluación.

Un tipo de mecanismo podría ser las técnicas de evaluación (p.ej. técnicas de evaluación de modelos) y herramientas que implementen estas técnicas.

Teniendo el marco y sus aspectos constituyentes ya descritos, es necesario determinar el estado actual en el contexto de los patrones de privacidad, y definir líneas de acción.

5.3. Estado actual de los aspectos de aplicación en los patrones de privacidad

Tras efectuar la búsqueda en el dominio de los patrones de privacidad, no se ha encontrado estudio alguno que analice la aplicación de los patrones de privacidad y tampoco los aspectos de aplicación. Sin embargo, hay contribuciones separadas que, por un lado, reportan defectos existentes en los patrones de privacidad, y, por otro lado, proponen soluciones que equivalen a las mejoras dentro del marco propuesto. A continuación, se listan los defectos y mejoras encontradas en los patrones de privacidad dentro de cada aspecto del marco establecido.

Ámbito. – Un inconveniente existente es el uso indiscriminado del término "patrón de privacidad" para referirse a diferentes tipos que pueden aplicarse a diferentes tareas dentro del proceso de desarrollo de software (p.ej. al análisis de requisitos, al diseño, a la definición de proceso [Hazeyama17]). El referirse a los diferentes tipos de patrones de privacidad como si fueran los mismos puede ser el desencadenante de otros defectos relacionados a la heterogeneidad en los patrones dentro de un mismo ámbito. Por ejemplo, en su descripción o en su forma de representación.

Estructura. – En este caso no se han encontrado defectos reportados; pero sí, una plantilla común establecida en el repositorio ubicado en *privacypatterns.org* [Doty11b] y acordada por la comunidad de investigadores que ahí colaboran. De modo que los autores pueden usar dicha plantilla y publicar ahí sus patrones para que los desarrolladores puedan acceder a ellos. Esta solución contribuye a normalizar los patrones de privacidad y evita los inconvenientes de la heterogeneidad de estructura de los contenidos.

Comprensión. – En este aspecto se han encontrado la mayoría de defectos; entre ellos se incluyen descripciones muy abstractas que carecen del nivel de detalle y claridad apropiados, la variación en el nivel de abstracción y descripción entre patrones de privacidad en un mismo ámbito, y los diferentes estilos de redacción y vocabularios utilizados al escribir. Respecto a la representación, hay mucha variedad en las formas de representación en patrones del mismo tipo; por ejemplo, se usan modelos, interfaces gráficas de usuario, procesos e incluso código. A esto se suma el hecho de que los modelos están infrutilizados a pesar de ser los medios habituales de representación en el diseño de sistemas. Estos inconvenientes dificultarían el entendimiento de los patrones y su posterior instanciación.

La plantilla de patrones definida en [Doty11b] contiene una sección opcional para colocar ejemplos motivacionales, que debería mostrar cómo aplicar un patrón. Sin embargo, muchos de esos ejemplos siguen presentando inconvenientes como la heterogeneidad y la falta de representaciones más útiles. De hecho, medios de representación útiles han sido poco investigados a pesar de las ventajas que podrían proveer a la tarea de instanciación. Estos podrían ayudar a mejorar la comprensión a través del uso de elementos familiares al desarrollador, podrían ser una forma de normalización de las descripciones heterogéneas al utilizar notaciones comunes, o podrían ser utilizados en técnicas de transformación de modelos usadas en la ingeniería dirigida por modelos [Schmidt06], etc. En [Suphakul17] se presenta un estudio en el que se propone una serie de patrones utilizando UML para representar las soluciones.

Clasificación. – En este aspecto vale la pena destacar dos esquemas de categorización por su utilización en grandes colecciones de patrones: aquel reportado en [Drozd16a] y que usa los principios e instrucciones de la norma ISO/IEC 29100 [ISO/IEC29100] y aquel que usa una serie de estrategias y tácticas [Hoepman14] [Colesky16] derivadas de los principios de la ISO/IEC 29100 [ISO/IEC29100] y de las directrices de la Organización de Cooperación y Desarrollo Económicos (OECD) [OECD80]. El esquema basado en estrategias y tácticas ha sido adoptado en el repositorio en [Doty11b].

Además de estos esquemas de clasificación también existen otras propuestas, como aquellas basadas en dominios de aplicación; sin embargo, no son un medio organizado ya que se limitan a utilizar una variedad de etiquetas sin límite impuesto.

Interconexión. – Dentro de este aspecto se engloban dos propuestas de lenguajes de patrones, el uno para sistemas de interacción entre usuarios [Schümmer04] y el otro para proveer anonimidad a las comunicaciones [Hafiz13]. El resto de contribuciones ha llegado a establecer catálogos de patrones, pero hace falta evolucionarlos hacia sistemas o lenguajes de patrones.

Calidad. –Aquí nos encontramos con el inconveniente de no contar con ninguna clase de medios o marcos metodológico para evaluar la calidad de los patrones de privacidad. De momento, hasta donde sabemos, el único medio para evaluar los patrones son los eventos PLOP (p.ej. EuroPLOP en el caso europeo), que son espacios donde los expertos en el área de patrones realizan un proceso de evaluación por pares y de mejora iterativa a las contribuciones aceptadas conocida como pastoreo.

Almacenamiento. – En este aspecto se puede aplicar como mejoras el uso de un esquema común de codificación y la utilización repositorios comunes. Ambos tipos de contribuciones pueden ir de la mano para facilitar un mejor almacenamiento y acceso. Precisamente el repositorio establecido en [Doty11b] va en camino de ser uno común, y de momento integra, entre otros, catálogos como [PRIPARE15] o el reportado en [Drozd16a]. Este repositorio común guarda un conjunto de patrones codificados utilizando el lenguaje de marcado ligero “Markdown” (.md).

Selección. – Dentro de este aspecto, se han encontrado dos contribuciones presentadas como herramientas para seleccionar los patrones de privacidad más apropiados. Una de ellas es el catálogo en línea propuesto en [Drozd16a] que aprovecha su esquema de categorización jerárquica para abordar el patrón más adecuado. La otra es una herramienta para la selección automática de patrones de privacidad [Pearson10] basada en reglas de selección.

Instanciación. – En este aspecto hacemos referencia a una sección dentro de la plantilla definida en [Doty11b] y que se usaría para definir las guías del patrón. No se han encontrado propuestas metodológicas adicionales como técnicas o herramientas. Aunque los repositorios han sido en ocasiones referidos como herramientas de aplicación, bajo el marco establecido en este trabajo, son considerados en el aspecto de almacenamiento.

Evaluación. – En este aspecto no se han encontrado propuestas.

En general, los defectos y mejoras que se han encontrado en el dominio de los patrones de privacidad están alineados con aquellos encontrados en el de los patrones de seguridad. Sin embargo, existe una diferencia en el nivel de cobertura de los aspectos; esto es, hay diferencia entre la cantidad de reportes de defectos, así como las mejoras. En el dominio de los patrones de seguridad hay más contribuciones realizadas precisamente debido a su mayor nivel de madurez.

5.4. Guía para mejorar la instanciación de los patrones de privacidad

Una vez presentado el estado de los diferentes aspectos de aplicación de los patrones de privacidad, es necesario analizar las mejoras en las que la investigación en el área debería enfocarse de cara a fomentar una instanciación exitosa de los patrones.

Para orientar el análisis, la Figura 5.3 muestra el marco de referencia para la aplicación incluyendo los aspectos organizados acorde a las etapas del ciclo de vida de los patrones y las mejoras ya reportadas en el dominio de los patrones de seguridad (recuadros redondeados). Las mejoras están colocadas bajo el aspecto al cual contribuyen principalmente y apuntan con flechas a otros aspectos a los cuales también podrían ayudar.

Para mapear las mejoras existentes en el dominio de los patrones de privacidad, se ha agregado una capa de color que indica el nivel de cobertura; el blanco muestra que no hay contribuciones, el celeste ligero indica que hay una baja cantidad de contribuciones pero que no se tiene reportes de uso o de experiencias empíricas, y el celeste fuerte señala una baja cantidad de contribuciones pero que se conoce un uso moderado en otras propuestas de investigación.

Como se puede notar en la Figura 5.3, existe una clara concentración de contribuciones (más recuadros pintados) en las etapas dentro del proceso de extracción. Esto guarda coherencia con el hecho de que es necesario primero construir y desarrollar patrones para luego aplicarlos. Sin embargo, esta posición no prohíbe el analizar el proceso de aplicación para desarrollar patrones con características que favorezcan posteriormente a una instanciación exitosa. Es más, este marco muestra que la instanciación no es una tarea aislada, y que, para fomentar una instanciación exitosa, se requiere tomar medidas desde las etapas del proceso de extracción.

Dentro del proceso de extracción, la mayoría de mejoras se han centrado en los aspectos de *estructura*, *comprensión*, *clasificación* y *almacenamiento*; y se ha prestado poca atención al *ámbito*, a la *interconexión* y a la *revisión*, lo que ha dado lugar a que los autores reciban poca ayuda para discriminar entre diferentes tipos de patrones de privacidad, para relacionar los patrones y para garantizar su calidad. Estos aspectos que han recibido poca atención deberían ser considerados para futuras contribuciones.

Figura 5.3: Estado de las mejoras a la aplicación de los patrones de privacidad

5.4.1. Priorización de las mejoras

Además del criterio de existencia o no de propuestas dentro de los aspectos, aquí se definen dos perspectivas para determinar el impacto que una propuesta de mejora podría tener. Por una parte, determinamos la cantidad de aspectos cuyos defectos podrían ser solventados por una determinada mejora; y por otra, nos fijamos en la utilidad que una mejora tendría de cara a ser usada por otras.

Considerando la primera perspectiva, las mejoras que más aspectos ayudan a cubrir al solucionar sus defectos (representadas en la Figura 5.3 a través de los recuadros con mayor tamaño) son:

- Una taxonomía de tipos de patrones que defina y separe el ámbito en el cual se aplica cada tipo de patrón. Además, el tener un conjunto de patrones agrupados en categorías específicas ayuda a controlar los defectos relacionados a la heterogeneidad en la estructura y descripciones de los patrones, aspectos *estructura* y *comprensión* respectivamente. Asimismo, una taxonomía vista como un esquema de clasificación común (aspecto *clasificación*) organiza los patrones y sirve como medio para aproximarse a los patrones más adecuados a usar (aspecto *selección*).
- Un esquema de clasificación común que resuelva el defecto de tener una diversidad de categorizaciones. Del mismo modo, puede contribuir al aspecto *interconexión* ya que permite agrupar patrones relacionados dentro de una categoría y al aspecto *selección* ya que provee de un nivel de aproximación hacia un grupo de patrones apropiados para resolver un problema determinado. Igualmente, dado que agrupa patrones dentro de una categoría puede ayudar a evitar los defectos de heterogeneidad en las descripciones (aspecto *comprensión*).

- Las relaciones que permitan conectar los patrones entre sí. El tener patrones relacionados podría ayudar a mejorar la comprensión de los mismos a través del contexto (aspecto *comprensión*). Así también, son medios explícitos para navegar en colecciones de patrones, ayudando por tanto en el aspecto de *selección*.
- Los sistemas y lenguajes que podrían ayudar a solventar muchos defectos debido a las características que debe cumplir un conjunto de patrones para alcanzar ese nivel. Sus propuestas podrían solventar defectos relacionados a la heterogeneidad en la *estructura* al definir una plantilla común, pueden agrupar los patrones al tener un esquema común de *clasificación* y un solo repositorio (aspecto *almacenamiento*), pueden ayudar en el aspecto *selección* a través de las relaciones establecidas entre sus patrones y también aportarían a la *instanciación* ya que deberían proveer guías de implementación.

Desde la perspectiva de la utilidad, algunas mejoras pueden ser usadas, o incluso ser un requisito, en la construcción de otras. Las mejoras más útiles (representadas en la Figura 5.3 en la parte superior de las mejoras alineadas en cada aspecto) son:

- Una taxonomía de tipos de patrones que podría ayudar a casi todas las otras mejoras debido a que provee una organización esencial de los dominios en los cuales tienen injerencia determinados tipos de patrones de privacidad, y, por lo tanto, sus mejoras. Por ejemplo, patrones de privacidad orientados a un ámbito de diseño arquitectural hacen que se usen medios de representación apropiados como los modelos arquitecturales.
- Los medios de representación útiles ya que pueden utilizarse al definir las secciones de un patrón, como la solución, los ejemplos y las pistas de evaluación; incluso podría estar definido dentro de una plantilla común. Al ser usado como medio de normalización, su existencia puede ser considerada como parámetro de calidad. Además, pueden ser usados por los medios de instanciación dentro de los sistemas y lenguajes de patrones. Asimismo, el uso de representaciones adecuadas puede ayudar a definir técnicas, guías y herramientas sustentadas en ellas; por ejemplo, aquellas basadas en transformaciones de modelos.
- Un esquema de clasificación común y el establecimiento de relaciones son requisitos para el desarrollo de sistemas y lenguajes de patrones. Asimismo, pueden formar parte de una plantilla común. Además, su existencia puede ser un criterio de calidad dado que garantizan organización; y también, pueden ser usadas dentro de técnicas y herramientas de selección. En el caso de una clasificación, hace que sea posible el acercarse a través de una categoría, a un grupo de patrones apropiados para resolver un problema. En el caso de las relaciones, puede permitir la navegación entre los patrones dentro de un conjunto.

Al cruzar la información en torno a la existencia o no de contribuciones, la capacidad de resolución de defectos y la utilidad para construir otros elementos, las mejoras más prioritarias y que más impacto podrían causar son: la taxonomía de tipos de patrones, el establecimiento de relaciones y los sistemas de patrones.

Finalmente, hay que mencionar que elementos como los lenguajes de patrones y los medios de representación útiles se podrían abordar en investigaciones posteriores ya que de momento hay contribuciones escasas. Respecto a los lenguajes de patrones, solo conocemos dos [Schümmer04][Hafiz13], y estos se aplican por naturaleza a dominios muy específicos; por lo que se requiere de más lenguajes de patrones para cubrir sistemas de información en varios dominios de aplicación relevantes (p.ej. sistemas móviles, de geolocalización, de salud, etc.). Considerando los medios de representación, en [Suphakul17] se ha empezado a explorar el uso de UML en la representación de patrones de privacidad; sin embargo, no se han explorado otras formas de representación como los diagramas de flujos de datos (DFD) que están siendo propuestos en la elicitación de requerimientos con criterios de privacidad [Deng11].

5.4.2. Próximos pasos

Previamente se ha establecido los tres tipos de mejoras que son más prioritarios y que más impacto pueden causar para conseguir la instanciación exitosa de un patrón. De estos, aunque no hay una propuesta formal de los tipos de patrones de privacidad, en [Hazeyama17] se hace una primera aproximación. Los investigadores describen rápidamente diferentes colecciones de patrones de privacidad, denotando que existen diferentes tipos; sin embargo, no hacen un mayor análisis o agrupación. Consideramos que, a la vez que se denota la problemática, es también una primera noción de cómo organizar los tipos de patrones de privacidad.

Por otro lado, no ha habido propuestas de cómo relacionar los patrones de privacidad de manera uniforme y sistemas de patrones de privacidad. Precisamente, este tipo de contribuciones encaja con la coyuntura actual de investigación en el ámbito de patrones de privacidad; donde hay varias propuestas esparcidas y un repositorio común que alberga un grupo de patrones de privacidad que solo ha llegado al nivel de ser considerado catálogo. Es por ello que hemos concentrado nuestros esfuerzos en proveer un medio para el **establecimiento de relaciones** entre patrones de privacidad y en la creación de **sistemas de patrones** de privacidad. Consideramos que son parte fundamental para mejorar la aplicación de los patrones y para permitir la evolución de los grupos de patrones de privacidad.

5.5. Resumen del capítulo y contribuciones originales

Este capítulo propone un marco de referencia que ayuda a comprender cuáles son los elementos que afectan la instanciación exitosa de los patrones de privacidad y en qué etapa dentro del ciclo de vida de los patrones lo hacen. Se describe la metodología seguida para construirlo a partir del conocimiento previo en el dominio de los patrones de seguridad, que usa esencialmente una búsqueda sistemática de información en una base de datos extensa y la teoría fundamentada para construcción del conocimiento. Asimismo, se detalla el procedimiento seguido para analizar el estado del dominio de los patrones de privacidad de acuerdo al marco de referencia. Las contribuciones directas de este estudio son:

- **Un marco de referencia para la aplicación de patrones de privacidad** compuesto por un conjunto de aspectos que agrupan defectos y mejoras que pueden afectar la instanciación exitosa de un patrón, organizados según las etapas del ciclo de vida de un patrón.
- La descripción del estado actual de los defectos y mejoras existentes en la investigación de patrones de privacidad.

- Un análisis de aquellas mejoras que más impacto tendrían de cara a fomentar la aplicación de los patrones de privacidad.

Este marco brinda una perspectiva completa y muestra que la instanciación no es una tarea aislada, y que, para fomentar una instanciación exitosa, se requiere tomar medidas desde las etapas iniciales del ciclo de vida de los patrones. El desarrollo de mejoras dentro de los diferentes aspectos identificados permitirá fomentar una aplicación exitosa de los patrones de privacidad con resultados positivos. A posteriori, esto podría demostrar su utilidad y fomentar su adopción.

A pesar de que queda mucho por investigar en esta área, el análisis realizado ha permitido definir que las contribuciones que más impacto tendrían son aquellas enfocadas en los medios para el **establecimiento de relaciones** entre patrones de privacidad y en la creación de **sistemas de patrones** de privacidad. Estas son abordadas en los siguientes capítulos.

Los resultados de este estudio se han publicado en el artículo “A framework and roadmap for enhancing the application of privacy design patterns” [Caiza20] que fue presentado en la conferencia “The 35th ACM/SIGAPP Symposium On Applied Computing” (Ranking CORE⁹ B).

Este estudio se ha llevado a cabo en colaboración con investigadores del proyecto “TRUESSEC.eu - TRUst-Enhancing certified Solutions for SEcurity and protection of Citizens’ rights in digital Europe”¹⁰ dentro del programa de investigación e innovación Horizonte 2020 de la Unión Europea.

En la fase preliminar de la realización de este estudio se ha llevado a cabo la dirección de un trabajo fin de máster con el título “A systematic literature review on the applicability of security patterns” [García18]. Este trabajo obtuvo la distinción de **Matrícula de Honor** en el Máster Universitario en Ingeniería de Redes y Servicios Telemáticos de la Universidad Politécnica de Madrid.

---

⁹ Ranking Computing Research and Education Association <http://portal.core.edu.au/conf-ranks/>

¹⁰ TRUESSEC.eu <https://cordis.europa.eu/project/id/731711>

Capítulo 6

Relaciones entre patrones de privacidad

En los capítulos previos se ha manifestado la relevancia de investigar el proceso de aplicación de los patrones de privacidad de cara a probar las características y beneficios que se les atribuyen. Lo cual permitirá dar sustento a las propuestas construidas sobre ellos (p.ej. los métodos que ya constan en recomendaciones de ciberseguridad [Danezis14] [AEPD19]), y fomentará su adopción por los desarrolladores.

En el Capítulo 5 se ha demostrado que la aplicación no se reduce a una simple tarea de instanciación; sino que, comprende varios elementos que pueden afectarla de manera negativa (p.ej. al tener estructuras heterogéneas) o positiva (p.ej. al usar una estructura común), y que se ubican a lo largo de todo el ciclo de vida de un patrón. Estos elementos han sido organizados en aspectos dentro de un marco de referencia y han sido priorizados. Las relaciones entre los patrones son una de las mejoras prioritarias que más impacto podrían causar.

Adicionalmente, el establecimiento de relaciones entre los patrones de privacidad se acopla a las condiciones actuales de los catálogos existentes. Por un lado, hay varios conjuntos de patrones de privacidad dispersos [Schümmer04][Lobato09][Chung2004][Hafiz13][Fischer-Hübner10][Drozd16a][PRIPARE15][Colesky16], con sus propias características y defectos. Por otro lado, hay un esfuerzo común (compuesto por investigadores de Clever, Escuela Politécnica Nacional, Inria, Radboud University, UC Berkeley School of Information, Ulm University, Universidad Politécnica de Madrid, y Vienna University of Economics and Business) para integrar los catálogos dispersos en un repositorio común en *privacypatterns.org* [Doty11b], y para evolucionarlos.

Parte permitir esta evolución se debe interconectar los patrones de privacidad a través del establecimiento de relaciones; sin embargo, hay problemas a solventar. Los catálogos de manera individual ya presentan varios defectos que incluyen la ausencia de relaciones, inconsistencias y descripciones laxas entre aquellas pocas establecidas. Estos empeoran por cuestiones de heterogeneidad entre los catálogos. Por ello se hace necesario un esquema normalizado para facilitar el establecimiento de las relaciones entre los patrones dentro de los catálogos individuales y en el repositorio común.

Este capítulo propone **una taxonomía de tipos de relaciones para definir conexiones entre patrones de privacidad**, que ayudaría a interconectar los patrones, lo que aportaría a su proceso de aplicación. Esta taxonomía funciona como una caja de herramientas que ayuda a los investigadores a comprender, seleccionar y establecer las relaciones correctas entre los patrones de privacidad. Además, la taxonomía ya aplicada podría colaborar al entendimiento de los desarrolladores a través del contexto, e incluso los podría guiar hacia aquellos más adecuados para resolver un problema. Esta taxonomía ha sido utilizada para construir dos sistemas de patrones (Capítulo 7).

La estructura de este capítulo se ha organizado de modo que en la Sección 1 se describe la metodología seguida, para en la Sección 2 presentar y explicar cada tipo de relación dentro de la taxonomía propuesta. En la Sección 3 se realiza un análisis profundo donde se valida la taxonomía, se desvelan las limitaciones de los esquemas previos, se explican los beneficios de la taxonomía, y se mencionan vías de extensión a la misma. Finalmente, en la Sección 4 se presenta el resumen y las contribuciones originales.

6.1. Metodología

Este estudio se lleva a cabo adoptando una estrategia de triangulación en la cual, por una parte, se analizan tipos de relaciones genéricos y más maduros, y por otra, se mapean las conexiones¹¹ (aquellas existentes) de los catálogos de patrones de privacidad individuales que se están integrando. Esto se lleva a cabo en un proceso en dos fases (Figura 6.1).

Figura 6.1: Etapas en el desarrollo de la taxonomía de tipos de relaciones

En la primera fase se analizaron los tipos de relaciones propuestos en los trabajos de [Noble98] y de [Meszaros97] para componer una primera propuesta de taxonomía. El primer trabajo es ofrecido por EuroPLOP [EuroPLOP04], uno de los eventos más importantes dentro de la comunidad de expertos en patrones, para apoyar a los autores de patrones y presenta un conjunto genérico de tipos de relaciones. Mientras que el segundo trabajo plantea tipos de relaciones para el área más específica, pero muy madura y de amplia adopción, de los patrones de diseño orientado a objetos.

Los tipos de relaciones de ambos estudios fueron analizados e integrados en una sola taxonomía. En este proceso se solventaron discrepancias y se erigieron definiciones alineadas. Entre las discrepancias se encontraron definiciones de tipos de relaciones referidas con nombres diferentes (p.ej. especializar y refinar), relaciones abordadas desde la perspectiva contraria (p.ej. conducir a y establecido por), relaciones cuya definición podría contener a otras (p.ej. la definición de la relación estar-en-conflicto de [Noble98] podría incluir elementos alternos y similares), etc.

En la segunda fase se realizaron dos tareas: recolección y mapeo. En la primera se recolectaron las relaciones expresadas en los catálogos individuales. Esto desveló los problemas existentes: algunos catálogos no expresaban en absoluto las relaciones, otros proporcionaban soluciones improvisadas, otros simplemente expresaban la existencia de una relación sin más detalles; y, además, no siempre tenían esquemas normalizados. La segunda tarea se orientó a mapear las relaciones actuales existentes en los catálogos individuales a la taxonomía de los tipos de relaciones propuesta en la primera fase. Mientras se llevaba a cabo el mapeo, se mejoraron y perfeccionaron las definiciones de los tipos de relaciones de la taxonomía propuesta.

6.2. Taxonomía de relaciones

Al finalizar el procedimiento previamente descrito, se han identificado nueve tipos de relaciones (Tabla 6.1) aplicables tanto a los catálogos individuales de patrones de privacidad como al catálogo común que está siendo conformado. Cada tipo de relación se ha concebido (y se ha descrito) para relacionar dos patrones; sin embargo, las

---

¹¹ Para facilitar la lectura, se usa en ocasiones la palabra conexión para referirse a una relación, la cual es una instancia de un tipo de relación.

relaciones n-arias también podrían usar esta propuesta, dado que podrían descomponerse en relaciones binarias.

A continuación, se explican los tipos de relaciones propuestos. Para ello se establece su definición, se exponen las características de la relación (p.ej. respecto a su simetría), y se brinda un ejemplo del tipo de relación propuesto usando patrones pertenecientes a los cuatro catálogos que actualmente expresan conexiones¹² [Fischer-Hübner10][Hafiz13] [PRIPARE15][Colesky16]. La Tabla 6.2 muestra una descripción breve de los patrones de privacidad usados para los ejemplos.

Tabla 6.1: Tipos de relaciones para los patrones de privacidad

Tipo de relación	Descripción
Alternativa	Cuando existen dos patrones que resuelven el mismo problema en contextos superpuestos a través de soluciones diferentes.
Complemento	Cuando hay dos patrones que resuelven sus propios problemas, pero que pueden trabajar juntos (complementándose mutuamente) para proveer una solución con más alcance que la de los patrones individuales.
Conducción	Cuando la aplicación de un patrón conduce a un nuevo problema, que podría ser resuelto por otro patrón.
Conflicto	Cuando hay dos patrones, cuyas soluciones son mutuamente excluyentes (independientemente de los problemas y contextos).
Especificación	Cuando un patrón resuelve el mismo problema que otro, en un contexto más específico y proporcionando una solución más concreta.
Extensión	Cuando un patrón mejora la solución de otro. El patrón que se está mejorando debe existir antes que el otro (hay una fuerte dependencia entre ellos).
Requerimiento	Cuando el contexto de un patrón requiere el cumplimiento de una condición que podría cumplirse con la aplicación de otro patrón.
Similitud	Cuando hay dos patrones con problemas o soluciones similares.
Uso	Cuando la solución de un patrón utiliza la de otro dentro de su procedimiento para resolver su propio problema.

El tipo de relación **alternativa**, alineado con la definición en [Meszaros97] puede utilizarse cuando dos patrones presentan soluciones diferentes, exclusivas o no, para resolver el mismo problema en contextos superpuestos o iguales. Este tipo de relación es simétrico; es decir, si X es una **alternativa** a Y, entonces, Y es una **alternativa** a X. El catálogo en [PRIPARE15] presenta casos de aplicación de este tipo de relación; por ejemplo, el patrón *Ofuscación de Medición con Ruido Agregado* es una **alternativa** al patrón *Complemento de Privacidad Confiable*. Ambos patrones comparten el mismo contexto (un proveedor de servicios que obtiene mediciones continuas de algún atributo del servicio brindado vinculadas a datos de identificación del usuario) y el mismo problema (la posibilidad de revelar demasiada información del usuario cuando las mediciones se repiten durante mucho tiempo). La *Ofuscación de Medición con Ruido Agregado*, se basa en la ofuscación de los datos a través de ruido en el aparato de medición antes de su transmisión. El *Complemento de Privacidad Confiable*, por otro lado, hace uso de un dispositivo intermedio confiable controlado por el usuario para el procesamiento, y se basa en la agregación de los datos.

¹² Para facilitar la lectura, se usa la palabra conexión para referirse a una relación encontrada en los catálogos individuales.

Tabla 6.2: Descripción breve de los patrones de privacidad usados en los ejemplos

Patrón	Breve descripción
<i>Ofuscación de Medición con Ruido Agregado</i> [PRIPARE15]	Propone la adición de ruido a los datos para ofuscarlos (manteniendo la validez de los datos a largo plazo) para contrarrestar la divulgación de información en transmisiones recurrentes vinculadas a datos personales.
<i>Complemento de Privacidad Confiable</i> [PRIPARE15]	Propone agregar las mediciones de un consumidor durante un lapso de tiempo previo a enviarlos al sistema de facturación para evitar divulgar información en transmisiones recurrentes vinculadas a datos personales.
<i>Panel de Concienciación de Privacidad en Espacios Colaborativos</i> [Fischer-Hübner10]	Tiene por objeto hacer que un usuario esté más informado proporcionándole información sobre la audiencia de sus contribuciones, e incluso indicando cuando tienen información adicional sobre él.
<i>Opciones de Privacidad en Redes Sociales</i> [Fischer-Hübner10]	Proporciona al usuario el control sobre la audiencia de sus contribuciones.
<i>Cifrado en Capas</i> [Hafiz13]	Sugiere realizar un cifrado de paquetes de datos en capas dentro de una red mixta.
<i>Relleno de Longitud Constante</i> [Hafiz13]	Hace que el tamaño de todos los paquetes de datos sea el mismo. Para ello, se puede recurrir a utilizar el relleno para incrementar el tamaño del paquete.
<i>Enrutamiento por Lotes</i> [Hafiz13]	Consiste en recoger los paquetes entrantes en una red mixta y despacharlos como un lote.
<i>Espera Aleatoria</i> [Hafiz13]	Sugiere usar tiempos de espera aleatorios previo a enviar los datos.
<i>Visualización de Coincidencias de Políticas</i> [Fischer-Hübner10]	Usa piezas de rompecabezas para proveer información visual sobre la medida en la cual una política de privacidad coincide con la configuración de privacidad del usuario antes de la divulgación de sus datos personales. De este modo, proporciona al usuario una forma más precisa de observar las implicaciones de privacidad antes de lograr su consentimiento.
<i>Consentimiento Informado</i> [Fischer-Hübner10]	Propone el uso de algunos conceptos de interacción humano-máquina para obtener el consentimiento informado e inequívoco del usuario en situaciones de divulgación de datos personales.
<i>Visualización de Política de Privacidad</i> [PRIPARE15]	Tiene como objetivo proporcionar al usuario información sobre qué datos se están revelando, a quién y con qué propósitos, todo ello evitando molestarle con información innecesaria.
<i>Visualización Dinámica de la Política de Priv.</i> [Fischer-Hübner10]	Provee información adicional (sobre los datos que se revelan y las posibles consecuencias) en cajas al colocar el cursor sobre elementos gráficos.
<i>Gestión de Obligaciones</i> [PRIPARE15]	Gestiona el ciclo de vida de la información en entornos en los que los datos pueden ser manejados por múltiples partes, basándose en las preferencias individuales del usuario y en las políticas de la organización.
<i>Políticas Adhesivas</i> [PRIPARE15]	Propone adherir a los datos de usuario, políticas (legibles por una máquina) que definen el uso y las obligaciones sobre los mismos.
<i>Pasarela de Agregación</i> [PRIPARE15]	Lista tareas para intercambiar datos sin revelar más información de la necesaria: cifrar las mediciones, transmitir los datos cifrados a un tercero de confianza, agregar un grupo de mediciones y enviar los datos.
<i>Redacción Consciente sobre Privacidad</i> [Fischer-Hübner10]	Establece el uso de frases fáciles de comprender para expresar los términos relacionados a la privacidad.
<i>Íconos de Privacidad</i> [Fischer-Hübner10]	Establece el uso de íconos, en línea con un propósito determinado, en el área de la seguridad y privacidad.
<i>Íconos para Política de Privacidad</i>	Propone el uso de íconos para mostrar cómo los datos son recolectados, procesados, almacenados, y con qué propósitos.

El tipo de relación **complemento**, está basado en las definiciones previas en [Noble98] (combina) y en [Meszaros97] (complementado por), y expresa que dos patrones pueden trabajar juntos (complementándose entre sí) para resolver un problema mayor no abordado por ninguno de los patrones individualmente. No hay necesidad de que los dos patrones tengan que implementarse juntos debido a que cada uno puede resolver su propio problema de manera individual. Sin embargo, al usar ambos patrones, la solución de cada patrón podría mejorarse en el mismo u otro contexto. Este tipo de relación es simétrico. Se han encontrado instancias de este tipo de relación en todos los catálogos considerados. Por ejemplo, en [Fischer-Hübner10] El patrón *Panel de Concienciación de Privacidad en Espacios de Trabajo Colaborativos* **complementa** al patrón *Opciones de Privacidad en Redes Sociales*. El primero tiene como objetivo informar al usuario sobre qué audiencia tienen sus contribuciones, mientras el patrón *Opciones de Privacidad en Redes Sociales* brinda al usuario control para definir la audiencia para una publicación. Ambos patrones, a la vez que resuelven sus propios problemas, pueden trabajar juntos para mejorar el diseño de una aplicación en un entorno colaborativo.

El tipo de relación **conducción**, acorde a la definición en [Meszaros97], se utiliza cuando la aplicación de un patrón (para resolver un problema inicial) conduce a un nuevo problema, que podría ser resuelto por otro patrón. El patrón inicial y aquel para resolver el nuevo problema deberían aplicarse conjuntamente para proveer una solución completa sin dejar problemas por resolver. Este tipo de relación es asimétrico (si X **conduce a** Y, no es posible que Y **conduce a** X). Hay ejemplos de este tipo de relación en varios de los catálogos [PRIPARE15][Hafiz13] [Colesky16]. Por ejemplo, en [Hafiz13], el patrón *Cifrado en Capas* **conduce a** un problema que podría ser resuelto por el patrón *Relleno de Longitud Constante*. El *Cifrado en Capas* usa criptografía de clave asimétrica y simétrica, lo que produce una variación en el tamaño de los datos, lo que a su vez puede revelar información a terceros. El *Relleno de Longitud Constante* aborda este nuevo problema haciendo que el tamaño de los datos sea el mismo.

El tipo de relación **conflicto** se aplica cuando dos patrones tienen soluciones que no pueden ser implementadas juntas, independientemente de sus problemas y contextos. Esta definición es diferente a la propuesta en [Noble98], cuya definición se ajusta al tipo alternativa en nuestra taxonomía propuesta. Este tipo de relación es simétrica. Hay casos de este tipo de relación en los catálogos propuestos en [PRIPARE15] y [Hafiz13]. Por ejemplo, en [Hafiz13], el patrón *Enrutamiento por Lotes* está en **conflicto** con el patrón *Espera Aleatoria*, ya que el primero tiene su fundamento en esperar hasta que se complete un lote de paquetes, mientras que *Espera Aleatoria* no puede esperar por lotes completos ya que se basa en tiempos aleatorios para enviar los datos.

El tipo de relación **especificación**, en línea con [Meszaros97] y [Noble98], se aplica cuando un patrón resuelve el mismo problema que otro, en un contexto más afinado y proporcionando una solución más específica. El patrón que se está especificando puede ser más general, simple o abstracto [Noble98]. Este tipo de relación es asimétrico y transitivo (es decir, si Y especifica a X y a su vez Z especifica a Y, entonces Z especifica a X). Hemos encontrado ejemplos de este tipo de relación en los catálogos reportados en [Fischer-Hübner10] y [PRIPARE15]. Por ejemplo, en [Fischer-Hübner10] encontramos que el patrón *Visualización de Coincidencias de Políticas* **especifica** el patrón *Consentimiento Informado*. El patrón *Consentimiento Informado* es genérico y tiene por objeto obtener el consentimiento informado e inequívoco del usuario en eventos asociados a la divulgación de datos personales, a través de medios basados en la interacción humano-máquina. Por su parte, el patrón *Visualización de Coincidencias de Políticas* ofrece una solución más específica, visual y basada en una metáfora de

rompecabezas. Esta se usa en situaciones en las que un usuario ha definido sus preferencias de privacidad y desea saber hasta qué punto una política de privacidad se adhiere a dichas preferencias antes de revelar sus datos personales.

El tipo de relación **extensión** se utiliza cuando la solución de un patrón mejora (extiende) la solución de otro, yendo más allá de usarla. La existencia del patrón que se está mejorando es necesaria para la definición y aplicación del otro. Esto no funciona a la inversa, ya que el patrón extendido podría definirse y aplicarse solo. Este tipo de relación es asimétrico y transitivo. Hay ejemplos de este tipo de relación en varios catálogos [Fischer-Hübner10][Hafiz13][Colesky16]. Por ejemplo, el patrón *Visualización Dinámica de la Política de Privacidad* **extiende** al patrón *Visualización de Política de Privacidad*, ya que proporciona información adicional a su solución utilizando información visual extra.

El tipo de relación **requerimiento**, en consonancia con la definición en [Noble98], se utiliza cuando antes de aplicar un patrón, se necesita cumplir alguna condición en su contexto. El cumplimiento de esa condición se puede lograr al aplicar otro patrón. Este tipo de relación no es el inverso del tipo de relación **conducción**. Este tipo de relación es asimétrica y transitiva. Un ejemplo de este tipo de relación se puede encontrar en [PRIPARE15] cuando el patrón *Gestión de Obligaciones* **requiere** que se implemente el patrón *Políticas Adhesivas* previamente. El patrón *Gestión de Obligaciones* controla la administración del ciclo de vida de la información en función de las preferencias individuales y políticas de una organización. Por ello se **requiere** establecer previamente las preferencias individuales, y para ello se puede usar el patrón *Políticas Adhesivas*.

El tipo de relación **similitud**, basado en la definición provista en [Noble98], contiene los casos en los que ambos patrones abordan problemas similares o proveen soluciones similares. Incluso, consideramos que este tipo de relación podría usarse para aquellos casos en los que los problemas son iguales, pero los tipos de relación **alternativa** o **especificación** no se ajustan. Este tipo de relación es simétrica. Existen ejemplos de este tipo de relación en los catálogos [Fischer-Hübner10][PRIPARE15][Colesky16]. Por ejemplo, en [PRIPARE15] tenemos que el patrón *Pasarela de Agregación* es **similar** al patrón *Complemento de Privacidad Confiable*, ya que ambos utilizan una solución basada en un dispositivo intermedio de confianza.

Finalmente, el tipo de relación **uso**, se ha inferido de [Meszaros97] y es compatible con la definición de [Noble98], se aplica cuando la solución de un patrón utiliza la de otro como parte del procedimiento seguido para resolver su problema. Los patrones no dependen del otro ya que cada uno resuelve su propio problema. Este tipo de relación es asimétrica y transitiva. Hemos encontrado ejemplos de este tipo de relación en todos los catálogos. Por ejemplo, en [Fischer-Hübner10] podemos afirmar que el patrón *Visualización de Política de Privacidad* podría **usar** el patrón *Redacción Consciente sobre Privacidad* para redactar usando términos comprensibles al proveer información sobre sus datos personales al usuario final.

6.3. Análisis de la taxonomía de tipos de relaciones

El mapeo entre la taxonomía finalmente propuesta y las conexiones establecidas en los catálogos individuales resulta revelador (Tabla 6.3). Por un lado, muestra que la taxonomía es aplicable a los catálogos individuales, por lo cual queda validada, y, por otro, expone las limitaciones de los esquemas de conexiones que actualmente tienen dichos catálogos.

En las siguientes secciones, se explica la validación de la taxonomía, cuyo proceso ha permitido desvelar las limitaciones encontradas en los esquemas usados en los catálogos individuales. Esto permite reflexionar sobre las ventajas y utilidad de la taxonomía propuesta de cara a ser usada por otros investigadores o para ayudar en el proceso de aplicación por parte de los desarrolladores. Finalmente, se presentan dos modos en los cuales la taxonomía podría extenderse.

6.3.1. Validación de la taxonomía propuesta en los catálogos individuales

La Tabla 6.3 muestra el resultado del proceso de mapeo entre la taxonomía final propuesta y los cuatro catálogos que expresan conexiones [Fischer-Hübner10][Hafiz13][PRIPARE15][Colesky16]. La primera columna muestra los tipos de relación propuestos en la taxonomía y las otras columnas corresponden a cada uno de los catálogos considerados, y contienen la expresión usada para definir una relación y el número de veces que ha aparecido.

Tabla 6.3: Mapeo de los tipos de relaciones propuestos a los catálogos existentes

Tipo de relación	Conexiones en:			
	[Fischer-Hübner10]	[PRIPARE15]	[Hafiz13]	[Colesky16]
Alternativa		Relacionado (3)		
Complemento	Relacionado (5)	De apoyo (8) Relacionado (3)	Se usa con (2) Sigue (1) Otras expresiones (2)	Relacionado (12)
Conducción		Relacionado (1)	Sigue (1)	Relacionado (1)
Conflicto		En conflicto (2)	Otras expresiones (1)	
Especificación	Relacionado (1)	Relacionado (1)		
Extensión	Relacionado (2)		Se usa con (1) Crea un X más fuerte (1) Crea un mejor (1) Otras expresiones (3)	De apoyo (1) Relacionado (3)
Requerimiento		Relacionado (1)		
Similitud	Relacionado (1)	Relacionado (1)		Relacionado (1)
Uso	Relacionado (3)	Relacionado (1) De apoyo (6)	Adopta (2) Sigue (1) Otras expresiones (2)	Relacionado (8)

Para poder determinar el tipo de relación más apropiado, se identificaron primero las conexiones ya expresadas dentro de los catálogos. Se tomó nota de la existencia de una relación entre dos patrones, y en aquellos casos en los que las conexiones eran más descriptivas (por ejemplo, proveyendo nombres como “es usado con”) también se apuntó la expresión usada, para entender el tipo de relación que podría existir. Posteriormente, se procedió a estudiar el problema, el contexto y la solución (en ocasiones las consecuencias) de los dos patrones relacionados para extraer posibles semejanzas o diferencias. Con esta información se procedió a localizar y seleccionar el tipo de relación más apropiado dentro de nuestra taxonomía. Por ejemplo, en [Fischer-Hübner10] el patrón *Panel de Concienciación de Privacidad en Espacios Colaborativos* lista al patrón *Opciones de Privacidad en Redes Sociales* como relacionado; mientras

que después de nuestro análisis hemos mapeado esta relación como a una del tipo **complemento**.

Este ejercicio ha permitido probar la validez de la taxonomía en cada uno de los catálogos. Se ha considerado como validado, cada tipo de relación instanciado al menos una vez en los referidos catálogos individuales.

Vale la pena notar que, al identificar las conexiones en los catálogos estudiados, fue necesario realizar un ajuste para representarlas. Por ejemplo, hay conexiones que se muestran únicamente al listar los nombres de los patrones en secciones tituladas como "patrones relacionados" [Fischer-Hübner10][PRIPARE15][Colesky16], "patrones de apoyo" [Colesky16][PRIPARE15] o "patrones en conflicto" [PRIPARE15]. Estos títulos de secciones se han utilizado como nombres de las conexiones "relacionado", "de apoyo" y "en conflicto" en la Tabla 6.3. Asimismo, el conjunto de patrones en [Hafiz 2013] no utiliza una forma uniforme para expresar las conexiones. En este caso, se han seleccionado aquellos términos que representan claramente una relación ("se usa con", "sigue", "crea un X más fuerte", etc.). Para los casos en los que hay expresiones muy largas para definir una relación, hemos usado el comodín "otras expresiones".

6.3.2. Limitaciones de los esquemas de relación actuales

El ejercicio de mapeo realizado desvela varios problemas que tienen los esquemas para definir las relaciones en los catálogos individuales, siendo en ocasiones muy laxos, no estando normalizados o provocando inconsistencias.

Los catálogos en [Fischer-Hübner10] y [Colesky16] usan un esquema muy laxo que carece de la suficiente especificidad para saber cómo se relacionan los patrones. Ahí únicamente se usa una sección titulada "patrones relacionados" para listar los nombres de los patrones que tienen cierta relación. Las conexiones definidas en dicha sección dentro de ambos catálogos se han mapeado a cinco de los tipos de relaciones de la taxonomía propuesta. Esto implica que el emplear una sección de "patrones relacionados" es una forma muy general de representación, lo cual limita la usabilidad de estos catálogos. Por ejemplo, cuando dos o más patrones relacionados podrían ser candidatos a usarse, se desconocería las implicaciones particulares de cada caso.

El catálogo presentado en [PRIPARE15] utiliza las secciones "patrones relacionados", "patrones de apoyo" y "patrones en conflicto" para mostrar las relaciones. Sin embargo, las dos primeras aún no son lo suficientemente específicas y definidas e incluso pueden ser inconsistentes. Por ejemplo, las conexiones en la sección "patrones relacionados" pueden expresarse utilizando siete tipos diferentes de relaciones de nuestra taxonomía, lo cual indica que no son lo suficientemente específicas. Asimismo, las conexiones en las secciones "patrones relacionados" y "patrones de apoyo" podrían representarse como relaciones del tipo complemento o de uso, lo cual indica inconsistencia.

Finalmente, el esquema para relacionar utilizado en [Hafiz13] da cabida para que existan inconsistencias, ya que ha habido casos en los que el autor utiliza más de tres expresiones diferentes para referirse a un mismo tipo de relación.

6.3.3. Utilidad de la taxonomía

La taxonomía propuesta proporciona un esquema específico y uniforme para representar las relaciones entre los patrones que solventaría los defectos actuales, y que puede ser usado tanto en los catálogos individuales como en el común.

Además, la taxonomía propuesta será útil tanto para los autores de patrones de privacidad como para los desarrolladores que los utilicen. Los autores podrían definir

más fácilmente las relaciones entre los patrones del catálogo común que se está conformando al disponer de la taxonomía funcionando como una caja de herramientas.

Por su parte, los desarrolladores podrían recibir soporte al proceso de aplicación de los patrones de privacidad. La taxonomía ayudaría a entender cómo se relacionan e interactúan los patrones, lo cual ayudaría al entendimiento individual y colectivo a través del contexto grupal. Asimismo, los desarrolladores podrían seguir las relaciones entre los patrones, de una manera más consciente y argumentada, para encontrar el patrón o patrones más adecuados para resolver problemas más complejos o para proponer mejores soluciones. Incluso, el usar una codificación apropiada para estas relaciones podría ayudar a una futura explotación automatizada del catálogo común de patrones a través de herramientas de navegación.

Por otro lado, el uso de los tipos de relaciones que proponemos puede complementar los esquemas de clasificación [Drozd16a][Hoepman14] usados actualmente, para proporcionar un esquema de organización más completo. Esto se debe a que los esquemas de clasificación pueden ser útiles para identificar un conjunto de patrones que pertenecen a una categoría, y las relaciones establecidas brindan información más específica sobre las conexiones entre sus patrones. Por ejemplo, la Figura 6.2 muestra un conjunto de patrones (y sus relaciones) que pertenecen al principio de privacidad "Consentimiento y Elección" según la clasificación propuesta en [Drozd16a], y considerando las descripciones hechas en la propuesta de [Fischer-Hübner10]. El contar con esquemas normalizados para establecer las relaciones y para clasificar los patrones puede dar cabida al establecimiento de técnicas y herramientas que aprovechen sus ventajas y permitan localizar y seleccionar los patrones de privacidad más adecuados para un problema en particular.

Finalmente, y muy importante, la taxonomía propuesta ha hecho que sea posible interconectar los patrones que han sido integrados dentro del catálogo común en *privacypatterns.org* [Doty11b], y así, ha permitido su avance hacia dos sistemas de patrones de privacidad (para informar y controlar el procesamiento de datos personales) que se abordan en el siguiente capítulo.

Figura 6.2: Ejemplo de aplicación de la taxonomía de tipos de relaciones

6.3.4. Extensiones futuras

La taxonomía propuesta puede ser extendida a través de las definiciones de características que tienen los diferentes tipos de relaciones propuestos como a través de la inclusión de nuevos tipos de relaciones.

En el primer caso, mientras se realizaba esta contribución, se ha notado que los tipos de relaciones podrían tener un atributo que indique la obligación, prohibición o posibilidad de que el patrón de destino en una relación se aplique. Por ejemplo, cuando el patrón *Cifrado en Capas* **conduce** al patrón *Relleno de Longitud Constante* para resolver el problema de los tamaños diferentes de los datos, se obliga a que el segundo patrón sea también aplicado, ya que, de lo contrario se dejaría un problema no resuelto. Al contrario, en una relación del tipo **conflicto**, se prohíbe que los dos patrones en conflicto se apliquen a la vez. En un tercer caso, en la relación en la que el patrón *Visualización de Política de Privacidad* podría **usar** el patrón *Redacción Consciente sobre Privacidad* no se expresa ni obligatoriedad ni prohibición, sino una opción. Aún se requiere definir la naturaleza y valores de este posible atributo, pero se considera que podría tener mucha utilidad de cara a construir técnicas o herramientas para navegación automatizada dentro de los catálogos, sistemas o lenguajes de patrones. Una de las posibles opciones que podría ser útil para definir los valores es aplicar la lógica deóntica que considera conceptos como autorizado, prohibido, obligatorio o indiferente.

Respecto a la inclusión de nuevos tipos de relaciones, estos podrían aparecer a medida que se estudien más patrones de privacidad o catálogos de los mismos, o mientras se vaya mejorando la taxonomía. Incluso la propuesta de [Noble98] contiene más tipos de relaciones (variante, variante-que-usa, secuencia de elaboración, etc.) que aún no hemos encontrado en los patrones estudiados. Es necesario seguir investigando para poder afirmar si son aplicables o no a los patrones privacidad.

6.4. Resumen del capítulo y contribuciones originales

Este capítulo ha descrito un estudio que, guiado por una estrategia de triangulación, llega a componer una taxonomía de tipos de relaciones que permiten definir conexiones entre patrones de privacidad. Se ha descrito el procedimiento seguido hasta conformar la propuesta de taxonomía, se han explicado las definiciones de cada uno de los tipos de relaciones dentro de ella, y se ha validado la taxonomía a través de un mapeo entre los diferentes tipos de relaciones y las conexiones existentes en cuatro catálogos.

Las contribuciones principales de este estudio son:

- **Una taxonomía de tipos de relaciones para definir conexiones entre patrones de privacidad.** Esta taxonomía está compuesta por nueve tipos de relaciones: alternativa, complemento, conducción, conflicto, especificación, extensión, requerimiento, similitud y uso.
- Una validación de la idoneidad de la taxonomía a través del mapeo de los tipos de relaciones propuestos y las conexiones existentes en los catálogos de patrones de privacidad dispersos, que de momento son laxas e inconsistentes.

Esta taxonomía permite dar un paso importante en el esfuerzo que realizan los investigadores en el área de los patrones de privacidad para llegar a consolidar un catálogo común y evolucionarlo. Además, en línea con la propuesta del Capítulo 5, se espera que el establecimiento de las relaciones ayude a mejorar la comprensión de los patrones y actúe como medio para futuras contribuciones en el ámbito de la búsqueda y selección de patrones.

El establecimiento de relaciones (de manera normalizada al usar la taxonomía) ha permitido que un grupo numeroso de patrones almacenados en el repositorio común *privacypatterns.org* [Doty11b] evolucione a dos sistemas de patrones para informar y controlar el procesamiento de datos. Estos son abordados en el Capítulo 7.

La taxonomía producto de este estudio se ha publicado en el artículo “Organizing design patterns for privacy: A taxonomy of types of relationships” [Caiza17] que fue presentado en la conferencia “22nd European Conference on Pattern Languages of Programs - EuroPLOP ’17” (Ranking CORE B).

Parte de los resultados de este estudio se han publicado en el artículo “Towards organizing the growing knowledge on privacy engineering” [DelAlamo18] que es parte del libro “Privacy and Identity Management. The Smart Revolution” de las series IFIP Advances in Information and Communication Technology (SJR¹³: Q3).

Finalmente, este estudio se ha llevado a cabo en colaboración con investigadores del proyecto “TRUst-Enhancing certified Solutions for SEcurity and protection of Citizens’ rights in digital Europe” dentro del programa de investigación e innovación Horizonte 2020 de la Unión Europea.

---

¹³ Factor de impacto SCImago Journal Rank <https://www.scimagojr.com/>

Capítulo 7

Sistemas de patrones de privacidad

En los capítulos previos se ha manifestado la relevancia de investigar el proceso de aplicación de los patrones de privacidad de cara a probar sus características y beneficios. Lo cual busca brindar soporte a las propuestas y recomendaciones basadas en ellos (p.ej. métodos de ingeniería de privacidad) e impulsar su adopción por los desarrolladores. En el Capítulo 5 se ha demostrado que el proceso de aplicación de un patrón no se reduce a una instanciación aislada; sino que, esta depende un conjunto de elementos que vienen desde la definición de los patrones (p.ej. usar una plantilla común). Estos elementos han sido agrupados en diferentes categorías (aspectos de aplicación), y ubicados a lo largo del ciclo de vida de un patrón.

Dentro de los elementos que más impacto positivo podrían causar se encuentran el establecimiento de conexiones entre los diferentes patrones de privacidad y la elaboración de sistemas de patrones de privacidad. Ambos pertenecientes al aspecto de organización, de acuerdo al marco de referencia para la aplicación de patrones descrito en el Capítulo 5. El Capítulo 6 ha presentado una taxonomía de nueve tipos de relaciones como un elemento que facilita y normaliza el establecimiento de conexiones entre los patrones de privacidad. Además, se ha validado su aplicabilidad a varios conjuntos de patrones de privacidad conexiones [Fischer-Hübner10][Hafiz13][PRIPARE15][Colesky16] que existen en la literatura. De este modo, esa contribución supone una ayuda importante para cualquier investigador o autor de patrones de privacidad; en especial para el grupo de investigadores que estamos conformando un repositorio común en *privacypatterns.org* [Doty11b].

Los patrones del repositorio común tienen ciertas características (p.ej. un esquema de clasificación común, estar almacenados en un mismo lugar) que les han hecho alcanzar el nivel de catálogo acorde a criterios ampliamente reconocidos [Buschmann96]. Sin embargo, aún hace falta establecer relaciones consistentes y descriptivas entre todos sus patrones y definir guías de instanciación. Estas carencias impiden la evolución del catálogo hacia ser un sistema de patrones que permita resolver problemas más complejos de diseño a través de la provisión de un grupo de patrones apropiados.

En este estudio se solventan las carencias de un grupo de 51 patrones del repositorio común en *privacypatterns.org* para llegar a conformar **dos sistemas de patrones de privacidad**: uno (con 31 patrones) para informar al usuario sobre el procesamiento de sus datos personales (en adelante **sistema para informar**) y otro para permitirle controlar dicho procesamiento (en adelante **sistema para controlar**). Para ello se estudia el repositorio común, se identifican las carencias usando los requisitos expresados en [Buschmann96], y se las solventan. Como parte fundamental del proceso de creación de estos, se aplica la taxonomía de tipos de relaciones propuesta previamente en el Capítulo 6.

La estructura de este capítulo se ha organizado de modo que en la Sección 1 se describe la metodología seguida. En la Sección 2 se presenta el sistema para informar y en la Sección 3 el sistema para controlar. A continuación, en la Sección 4 se explica el cumplimiento de los requisitos para llegar a ser sistemas y en la Sección 5 se analizan las implicaciones de los sistemas en la aplicación de los patrones de privacidad. Finalmente, la Sección 6 expone el resumen y las contribuciones originales.

7.1. Metodología

El proceso seguido (Figura 7.1) busca el cumplimiento de los requerimientos [Buschmann96] que un conjunto de patrones debería satisfacer para ser considerado como un sistema. En una primera etapa se estudian los patrones del repositorio común en *privacypatterns.org* [Doty11b] para determinar los requisitos que ya se cumplen y aquellos que aún deben ser alcanzados. A continuación, se hacen las modificaciones adecuadas para cada uno de los patrones de los dos sistemas a construir. Un componente principal en esta etapa es el establecimiento de las relaciones entre los patrones de privacidad haciendo uso de la taxonomía de tipos de relaciones (propuesta en el Capítulo 6).

Figura 7.1: Etapas en el desarrollo de los sistemas

En [Buschmann96] se expone un conjunto de seis requisitos que un conjunto de patrones debe cumplir para llegar a un nivel de sistema.

- R1. Este debe comprender una **base suficiente** de patrones que permitan definir la estructura fundamental de un sistema de información, que ayuden a refinar dicha estructura y también a implementarla.
- R2. Este debe presentar todos sus patrones de manera **uniforme**, expresando de manera clara su esencia y detalles, y permitiendo su comparación.
- R3. Este debe definir **guías de implementación** (instanciación) de los patrones de cara a construir un sistema de información.
- R4. Este debe tener un **esquema de organización** para todos los patrones que lo constituyen de tal manera que permita localizar rápidamente un patrón u otros relacionados dentro de una misma categoría.
- R5. Este debe exponer todas las **interconexiones** que puede haber entre los patrones que lo constituyen.
- R6. Este debe soportar su propia **evolución**, permitiendo que los patrones puedan actualizarse, eliminarse o se puedan agregar nuevos según corresponda.

7.2. Conformación de los requisitos para ser sistemas

Las características actuales de los patrones de privacidad en el repositorio común hacen que puedan ser considerados en conjunto como un catálogo, y, aunque se han realizado otras mejoras, aún hacen falta características importantes que les permitan evolucionar para llegar al nivel de sistema.

A continuación, se explica en primer lugar la conformación de aquellos requisitos que permiten al conjunto de patrones alcanzar el nivel de catálogo; para posteriormente, indicar los cambios realizados como parte de la conformación de requisitos para alcanzar el nivel de sistema.

7.2.1. Requisitos existentes (alcanzan para ser un catálogo)

Las características que les permiten ser considerados como un catálogo, sin aún alcanzar el nivel de sistema, incluyen el reunir una cantidad considerable de patrones descritos usando una plantilla común, clasificados bajo una misma forma de categorización basada en estrategias de diseño [Hoepman14], codificados usando un mismo esquema (Markdown¹⁴) y almacenados en el mismo repositorio [Doty11b].

Base suficiente (R1)

Previo a la realización de este estudio el repositorio contaba con 32 patrones de privacidad publicados y un grupo de propuestas de patrones pendientes que totalizaban 86. Actualmente, 72 patrones de privacidad ya están publicados, incluyendo todos aquellos que conforman los dos sistemas propuestos. Estos patrones provienen de otros catálogos como [Fischer-Hübner10][Hafiz13][PRIPARE15] (Capítulo 2).

Del total de patrones, 31 pertenecían a la estrategia Informar, 20 a la estrategia Controlar, y el resto estaban distribuidos entre las otras 6 estrategias. Las otras estrategias tenían cantidades bastante variables de patrones, yendo desde 1 en la estrategia Abstraer hasta 11 en la de Ocultar. De esta última, la mayoría de patrones ya pertenecían al lenguaje propuesto en [Hafiz13]. Por ello, se escogieron las categorías Informar y Controlar para elaborar sus sistemas de patrones. Además, la cantidad de patrones en ambas estrategias está en línea con las de otras propuestas de sistemas; por ejemplo, 19 en un sistema sobre arquitectura de software [Buschmann96] y 15 en uno sobre arquitecturas seguras [Preschern15].

Descripción uniforme (R2)

Al inicio de este estudio el repositorio común contaba con una plantilla común donde se especificaban campos obligatorios y opcionales que un patrón a ser incluido debía tener.

La plantilla común busca normalizar la estructura de los patrones y define un grupo de campos que cada patrón obligatoriamente debería tener (*nombre, contexto, problema, solución, consecuencias y ejemplos*) y un grupo que son opcionales (*resumen, otros nombres para el patrón, estructura de la solución, detalles de implementación, limitaciones, usos conocidos, patrones relacionados, y un campo para colocar otra información a tener en cuenta*).

Esta plantilla ha sido actualizada de cara a favorecer la implementación de los patrones, esto se explica en detalle en la Sección 7.2.2.

Esquema de organización (R4)

Los patrones del repositorio común están organizados a través de un esquema de clasificación que usa en un primer nivel las estrategias de diseño de privacidad propuestas en [Hoepman14]. En un siguiente nivel de categorización, los patrones están clasificados acorde a las tácticas que se han propuesto dentro de cada estrategia [Colesky16][Hoepman18].

---

¹⁴ Lenguaje de marcado ligero que fomenta la legibilidad (<https://markdown.es/>).

La estrategia *Informar* incluye las tácticas que permiten *proveer* (información de qué datos personales, cómo y por qué se procesan), *explicar* (en profundidad o argumentar por qué se procesan los datos personales) y *notificar* (a los usuarios en tiempo real cuando hay alguna acción o procesamiento en sus datos personales) a los interesados sobre detalles del procesamiento de datos que se va a llevar o se está llevando a cabo.

La estrategia *Controlar* comprende las tácticas que permiten al interesado *consentir* (explícitamente el procesamiento de los datos personales), *elegir* (los datos personales que se van a proveer), *actualizar* (los datos personales recolectados) y *retirar* (o eliminar) sus datos personales de algún servicio o producto que los está utilizando.

La Figura 7.2 y Figura 7.3 muestran los patrones categorizados por las tácticas dentro de sus respectivas estrategias de diseño a través de diferentes colores.

Evolución (R6)

Ambos sistemas de patrones están publicados en el repositorio común en *privacypatterns.org* [Doty11b], cuyos archivos fuente están disponibles en GitHub [Doty11a]. Esto hace posible que sean accedidos y modificados desde cualquier lugar. El esfuerzo común de investigadores está abierto a nuevas contribuciones a los sistemas, que deben ser realizadas a través de GitHub, para posteriormente, ser revisadas por los administradores del repositorio. Esto brinda la posibilidad de poder agregar, actualizar, eliminar, y, por lo tanto, evolucionar los patrones de privacidad propuestos en estos sistemas.

La evolución no solo implica aumentar o cambiar los patrones; también puede implicar actualizaciones en los modelos conceptuales detrás de los patrones; por ejemplo, cambios en las estrategias, tácticas o en los tipos de relaciones usados. Esto es perfectamente factible, ya que no son esquemas cerrados y se podrían incluir nuevas categorías dentro de cada una de ellas.

7.2.2. Requisitos cumplimentados para alcanzar el nivel de sistema

Las características previas no son suficientes para que el conjunto de patrones alcance el nivel de sistema. Aún se requiere que se solventen carencias en torno a la implementación y a las relaciones entre los patrones.

Respecto a la implementación, varios patrones no llegan a definir una guía de cómo llevarla a cabo, o resultan ser incipientes, o se tienen diferentes niveles de abstracción cuando las expresan. El que la plantilla común defina el campo de *implementación* como opcional tampoco ayuda a que siempre se coloque información sobre la instanciación de los patrones.

Respecto a las relaciones, dependiendo del catálogo inicial al cual pertenecían, los patrones ya presentaban varios defectos como la ausencia de relaciones, inconsistencias entre las relaciones establecidas y descripciones laxas de las mismas (analizadas en el Capítulo 6). Estos se acrecientan debido a la heterogeneidad entre los diferentes catálogos integrados.

A continuación, se detallan las medidas adoptadas para cumplir con los requerimientos para llegar al nivel de sistema: se actualiza la plantilla común, se definen guías de instanciación fundamentales y se establecen exhaustivamente relaciones claras, específicas y consistentes usando la taxonomía propuesta en el Capítulo 6.

Descripción uniforme (R2)

La plantilla existente no brindaba las facilidades suficientes para cumplimentar todos los requisitos para llegar al nivel de sistema. Por ello, se han propuesto cambios que permiten soportar completamente los requisitos, que han sido aceptados y ya están publicados en el repositorio común.

Los cambios en la nueva plantilla incluyen la adición de un campo obligatorio llamado *fuerzas*, que describe las dificultades que presenta el contexto para resolver un problema. Además, cambiamos el nivel de opcionalidad de los campos que brindaban detalles de *implementación* y de los *patrones relacionados*, de ser opcionales a ser obligatorios. Estos cambios hacen que aquellos patrones que sigan la plantilla establecida definan guías de instanciación y las interconexiones que tengan con otros patrones; ayudando así a cumplir con los requisitos para llegar a ser un sistema. Durante el desarrollo de este estudio, solo se ha tratado de modificar o agregar los campos mencionados y no se ha eliminado ninguno de cara a evitar distorsionar la información que contenían los patrones ya publicados en el catálogo. La Tabla 7.1 muestra la plantilla común actualizada usada en los sistemas de patrones y que puede ser accedida a través de GitHub [Doty11a].

Tabla 7.1 Plantilla común usada para los patrones dentro de los sistemas

Tipo de campo	Descripción
Campos obligatorios	Nombre [descriptivo del patrón]
	Contexto [en el cual se desarrolla el patrón]
	Problema [que se soluciona]
	Fuerzas [que hacen que el problema sea difícil de solucionar]
	Solución [al problema]
	Implementación [de la solución]
	Patrones relacionados [a este patrón]
Campos opcionales	Resumen [del patrón]
	También conocido como [otros nombres del patrón]
	Estructura [de la solución del patrón]
	Ejemplos [motivacional de cómo implementar el patrón]
	Usos conocidos [del patrón en la práctica real]
	Consecuencias [al aplicar la solución]
	Limitaciones [tras la implementación del patrón]
	Mirar también [otra información no descrita previamente]

Guías de implementación (R3)

Para cumplir con este requisito, se brinda información fundamental de cómo cada uno de los patrones debe ser implementado. Para tal fin, se han hecho cambios en la plantilla y se ha completado la información faltante.

Respecto a la plantilla, se ha cambiado el campo *implementación* a obligatorio para garantizar que existan las guías fundamentales de implementación. Además, en línea con [Buschmann96] y [Meszaros97], dado que la presencia de otros campos también favorecería la implementación de un patrón, se han hecho obligatorios los campos

fuerzas y patrones relacionados (la Sección 2.5 provee información sobre los patrones y sus campos). El campo *fuerzas* ayuda a que el desarrollador entienda el contexto, problema y las dificultades de resolución para que pueda comprender bien la situación. Asimismo, el campo *patrones relacionados* ayuda a obtener un entendimiento contextual en función de los tipos de relaciones que tiene con otros patrones.

Además de estos campos obligatorios, la existencia de otros opcionales como el de *ejemplos* y el de *usos conocidos* pueden ilustrar la implementación. Así también los campos *estructura*, *consecuencias* y *limitaciones* podrían brindar más detalles sobre la implementación y efectos (p.ej. la necesidad de aplicar otro patrón posterior) de la misma previo a tomar decisiones.

Teniendo en cuenta lo anterior, los patrones han sido revisados y modificados para garantizar que todos tengan los campos obligatorios, en aras de que se explique adecuadamente la situación del patrón y se indiquen guías fundamentales de implementación. También se han revisado los campos opcionales, para que estén en consonancia con el resto de campos y para evitar información redundante. En el caso de los patrones que ya cumplían con esas condiciones, no se realizan modificaciones. Vale la pena notar que las guías de implementación pueden mejorarse a medida que los patrones sigan evolucionando de manera individual o en su camino a conformar lenguajes.

Finalmente, como se concluyó en el Capítulo 5, el proceso de aplicación exitoso de un patrón no depende únicamente de las guías de instanciación, también se verá afectado por otros elementos (p.ej. uso esquemas de representación útiles).

Interconexiones (R5)

La definición de las interconexiones entre los patrones es uno de los requisitos esenciales que permiten a un catálogo evolucionar hacia un sistema [Buschmann96]. Su importancia radica en que pueden ayudar en la comprensión de los patrones, en el proceso de selección de aquellos más apropiados para resolver un problema en un determinado contexto, y en llegar a incluso poder resolver problemas complejos gracias al uso de grupos de patrones relacionados [Meszaros97].

Para establecer las diferentes interconexiones, se ha usado la taxonomía de tipos de relaciones propuesta en el Capítulo 6. En la versión actual de los patrones que conforman los sistemas propuestos, se han encontrado instancias de los tipos: *alternativa*, *complemento*, *especificación*, *similitud* y *uso*. A continuación, se colocan ejemplos de los tipos de relaciones encontrados.

- Alternativa (Dos patrones que resuelven el mismo problema). Es el único caso en el cual se ha encontrado solo una instancia; y esta se localiza en el sistema para *informar* entre los patrones **Aviso asíncrono** y **Aviso ambiental**. Ambos informan a los usuarios cuando hay un acceso a los datos personales; sin embargo, mientras **Aviso asíncrono** busca notificar a través de mensajes modales (inevitables) ante eventos específicos, **Aviso ambiental** busca hacerlo de manera seguida causando la mínima distracción. En este caso, no son excluyentes y ambos podrían ser implementados de manera conjunta.
- Complemento (Dos patrones que pueden trabajar conjuntamente). En el sistema para *controlar* se encuentra que el patrón **Desalentar estrategias generales** puede complementar al patrón **Negociación de la política de privacidad** durante un evento de compartición de datos. Al darse una compartición de datos,

por defecto puede cargarse la configuración de privacidad definida en el patrón **Negociación de la política de privacidad**, pero se puede mantener la posibilidad de que la persona pueda cambiar la configuración de manera granular, usando el patrón **Desalentar estrategias generales**, en ese instante.

- Especificación (Dos patrones que resuelven el mismo problema, pero el uno en un contexto más específico). En el sistema para *informar* se tiene que el patrón **Prevenir errores o reducir su impacto**, que busca prevenir la divulgación no intencional de datos o reaccionar ante ella, puede ser especificado por los patrones **Aviso asíncrono** y **Aviso ambiental**. Ambos patrones brindan soluciones específicas de prevención a través de mecanismos de notificación asíncronos o continuos.
- Similitud (Dos patrones similares en problema o solución). En el sistema para *informar* se tiene que los patrones **Espejos de privacidad** y **Panel de privacidad** son similares. Ambos establecen un lugar para proveer información respecto a los datos personales que son recolectados y generar cierta conciencia de ello. Mientras el patrón **Panel de privacidad** brinda más detalles técnicos, el patrón **Espejo de privacidad** está en un contexto socio-técnico más amplio.
- Uso (Un patrón que puede usar la solución de otra dentro de la suya). En el sistema para *controlar*, el patrón **Desalentar estrategias generales** puede usar el patrón **Lista de amigos** para definir con quienes se comparte la información de manera granular.

7.3. Sistema para informar sobre el procesamiento de datos

Legalmente, previo a procesar los datos personales de los interesados, se debe obtener una base legal que lo respalde. Por ejemplo, en el caso del RGPD una de dichas bases es el consentimiento por parte de los interesados¹⁵, Art. 6 del RGPD [EuroPC16]. El Artículo 4 (11) del RGPD define al consentimiento como "la manifestación de voluntad libre, específica, informada e inequívoca por la cual el interesado acepta, mediante una clara acción afirmativa, el tratamiento de sus datos personales". En este contexto, el consentimiento informado implica que el interesado conoce quien procesa, las actividades de procesamiento a realizar, el propósito del procesamiento, y que él puede retirar su consentimiento cuando desee. Ante esto, las organizaciones deben proveer dicha información, la cual debe ser descrita de manera inteligible, clara y fácil de entender.

La estrategia *Informar* provista en [Hoepman18] persigue “informar a los sujetos de datos acerca del procesamiento de sus datos personales de manera oportuna y adecuada”. De este modo se le brinda al interesado la posibilidad de tomar decisiones informadas respecto al procesamiento de sus datos.

En el catálogo común de patrones de privacidad hay 31 categorizados dentro de la estrategia de diseño *Informar*, los mismos que han sido utilizados para elaborar un sistema de patrones que permite diseñar sistemas de información con características para informar a los interesados sobre el procesamiento de sus datos personales. El sistema es mostrado en la Figura 7.2 y despliega los patrones con nombres descriptivos, categorizados acorde a las tácticas dentro de la estrategia *Informar*, e interconectados.

---

¹⁵ Individuo, persona, usuario que genera datos personales.

El recuadro en línea punteada denota la existencia de un patrón compuesto que puede usar los patrones en su interior para ofrecer una solución a un problema más amplio.

Los Anexos (Tabla A.3 – Tabla A.5) muestran un resumen breve de cada uno de los patrones de este sistema agrupados por tácticas. Además, el sistema se encuentra publicado en el repositorio común en *privacypatterns.org* [Doty11b].

Figura 7.2: Sistema de patrones de privacidad para la estrategia Informar

7.4. Sistema para controlar el procesamiento de datos

Únicamente informar a los interesados sobre el procesamiento de datos no es suficiente, este debe ir complementado con la posibilidad de controlar dicho procesamiento. Por ejemplo, el RGPD ha sido concebido precisamente para que los interesados tengan el control de sus datos personales [EuroC19]. Este control incluye la posibilidad de elegir qué datos se procesan, de rectificar los datos, de eliminar los datos, entre otros.

Acorde a [Hoepman18], la estrategia de diseño *Controlar* busca “proveer a los sujetos de datos el control adecuado sobre el procesamiento de sus datos personales”. Esta estrategia tiene en cuenta el hecho de que proteger la privacidad del interesado no

implica prohibir todo procesamiento de sus datos personales; sino, que la persona debe estar en capacidad de controlar y decidir lo que sucede con sus datos.

En el catálogo común de patrones de privacidad, hay 20 patrones dentro de la estrategia de diseño Controlar que han sido utilizados para estructurar un sistema de patrones que permite diseñar sistemas de información con las características apropiadas. El sistema es mostrado en la Figura 7.3 y muestra los patrones con nombres descriptivos, categorizados usando las tácticas, e interconectados. El recuadro en línea punteada denota la existencia de tres patrones compuestos que pueden usar los patrones en su interior para ofrecer soluciones a problemas más amplios.

Los Anexos (Tabla A.6 – Tabla A.9) muestran un resumen breve de cada uno de los patrones de este sistema agrupados por tácticas. Además, el sistema se encuentra publicado en el repositorio común [Doty11b].

Figura 7.3: Sistema de patrones de privacidad para la estrategia de control

7.5. Implicaciones en el proceso de aplicación

Un sistema de patrones permite mostrar fundamentalmente cómo se interconectan sus patrones constitutivos y cómo estos se podrían implementar [Buschmann96]. Adicionalmente, y considerando lo discutido en el Capítulo 5, sus características como el usar una plantilla común y tener un esquema común de categorización también contribuyen a lograr una instanciación exitosa. Esto hace que, por un lado, un sistema de patrones sea una de las mejoras que más impacto puede causar al proceso de

aplicación de los patrones; pero, por otro lado, aún hay algunos defectos en los patrones que deberían ser solventados.

7.5.1. Mejoras al proceso de aplicación

Las características que ahora tienen todos los patrones que pertenecen al sistema se traducen en un conjunto de ventajas para los desarrolladores que deseen aplicarlos. A continuación, las hemos agrupado según los aspectos de aplicación definidos en el Capítulo 5.

- **Estructura.** Ahora los desarrolladores cuentan con patrones que tienen una misma estructura (definida en una plantilla común) que uniformiza los campos que poseen los patrones y el contenido que deberían tener, y por lo tanto ayudan a solventar los problemas de heterogeneidad previamente existentes.
- **Organización y Selección.** Aquí es donde más impacto tienen las características logradas en los patrones de privacidad. El contar con un esquema común de categorización y un esquema de interconexión normalizado hace que el desarrollador disponga de un conjunto de patrones plenamente organizado. Debido al esquema de categorización utilizado (estrategias de diseño y tácticas), se hace factible el localizar directamente aquellos patrones que adheridos a una táctica determinada permiten llegar a cumplir con un objetivo plasmado en una estrategia, que en última instancia llega a satisfacer un principio de privacidad. Además, las relaciones hacen que se identifiquen aquellos otros patrones que podrían ayudar a mejorar una solución, o que podrían perjudicarla. Esto en conjunto hace posible localizar el grupo de patrones de privacidad más adecuado que pueda solucionar problemas de diseño más complejos. Finalmente, la categorización y las interconexiones entre los patrones harán que sea posible la navegación dentro del sistema, lo cual favorecería el desarrollo posterior de técnicas o de herramientas de selección.
- **Comprensión.** Las características en torno a la estructura común y los campos actualmente definidos pueden favorecer la comprensión de los patrones ya que, por un lado, ayudan a enfrentar la confusión generada por cuestiones de heterogeneidad, y por otro, el uso de campos como los ejemplos o usos conocidos pueden favorecer a entender a través del mirar “cómo se hace”. Asimismo, las relaciones entre los patrones ayudan a su comprensión a través del contexto que estas crean.
- **Almacenamiento.** La ventaja para el desarrollador es la de tener un solo lugar donde encontrar los patrones adecuados para dotar a un sistema de características que le permitan informar al interesado y darle control del procesamiento de sus datos. Actualmente, el sistema de patrones está integrado en el repositorio común y se presenta como una herramienta web interactiva que muestra los patrones y las conexiones entre ellos (Figura 7.4).
- **Instanciación.** El desarrollador cuenta con guías fundamentales para implementar la solución expresada en el patrón, las mismas que pueden seguir mejorando a través de la forma de colaboración que soporta el repositorio actualmente.

Por todo esto los sistemas constituyen una mejora muy relevante de cara a poder aplicar exitosamente los patrones de privacidad; sin embargo, aún quedan defectos por

solventar debido a ciertas limitaciones relacionadas con su definición misma y también por el estado actual de los patrones.

Patterns — Control

Figura 7.4: Repositorio común de patrones de privacidad

7.5.2. Limitaciones existentes y vías de mejora

Por definición, un sistema comprende un conjunto de patrones que ayudan a establecer la estructura fundamental de un sistema de información y no cubren cada aspecto de su desarrollo como idealmente lo haría un lenguaje de patrones [Buschmann96]. Por ello, diferente a lo que sucede con los lenguajes, los sistemas de patrones se enmarcan en dominios más abiertos. Los desarrolladores deben ser conscientes de que no se espera que se utilicen todos los patrones del sistema a la vez, sino que, deberían seleccionar los más apropiados [Buschmann96].

Además, el análisis realizado durante este estudio, especialmente durante el establecimiento de las relaciones, ha desvelado defectos y posibles vías de mejora en los patrones de privacidad del repositorio común. Estos defectos escapan al cumplimiento

de los requisitos para ser sistema que se abordan en este estudio. Estos están principalmente asociados a cuestiones internas en las descripciones de los patrones, que resultan ser en ocasiones muy abstractas o laxas, sobre o infra-descritas, y aún no usan medios de representación más concretos o familiares (en la mayoría de casos) que favorezcan a la comprensión de los desarrolladores.

Consideramos que los siguientes pasos deberían enfocarse en integrar aportes como un vocabulario común para describir cuestiones de privacidad como el propuesto en [Pandit19] y formas de representación gráficas y familiares para los desarrolladores como las que se están empezando a explorar en [Suphakul17] o las que se exploran para el área de elicitación de requisitos [Deng11]. El integrar ese tipo de contribuciones no es nada trivial, ya que se necesitaría, en primer lugar, analizar si se puede realizar una integración de ese tipo de medidas; posteriormente, la definición de los mecanismos para integrarlos; a continuación, las pruebas respectivas; para finalmente, realizar el cambio en los patrones del repositorio (más de 100). Sin embargo, sería un buen aporte de cara a facilitar la aplicación de los patrones de privacidad.

Finalmente, los cambios que se hagan sobre los patrones pueden hacer que sus relaciones cambien. Por ende, es necesario tener en cuenta que lo discutido en este estudio se aplica al estado de los patrones en el momento en el que fueron evolucionados del catálogo a los sistemas (2018 y 2019).

7.5.3. Integración en el proceso de diseño con requisitos de privacidad

Los sistemas presentados podrían ser integrados a los métodos que usan patrones de privacidad [Hoepman14][Kung14][Notario15] debido a que el agruparlos en un sistema no afecta la concepción de lo que es un patrón. En especial, los sistemas propuestos se adoptan mejor al enfoque de diseño propuesto en [Hoepman14] ya que ahí es donde se propone el uso de estrategias de diseño y de tácticas. Inicialmente, se debería identificar las estrategias y tácticas que mejor permitirían lograr un objetivo de diseño acorde a requerimientos de privacidad planteados. Posteriormente, y siempre y cuando las estrategias sean informar o controlar, se accedería a los sistemas propuestos. Dada su categorización por estrategias y tácticas se llegaría a obtener un grupo de patrones que podrían ser de utilidad si se ajustan a los requisitos planteados. De ser este el caso, a continuación, se podrían utilizar las relaciones para lograr estructurar una solución más completa a través de la inclusión de más patrones. Con un grupo apropiado de patrones, se establecería la solución de diseño a los requisitos inicialmente planteados.

7.6. Resumen del capítulo y contribuciones originales

Este capítulo propone dos sistemas de patrones de privacidad: uno que permite informar al interesado sobre el procesamiento de sus datos personales y otro que le posibilita controlar dicho procesamiento. Se describe el proceso seguido, fundamentalmente centrado en el establecimiento exhaustivo de relaciones y cumplimiento de un conjunto de requisitos para que un grupo de patrones alcance el nivel de sistema.

Las contribuciones directas de este estudio son:

- **Un sistema compuesto por 31 patrones de privacidad para informar al usuario sobre el procesamiento de sus datos personales.**
- **Un sistema compuesto por 20 patrones de privacidad que le permite al interesado controlar el procesamiento de sus datos personales.**

- Una validación de la taxonomía de tipos de relaciones propuesta en el Capítulo 6 a través del establecimiento de las conexiones entre los patrones de privacidad de los dos sistemas propuestos.
- La actualización de la plantilla acordada por los colaboradores del esfuerzo común de investigación en patrones de privacidad, mantenida en *privacypatterns.org* [Doty11b].

Ambos sistemas propuestos han sido integrados al repositorio que mantenemos los investigadores del esfuerzo común en patrones de privacidad [Doty11b]. Además, constituyen una mejora sustancial en la investigación sobre la aplicación de los patrones de privacidad. Las características del sistema hacen que se provean diferentes ayudas a los desarrolladores de cara a la aplicación de los patrones: facilitan la comprensión de los patrones, permiten localizar aquellos más apropiados de cara a resolver un problema de diseño, posibilitan el establecimiento de soluciones más completas, brindan guías fundamentales para la instanciación, entre otras; todo esto a través de un solo lugar de acceso.

El sistema de patrones para controlar ha sido publicado en el artículo “*A system of privacy patterns for user control*” [Colesky18] presentado en la conferencia “33rd Annual ACM Symposium on Applied Computing” (Ranking CORE B).

El sistema de patrones para informar al interesado ha sido publicado en el artículo “*A System of Privacy Patterns for Informing Users: Creating a Pattern System*” [Colesky19] que fue presentado en la conferencia “The 23rd European Conference on Pattern Languages of Programs - EuroPLoP ’18” (Ranking CORE B).

Este estudio se ha llevado a cabo en colaboración con investigadores del proyecto “*Patterns for Privacy - P4P*”¹⁶ dentro de la Agenda de Investigación de Ciberseguridad Nacional por la Organización de los Países Bajos para la Investigación Científica.

Así también, se ha colaborado con investigadores del proyecto “*TRUst-Enhancing certified Solutions for SEcurity and protection of Citizens’ rights in digital Europe*” dentro del programa de investigación e innovación Horizonte 2020 de la Unión Europea.

Finalmente, este estudio se llevó a cabo principalmente en una estancia de investigación en el centro de investigación “The Privacy & Identity Lab (PI.lab)” en Radboud University en los Países Bajos.

---

¹⁶ Patterns for Privacy - P4P <http://www.cs.ru.nl/~jhh/p4p/p4p.html>

Capítulo 8

Divulgación de resultados

Los resultados alcanzados en los distintos estudios llevados a cabo y detallados en los capítulos precedentes han sido utilizados para elaborar publicaciones aceptadas en revistas con factor de impacto Web of Science Journal Citation Reports¹⁷ (JCR) y SCImago Journal Rank¹⁸ (SJR) de Scopus, y en congresos internacionales de relevancia acorde al ranking Computing Research and Education Association¹⁹ (CORE).

Varios de los estudios se han llevado a cabo en colaboraciones con otros investigadores trabajando en el área de la privacidad, en ocasiones como parte de estancias de investigación internacionales. También se han dado colaboraciones a través de supervisiones y asesorías de trabajos fin de máster y fin de grado. Finalmente, los resultados y conocimientos adquiridos han sido utilizados para llevar a cabo presentaciones divulgativas sobre el área de investigación y cursos de capacitación en torno al diseño de sistemas con requisitos de privacidad usando patrones y sobre los procedimientos metodológicos aprendidos a lo largo de la realización de esta tesis. Todas las contribuciones pueden ser apreciadas en la Figura 8.1.

Figura 8.1: Mapa de contribuciones

¹⁷ <https://jcr.clarivate.com/>

¹⁸ <https://www.scimagojr.com/>

¹⁹ <http://portal.core.edu.au/conf-ranks/>

8.1. Diseminación de resultados directos de la tesis

8.1.1. Publicaciones en revistas

- JCR1 Caiza, J. C., Martin, Y.-S., Guaman, D. S., Del Alamo, J. M., & Yelmo, J. C. (2019). Reusable Elements for the Systematic Design of Privacy-Friendly Information Systems: A Mapping Study. *IEEE Access*, 7, 66512–66535. <https://doi.org/10.1109/ACCESS.2019.2918003>
JCR: 3,745 (Q1 - 2019)

El artículo reporta las contribuciones del Capítulo 3: el mapa de elementos reusables para el diseño de sistemas respetuosos con la privacidad, incluyendo su madurez y tendencias, y el estado actual de la investigación en el área de los patrones de privacidad.

Siendo el principal contribuyente del artículo, se aportó a la definición del estudio, su planificación y ejecución, al análisis y construcción del mapa de elementos reusables, y al análisis del estado actual del área de los patrones de privacidad.

- SJR1 Del Alamo, J. M., Martín, Y. S., & Caiza, J. C. (2018). Towards organizing the growing knowledge on privacy engineering. In *IFIP Advances in Information and Communication Technology* (Vol. 526, pp. 15–24). https://doi.org/10.1007/978-3-319-92925-5_2
SJR: 0,21 (Q3 - 2019)

El artículo contiene parte de las contribuciones del Capítulo 6: un ejemplo de aplicación de la taxonomía de tipos de relaciones entre patrones de privacidad.

Se contribuyó al artículo con la elaboración del ejemplo, la escritura de una sección del artículo y la revisión del mismo.

8.1.2. Publicaciones en congresos

- CORE1 Caiza, J. C., Del Alamo, J. M., & Guamán, D. S. (2020). A framework and roadmap for enhancing the application of privacy design patterns. *The 35th ACM/SIGAPP Symposium On Applied Computing*, 1297–1304. <https://doi.org/10.1145/3341105.3375768>
Ranking CORE B (2018)

El artículo reporta las contribuciones del Capítulo 5: el marco de referencia para la aplicación de patrones, la descripción de los defectos y mejoras en el ámbito de los patrones de privacidad, y el análisis de las mejoras más novedosas que mayor impacto causarían en el proceso de aplicación de los patrones de privacidad.

Siendo el principal contribuyente del artículo, se aportó con la definición del estudio, su planificación y ejecución, al análisis y construcción del marco de referencia, y al estudio y priorización de las mejoras en los patrones de privacidad.

- CORE2 Caiza, J. C., Martín, Y.-S., Del Alamo, J. M., & Guamán, D. S. (2017). Organizing design patterns for privacy: A taxonomy of types of relationships. *Proceedings of the 22nd European Conference on Pattern Languages of Programs - EuroPLoP '17*, 1–11. <https://doi.org/10.1145/3147704.3147739>
Ranking CORE B (2018).

El artículo reporta las contribuciones del Capítulo 6: la taxonomía de tipos de relaciones para los patrones de privacidad y un mapeo de la misma a catálogos de patrones de privacidad existentes.

Siendo el principal contribuyente del artículo, se aportó con la definición del estudio, el análisis y la construcción de la taxonomía, y la validación a través del mapeo a los catálogos existentes.

- CORE3 Colesky, M., Caiza, J. C., Del Álamo, J., Hoepman, J.-H., & Martín, Y.-S. (2018). A system of privacy patterns for user control. *33rd Annual ACM Symposium on Applied Computing*, 1150–1156. <https://doi.org/10.1145/3167132.3167257>
Ranking CORE B (2018)

El artículo reporta parte de las contribuciones del Capítulo 7: el sistema para permitirle al usuario controlar el procesamiento de sus datos personales y la actualización de la plantilla común para los patrones de privacidad.

Siendo uno de los contribuyentes principales, se aporta al diseño del estudio, a la edición y revisión de los patrones, y al establecimiento exhaustivo de las relaciones.

- CORE4 Colesky, M., & Caiza, J. C. (2019). A System of Privacy Patterns for Informing Users: Creating a Pattern System. *Proceedings of the 23rd European Conference on Pattern Languages of Programs - EuroPLoP '18*, 1–11. <https://doi.org/10.1145/3282308.3282325>
Ranking CORE B (2018)

El artículo reporta parte de las contribuciones del Capítulo 7: el sistema para informar al usuario sobre el procesamiento de sus datos personales.

Se contribuyó al artículo con el establecimiento de las interconexiones entre los patrones, y la revisión de los patrones y del artículo en general.

8.2. Publicaciones en el contexto de la tesis

8.2.1. Publicaciones en revistas

- JCR2 Guaman, D. S., Del Alamo, J. M., & Caiza, J. C. (2020). A Systematic Mapping Study on Software Quality Control Techniques for Assessing Privacy in Information Systems. *IEEE Access*, 8, 74808–74833. <https://doi.org/10.1109/access.2020.2988408>
JCR: 3,745 (Q1 - 2019)

El artículo reporta un estudio sistemático sobre las técnicas para llevar a cabo el control de la calidad del software en sistemas de información en el ámbito

de la privacidad. Se colaboró en la planificación y ejecución de este estudio a través de la aplicación de los conocimientos metodológicos obtenidos durante la realización del mapeo sistemático de literatura reportado en el Capítulo 3.

- SJR2 Guamán, D. S., del Alamo, J. M., Veljanova, H., Haselbacher, A., & Caiza, J. C. (2019). Ranking online services by the core areas of trustworthiness. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, 465–478. SJR: 0,15 (Q4 - 2019)

El artículo propone un marco para clasificar servicios en la web basados en nivel de fiabilidad y las preferencias de los usuarios. En este artículo se colaboró a través discusiones con los autores sobre cómo abordar la privacidad desde el punto de vista técnico y con la revisión del mismo.

8.3. Colaboración en proyectos de investigación

- COL1 **PDP4E²⁰ (2018 - 2021)**: Methods and tools for GDPR (General Data Protection Regulation) compliance through Privacy and Data Protection Engineering (Acuerdo No. 787034). Programa de investigación e innovación Horizonte 2020 de la Unión Europea. Este proyecto busca proporcionar a los ingenieros con métodos y herramientas para la creación de sistemas que cumplan con el RGPD, en particular en lo referente a la privacidad y protección de datos desde el diseño.

Se colaboró con el equipo UPM del proyecto en el análisis de los elementos de diseño (p.ej. métodos, modelos, herramientas) que se encuentran en el estado del arte y que han sido reportados en el Capítulo 3.

- COL2 **CLIP (2020 - 2021)**: Consolidación del Liderazgo en Investigación en Ingeniería de Privacidad. Programa V-PRICIT de Apoyo a la realización de Proyectos de I+D para jóvenes investigadores de la Comunidad de Madrid y la Universidad Politécnica de Madrid (Acuerdo APOYO-JOVENES-QINIM8-72-PKGQ0J). Este proyecto persigue la investigación en métodos, técnicas y herramientas para ayudar a los ingenieros software a lidiar con los problemas de privacidad y protección de datos e introducir soluciones adecuadas en los procesos de desarrollo de nuevos sistemas y servicios, en particular en las actividades de diseño y control de calidad.

En este marco, se ha colaborado en este proyecto para el avance de las actividades de diseño, al ser éste el objetivo principal de la tesis doctoral

- COL3 **TRUESSEC.eu²¹ (2017 - 2018)**: TRUst-Enhancing certified Solutions for SEcurity and protection of Citizens' rights in digital Europe (Acuerdo No. 731711). Programa de investigación e innovación Horizonte 2020 de la Unión Europea. Este proyecto abordó la certificación y etiquetado de las características de confiabilidad de productos y servicios digitales, entre las que se encontraban la privacidad y protección de datos.

---

²⁰ <https://cordis.europa.eu/project/id/787034>

²¹ <https://cordis.europa.eu/project/id/731711>

Se colaboró en las discusiones sobre cómo abordar los requisitos legales de protección de datos desde el punto de vista técnico, para cumplir con los requisitos de privacidad por diseño y por defecto, en línea con las contribuciones descritas.

COL4 P4P²² (2014 - 2018): Patterns for Privacy. Agenda de Investigación de Ciberseguridad Nacional por la Organización de los Países Bajos para la Investigación Científica (NWO/STW National Cyber Security Research Agenda –NCSRA II– programme). Este proyecto buscaba proveer a los diseñadores de sistemas un conjunto de herramientas que puedan aplicarlas en el diseño respetuoso con la privacidad. Principalmente se enfocaban en definir un catálogo extenso de patrones y herramientas basadas en los patrones para soportar el proceso de desarrollo de los sistemas.

Se colaboró para establecer una sinergia entre los logros de esta tesis con los llevados a cabo en el proyecto P4P para conseguir la evolución del catálogo de patrones hacia los dos sistemas reportados en el Capítulo 7.

8.4. Estancias de investigación

EST1 Escuela Politécnica Nacional

Esta estancia se llevó a cabo en las instalaciones de la Escuela Politécnica Nacional donde se planificó, preparó y ejecutó el experimento exploratorio reportado en el Capítulo 4. Además, se llevaron a cabo varios eventos de divulgación a través de un curso de capacitación (DIV4) sobre el diseño de sistemas con requisitos de privacidad, y de conferencias (DIV1 – DIV3) sobre la privacidad y sobre cómo llevar a cabo estudios de estado del arte.

EST2 Privacy and Identity Lab (PI.lab) – Radboud University

Esta estancia se llevó a cabo en las instalaciones de PI.lab en los Países Bajos, donde se trabajó conjuntamente con Jaap-Henk Hoepman (Director científico de PI.lab) y Michael Colesky (Coordinador del proyecto P4P). Durante esta estancia se llevó a cabo la colaboración COL4 donde se elaboraron los sistemas de patrones reportados en el Capítulo 7 y publicados en los artículos CORE3 y CORE4.

8.5. Trabajos fin de titulación

8.5.1. Trabajos fin de grado

SUP1 Guzmán, D., 2020. Desarrollo de un sistema de soporte en el proceso de experimentación con patrones de privacidad. Ingeniería en Electrónica y Redes de Información. Escuela Politécnica Nacional.

Este trabajo fin de grado dio como resultado la creación de una herramienta de soporte a la experimentación con patrones de privacidad. Esta herramienta fue usada durante el estudio experimental exploratorio reportado en el Capítulo 4.

---

²² <http://www.cs.ru.nl/~jhh/p4p/p4p.html>

8.5.2. Trabajos fin de máster

SUP2 García, B., 2018. A systematic literature review on the applicability of security patterns. Máster Universitario en Ingeniería de Redes y Servicios Telemáticos. Universidad Politécnica de Madrid.

Proyecto fin de máster premiado con la **Matrícula de honor** de esa cohorte.

Este trabajo fin de máster se constituyó en el trabajo preliminar para llegar a proponer el marco de referencia de aplicación de los patrones de privacidad reportado en el Capítulo 5.

8.6. Conferencias y presentaciones realizadas

DIV1 Caiza, J.C., 2020. Necesito hacer mi tesis, ¿cómo hago mi estudio del estado del arte? Escuela Politécnica Nacional. Quito – Ecuador.
<https://detri.epn.edu.ec/necesito-hacer-mi-tesis-como-hago-mi-estudio-del-estado-del-arte/>

Se planificó, preparó y llevó a cabo esta conferencia que recoge los conocimientos obtenidos durante los estudios de literatura realizados a lo largo de la tesis. La conferencia detalla las etapas y tareas a realizarse en estudios de literatura, y los distingue para nivel de pregrado y posgrado.

DIV2 Caiza, J.C., 2020. ¡Mi teléfono me escucha! La importancia de la privacidad y los datos personales. Escuela Politécnica Nacional. Quito – Ecuador.
<https://www.redineditaepn.com/wordpress/mi-telefono-me-escucha-la-importancia-de-la-privacidad-y-los-datos-personales/>

Se planificó, preparó y llevó a cabo esta conferencia que recoge los conocimientos de la importancia y contexto de la privacidad obtenidos a lo largo de la realización de esta tesis. La conferencia describe la privacidad, su contexto, importancia, y la perspectiva de ingeniería.

DIV3 Caiza, J.C., 2019. ¡Mi teléfono me escucha! La importancia de la privacidad y los datos personales. Escuela Politécnica Nacional. Quito – Ecuador.
<https://detri.epn.edu.ec/mi-telefono-me-escucha/>

Se planificó, preparó y llevó a cabo esta conferencia que recoge los conocimientos de la importancia y contexto de la privacidad obtenidos a lo largo de la realización de esta tesis. La conferencia describe la privacidad, su contexto, importancia, y la perspectiva de ingeniería.

8.7. Cursos y seminarios dictados

DIV4 Caiza, J.C., 2020. Diseño de sistemas con requisitos de privacidad. Escuela Politécnica Nacional y Universidad de las Américas. Quito – Ecuador.
<https://detri.epn.edu.ec/proteccion-de-datos-privacidad-desde-el-diseno/>

Se planificó, preparó y llevó a cabo este curso que recoge los conocimientos obtenidos a lo largo de toda esta tesis. El curso incluyó la introducción a la privacidad, el proceso de diseño basado en estrategias, y el diseño usando patrones de privacidad.

Capítulo 9

Conclusiones y trabajo futuro

El uso de los datos como materia prima en el mundo digital y escándalos de uso masivo e ilícito de los mismo son parte de un escenario en el que las organizaciones necesitan los datos para el funcionamiento de sus modelos de negocio y las personas están cada vez más conscientes y preocupadas por su privacidad. En este escenario complejo, legislaciones como el RGPD tratan de poner un orden a través de diferentes formas como la protección de datos desde el diseño.

Propuestas de herramientas, técnicas, métodos y teorías para cada una de las etapas en el desarrollo de sistemas respetuosos con la privacidad han venido incrementándose. Sin embargo, aquellas enfocadas en la actividad del diseño están emergiendo, son dispersas y aún no se ha estudiado su aplicación en la práctica.

Esta tesis doctoral se ha enfocado en contribuir al avance del conocimiento sobre el diseño de sistemas respetuosos con la privacidad para llevarlo a la práctica, enfocándose principalmente en los patrones de privacidad al ser uno de los elementos que más atención ha obtenido en el ámbito.

9.1. Conclusiones

Para cumplir con los requisitos para optar por la mención internacional en el título, establecidos en el Art. 33 del Reglamento de Elaboración y Evaluación de la Tesis Doctoral de la Universidad Politécnica de Madrid, se presentan a continuación las conclusiones en español e inglés.

9.1.1. Versión en español

Los resultados obtenidos en los estudios realizados en el marco de esta tesis han permitido ir respondiendo a cada una de las preguntas de investigación formuladas inicialmente. Esto ha permitido alcanzar el objetivo general planteado en la introducción (Capítulo 1): contribuir al conocimiento sobre el diseño de sistemas respetuosos con la privacidad haciendo avances hacia su implementación en la práctica.

El estudio sistemático de literatura detallado en el Capítulo 3 ha permitido elaborar **un mapa de los elementos existentes para el diseño de sistemas de información respetuosos con la privacidad**. El mapa muestra los elementos reportados en el estado del arte e incluyen modelos de referencia, estrategias, métodos, herramientas, patrones, entre otros (respondiendo con ello la pregunta de investigación *P1: ¿Cuáles son las propuestas existentes para el diseño de sistemas respetuosos con la privacidad?*). En particular, el área de investigación en patrones de privacidad recibe la mayor cantidad de contribuciones, habiendo varios trabajos que hacen nuevas propuestas de patrones, y de herramientas, métodos y estrategias de diseño usando estos nuevos patrones. Sin embargo, el área de investigación es aún incipiente y sus contribuciones se concentran en el desarrollo de nuevas propuestas. Prácticamente hay nula investigación en temas muy importantes como el proceso de aplicación (respondiendo así a la pregunta de investigación *P.1.2: ¿Cuáles con las tendencias y madurez de la investigación en los patrones de privacidad?*).

El proceso de aplicación es de gran relevancia de cara a probar las características y beneficios de los patrones de privacidad, que usualmente han sido asumidos por las

diferentes propuestas en el área, y para promover la transferencia de conocimiento de la investigación a la ingeniería. Por ello, en el Capítulo 4 se detalla un estudio experimental exploratorio para determinar la percepción de los desarrolladores y los beneficios de aplicar los patrones de privacidad. Los resultados no permiten ser concluyentes respecto a los beneficios atribuidos a los patrones de privacidad. Sin embargo, el **análisis de los instrumentos y oportunidades de experimentación en la aplicación de los patrones de privacidad** efectuado permite notar que se pueden aplicar los patrones de privacidad estudiados, pero esta aplicación es limitada debido a ciertas condiciones de los patrones (esto permite responder la pregunta *P2: ¿Los patrones de privacidad actuales reúnen las características suficientes para su aplicación?*). Estas condiciones vienen desde la misma etapa de definición de los patrones y podrían alterar su instanciación exitosa.

Para profundizar en estas ideas, el Capítulo 5 muestra un estudio llevado a cabo exclusivamente para encontrar y entender cuáles son las cuestiones que afectan el proceso de aplicación de los patrones de privacidad. De este modo, se llega a conformar un **marco de referencia para la aplicación de los patrones de privacidad** que está compuesto por aspectos que categorizan los elementos que perjudican (defectos) o que favorecen (mejoras) a una instanciación exitosa de los patrones de privacidad (respondiendo así a la pregunta de investigación *P3: ¿Cuáles son las cuestiones que afectan la aplicación de los patrones de privacidad?*). Se analiza entonces el estado actual de las mejoras ya reportadas en la investigación de patrones de privacidad (contestando a la pregunta de investigación *P4: ¿Qué medidas ya se han implementado para mejorar la aplicación de los patrones de privacidad?*). Considerando este análisis, se definen aquellas mejoras faltantes que más impacto podrían causar de cara a favorecer una instanciación exitosa de los patrones de privacidad: las interconexiones entre los patrones de privacidad y la creación de sistemas de patrones (Esto responde a la pregunta de investigación *P5: ¿Qué medidas deberían implementarse primero para causar más impacto?*).

De modo que en el Capítulo 6 se elabora una **taxonomía de tipos de relaciones para definir conexiones entre patrones de privacidad** que pertenecen a los catálogos individuales encontrados en la literatura y al repositorio común que se ha conformado en *privacypatterns.org*. La taxonomía podría ser versátil y útil tanto para los autores de patrones de privacidad como para los desarrolladores que usen patrones de privacidad. Para los autores, ya que funciona como una caja de herramientas que provee el tipo de relación más adecuado para establecer una conexión entre dos patrones. Por otro lado, las relaciones conformadas pueden brindar información específica a los desarrolladores para que establezcan un grupo de patrones más apropiados de cara a resolver un problema de diseño. La taxonomía es inicialmente validada en los catálogos individuales, y posteriormente, a través de su uso para la construcción de dos sistemas de patrones. De este modo se responde a las preguntas de investigación *P6. ¿Se pueden implementar las medidas prioritarias?* y *P6.1 ¿Cómo se pueden validar estas medidas?*

Finalmente, el Capítulo 7 muestra **dos sistemas de patrones de privacidad** elaborados: uno conformado por 31 patrones de privacidad, orientado a proveer información al usuario sobre el procesamiento de sus datos personales, y el otro, conformado por 20 patrones de privacidad, para permitirle controlar dicho procesamiento. Dentro de los requisitos que se cumplieron para que los patrones del repositorio alcancen el nivel de sistema, está el establecimiento exhaustivo de las interconexiones entre ellos. Por ende, la elaboración de los dos sistemas sirve como validación de la taxonomía de tipos de relaciones propuesta previamente. Además, los sistemas de patrones han sido

propuestos a los investigadores del esfuerzo común (en el área de investigación de los patrones de privacidad) para que se validen antes de su publicación en el repositorio. Actualmente, ambos sistemas de patrones, están publicados en el repositorio. Esto permite responder a las preguntas de investigación *P6. ¿Se pueden implementar las medidas prioritarias? y P6.1 ¿Cómo se pueden validar estas medidas?*

Las contribuciones realizadas a lo largo de la tesis han permitido contestar a todas las preguntas de investigación planteadas, y que en conjunto han permitido llegar a cumplir con el objetivo general.

9.1.2. Versión en inglés

The results obtained in the studies carried out as part of this thesis have made it possible to respond to each of the research questions initially formulated. They have led to the achievement of the general objective defined in the introduction (Chapter 1): contributing to knowledge about the design of privacy-aware systems by making advances towards its implementation in practice.

The systematic literature study, detailed in Chapter 3, has made it possible to create a **map of existing elements for the design of privacy-aware information systems**. The map shows the elements reported in the state of the art and includes reference models, strategies, methods, tools and privacy patterns, among others (answering the research question *P1: What are the existing proposals for the design of privacy-aware systems?*). The research area of privacy patterns has received most of the contributions, which include new pattern proposals, and tools, methods and design strategies using these new patterns. However, this research area is still incipient, and its contributions mainly focus on the development of new proposals. There is practically no research on very important issues such as the application process (answering the research question *P.1.2: What are the trends and maturity of privacy patterns research?*).

The application process is necessary to prove the characteristics and benefits attributed to privacy patterns and to promote knowledge transfer from research to engineering. Therefore, Chapter 4 details an exploratory experimental study to determine developers' perception and the benefits of applying privacy patterns. The results did not allow us to be conclusive about the benefits. However, **the analysis of the instruments and experimentation opportunities in the application of privacy patterns** lead us to note that the studied privacy patterns can be applied in a limited way. Their application is limited due to certain conditions of the applied patterns (answering the question *P2: Do the current privacy patterns have sufficient characteristics for their application?*), which may come even from the early stages of their development, such as a pattern's definition, and could affect their successful instantiation.

To further explore these ideas, Chapter 5 shows a study conducted exclusively to find and understand the issues affecting the process of applying privacy patterns. The result is **a framework for the application of privacy patterns**, which is composed of aspects categorizing elements that negatively (drawbacks) or positively (improvements) affect a successful instantiation (answering the research question *P3: What are the issues affecting the application of privacy patterns?*). Next, this study analyzed the current state of privacy pattern research by looking for improvements (answering the research question *P4: What measures have already been implemented to improve the application of privacy patterns?*). After this analysis, the most impactful improvements for a successful instantiation are defined: the interconnections between privacy patterns and the creation of pattern systems (answering the research question *P5: Which measures should be implemented first to provoke the most impact?*).

In Chapter 6, **a taxonomy of types of relationships to define connections between privacy patterns is proposed**. It works to establish relationships between patterns of the individual catalogues found in the literature and patterns from the common repository at *privacypatterns.org*. This taxonomy could be versatile and useful for authors of privacy patterns and for developers using privacy patterns. For authors, it works as a toolbox that provides the most appropriate type of relationship to use to establish a connection between two patterns. These connections can provide specific information to the developers so that they can identify an adequate set of patterns to solve a design problem. This taxonomy was initially validated in the individual catalogues, and later, it was used for the creation of two pattern systems. It answers the research questions *P6. Can the priority measures be implemented?* and *P6.1 How can these measures be validated?*

Finally, Chapter 7 shows **two privacy pattern systems**: one, consisting of 31 privacy patterns, that provides information to the data subjects about the processing of their personal data, and the other, consisting of 20 privacy patterns, that allows them to control such processing. The requirements fulfilled in order to advance the patterns within the common repository to the system level included the exhaustive establishment of interconnections between the patterns. Therefore, the development of the two pattern systems is a validation of the proposed taxonomy of types of relationships. In addition, the privacy pattern systems were proposed to the researchers of the common effort (in privacy patterns research) to be validated before their publication in the repository. Currently, both pattern systems are published there. This answers the research questions *P6. Can the priority measures be implemented?* and *P6.1 How can these measures be validated?*

The contributions made by this thesis allowed us to answer all of the research questions. Likewise, they have allowed us to reach the general objective.

9.2. Trabajos futuros

Como se ha mencionado a lo largo de esta tesis, la investigación en el área de los patrones de privacidad está en una etapa emergente; y aunque cada vez se propongan más patrones y elementos metodológicos basados en ellos, es indispensable que se llegue a probar en la práctica los beneficios que se les atribuye. En este sentido, hemos realizado valiosas contribuciones para avanzar el conocimiento sobre el proceso de aplicación de los patrones. Sin embargo, queda mucho por investigar tanto en el tema de aplicación, como en otros dentro de la investigación en los patrones de privacidad.

Aplicación de patrones. En el Capítulo 5 se estructuró un marco de referencia para el proceso de aplicación de los patrones y se analizó, en función de este, el estado actual de la investigación en los patrones de privacidad. Aunque se han cubierto los dos elementos que más impacto podrían causar considerando el estado de los patrones de privacidad al momento de llevar a cabo los estudios, quedan otros trabajos que podrían aportar mucho.

- La definición de una forma de representación útil para solucionar los defectos que escapan a las posibilidades de los sistemas de patrones (p.ej. heterogeneidad en la descripción y los diferentes niveles de abstracción entre patrones del mismo tipo). Se visualizan esquemas de representación útiles, como aquellos basados en modelos, que permitan uniformizar la descripción de los patrones de privacidad. Además, esto podría dar cabida a utilizar ingeniería basada en modelos. Por ejemplo, se podrían usar transformaciones basadas en modelos de

cara a automatizar el proceso de diseño o para facilitar la evaluación de los sistemas construidos con patrones de privacidad. Una idea que empezamos a explorar es que los patrones de privacidad actuales tanto en el apartado de descripción del problema como en el de la solución podrían revelar potenciales transformaciones a realizar.

- La creación de técnicas y herramientas de navegación y selección automatizadas o semi-automatizadas que guíen a los desarrolladores hacia un grupo de patrones más apropiados para resolver un problema de diseño en particular. Consideramos que los esquemas de estructura, categorización e interconexión uniformizados pueden sentar las bases para el desarrollo de ese tipo de herramientas. Actualmente, los repositorios en línea han sido presentados como herramientas que ayudan en la búsqueda y selección, pero su funcionalidad es limitada ya que se basa únicamente en la categorización de los patrones; y, además, en escenarios con una gran cantidad de patrones, su funcionamiento manual no ayudaría al desarrollador. Visualizamos herramientas que aprovechen las ventajas de navegación proporcionadas por el esquema de relaciones, incluso mejorado a través de propiedades de las relaciones que permitan determinar por dónde se puede ir o no en la navegación, o por donde resulta más conveniente hacerlo para un problema planteado.
- El aspecto de calidad de los patrones de privacidad no debe ser descuidado, ya que según se ha visto en los estudios realizados, varios patrones del repositorio *privacypatterns.org* aún tienen defectos. Aunque se sabe que los eventos PLOP son los más adecuados para recibir una evaluación de los patrones, según se ha visto en la literatura estudiada, la mayoría de propuestas de patrones que se publican no son enviadas allí. Se visualiza un marco de evaluación de la calidad de los patrones de privacidad, que complementariamente o como primer filtro a un evento PLOP, pueda ser usado de cara a reducir tiempos de procesamiento de las propuestas y para garantizar niveles de calidad satisfactorios.
- Posteriormente a las mejoras previamente comentadas, o de manera paralela, es necesario que a medida que se propongan nuevos patrones de privacidad, se realicen también más propuestas de sistemas y lenguajes. El desarrollo de este tipo de propuestas hará que se mejoren los patrones de privacidad, bien sea agrupados en dominios de aplicación más (lenguajes) o menos (sistemas) específicos. En cualquier caso, dado que hay un amplio espectro de dominios de aplicación para los sistemas de información que tratan datos personales, hay mucho espacio para proponer sistemas y lenguajes de patrones de privacidad.

Evaluación empírica. Todas las nuevas propuestas de elementos deberían ir acompañadas de estudios empíricos bien sea a nivel de laboratorio, o bien en entornos reales. Esto permitiría validarlas conceptualmente y en la práctica. En aras de promover este tipo de propuestas, luego del estudio exploratorio reportado en el Capítulo 4, hemos reunido un conjunto de instrumentos que pueden ser tomados como base para futuros estudios. De entre los elementos creados, destacamos la herramienta de experimentación construida y un compendio de diferentes métricas que se podrían utilizar agrupadas por persona, proceso o producto. Se visualizan estudios futuros que experimenten con las mejoras que se vayan haciendo a los patrones de privacidad y con los procesos de diseño respetuoso con la privacidad completos.

Minería de patrones de privacidad. En la coyuntura actual de la pandemia originada por el virus SARS-Cov-2 (causante de la enfermedad COVID-19), el tema de la privacidad está implícito al uso de herramientas tecnológicas de rastreo que se vuelven indispensables para localizar a las personas infectadas y establecer un cerco epidemiológico. Cuestiones de privacidad como la identificación de la persona, la localización, la determinación de hábitos, la compartición de datos personales, etc. son parte de un escenario complejo donde hay urgencia para construir este tipo de herramientas como parte de, y para definir, estrategias de salud que buscan salvaguardar la vida de la población.

Varias herramientas se están construyendo por empresas tecnológicas, operadoras, gobiernos, e incluyen interfaces de programación de aplicaciones (APIs), aplicaciones móviles, aplicaciones de identificación de datos biométricos, entre otras. En este contexto, varios patrones de desarrollo y de diseño se irán produciendo y por lo tanto se podrán minar.

Investigación para la enseñanza. Los patrones de privacidad, por definición están orientados a su uso por personas no expertas en privacidad, y precisamente por ello podrían ser explotadas en el proceso de enseñanza. Esto ha sucedido previamente en patrones en otros dominios como los de orientación a objetos o el de seguridad. En este sentido, se visualiza la investigación del uso de patrones de privacidad como medios para la enseñanza de ingeniería de privacidad, dominio de investigación y práctica profesional que está en pleno surgimiento.

Bibliografía

- [Abramov12] Abramov, J., Sturm, A., & Shoval, P. (2012). *Evaluation of the Pattern-based method for Secure Development (PbSD): A controlled experiment*. Information and Software Technology, 54(9), 1029–1043. <https://doi.org/10.1016/j.infsof.2012.04.001>
- [ACM12] Association for Computing Machinery. (2012). *The 2012 ACM computing classification system*. https://dl.acm.org/ccs/ccs_flat.cfm
- [AEPD19] Agencia Española de Protección de Datos. (2019). *A Guide to Privacy by Design*.
- [Alexander77] Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press.
<https://doi.org/10.1145/1529966.1529967>
- [Aljohani16] Aljohani, M., Hawkey, K., & Blustein, J. (2016). *Proposed privacy patterns for privacy preserving healthcare systems in accord with Nova Scotia's Personal Health Information Act*. In T. Tryfonas (Ed.), Human Aspects of Information Security, Privacy, and Trust. HAS 2016. Lecture Notes in Computer Science (Vol. 9750, pp. 91–102). Springer, Cham. https://doi.org/10.1007/978-3-319-39381-0_9
- [Alshammari18] Alshammari, M., & Simpson, A. (2018). *Privacy Architectural Strategies*. Proceedings of the 2018 Workshop on Privacy in the Electronic Society - WPES'18, 143–154.
<https://doi.org/10.1145/3267323.3268957>
- [Anthonysamy17] Anthonysamy, P., Rashid, A., & Chitchyan, R. (2017). *Privacy requirements: Present & future*. 2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering in Society Track, 13–22.
<https://doi.org/10.1109/ICSE-SEIS.2017.3>
- [Ayalon17] Ayalon, O., Toch, E., Hadar, I., & Birnhack, M. (2017). *How developers make design decisions about users' privacy: The place of professional communities and organizational climate*. Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, 135–138.
<https://doi.org/10.1145/3022198.3026326>
- [Bafandeh17] Bafandeh Mayvan, B., Rasoolzadegan, A., & Ghavidel Yazdi, Z. (2017). *The state of the art on design patterns: A systematic mapping of the literature*. Journal of Systems and Software, 125, 93–118. <https://doi.org/10.1016/j.jss.2016.11.030>

- [Beckers12] Beckers, K. (2012). *Comparing privacy requirements engineering approaches*. Seventh International Conference on Availability, Reliability and Security, ARES 2012, 574–581.
- [Bednar19] Bednar, K., Spiekermann, S., & Langheinrich, M. (2019). *Engineering Privacy by Design: Are engineers ready to live up to the challenge?* Information Society, 35(3), 122–142.
<https://doi.org/10.1080/01972243.2019.1583296>
- [Bier12] Bier, C., & Krempel, E. (2012). *Common privacy patterns in video surveillance and smart energy*. 7th International Conference on Computing and Convergence Technology (ICCT), 610–615.
- [Buschmann96] Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., & Stal, M. (1996). *Pattern-oriented software architecture: a system of patterns*. In John Wiley&Sons (Vol. 1). Wiley.
<https://doi.org/10.1192/bjp.108.452.101>
- [Cadwalladr18] Cadwalladr, C., & Graham-Harrison, E. (2018). *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*. The Guardian.
<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- [Caiza20] Caiza, J. C., Del Alamo, J. M., & Guamán, D. S. (2020). *A framework and roadmap for enhancing the application of privacy design patterns*. The 35th ACM/SIGAPP Symposium On Applied Computing, 1297–1304. <https://doi.org/10.1145/3341105.3375768>
- [Caiza17] Caiza, J. C., Martín, Y.-S., Del Alamo, J. M., & Guamán, D. S. (2017). *Organizing design patterns for privacy: A taxonomy of types of relationships*. Proceedings of the 22nd European Conference on Pattern Languages of Programs - EuroPLoP '17, 1–11. <https://doi.org/10.1145/3147704.3147739>
- [Caiza19] Caiza, J. C., Martin, Y.-S., Guaman, D. S., Del Alamo, J. M., & Yelmo, J. C. (2019). *Reusable Elements for the Systematic Design of Privacy-Friendly Information Systems: A Mapping Study*. IEEE Access, 7, 66512–66535.
<https://doi.org/10.1109/ACCESS.2019.2918003>
- [Carver10] Carver, J. C., Jaccheri, L., Morasca, S., & Shull, F. (2010). *A checklist for integrating student empirical studies with research and teaching goals*. Empirical Software Engineering, 15(1), 35–59.
<https://doi.org/10.1007/s10664-009-9109-9>
- [Cavacini15] Cavacini, A. (2015). *What is the best database for computer science journal articles?* Scientometrics, 102(3), 2059–2071.
<https://doi.org/10.1007/s11192-014-1506-1>

- [Cavoukian09] Cavoukian, A. (2009). *Privacy by design: The 7 foundational principles*. Information and Privacy Commissioner of Ontario.
- [CCCSA] Cloud Security Alliance. (2018). *Code of conduct for GDPR compliance*.
- [Chung04] Chung, E. S., Hong, J. I., Lin, J., Prabaker, M. K., Landay, J. A., & Liu, A. L. (2004). *Development and evaluation of emerging design patterns for ubiquitous computing*. Proceedings of the 2004 Conference on Designing Interactive Systems Processes, Practices, Methods, and Techniques - DIS '04, 233–242.
- [Clarke97] Clarke, R. (1997). *Introduction to dataveillance and information privacy, and definitions of terms*.
<http://rogerclarke.com/DV/Intro.html>
- [Colesky19] Colesky, M., & Caiza, J. C. (2019). *A System of Privacy Patterns for Informing Users: Creating a Pattern System*. Proceedings of the 23rd European Conference on Pattern Languages of Programs - EuroPLoP '18, 1–11. <https://doi.org/10.1145/3282308.3282325>
- [Colesky18] Colesky, M., Caiza, J. C., Del Álamo, J., Hoepman, J.-H., & Martín, Y.-S. (2018). *A system of privacy patterns for user control*. 33rd Annual ACM Symposium on Applied Computing, 1150–1156. <https://doi.org/10.1145/3167132.3167257>
- [Colesky16] Colesky, M., Hoepman, J.-H., & Hillen, C. (2016). *A critical analysis of privacy design strategies*. 2016 IEEE Security and Privacy Workshops (SPW), 33–40.
<https://doi.org/10.1109/SPW.2016.23>
- [Cysneiros04] Cysneiros, L. M., & Leite, J. C. S. do P. (2004). *Nonfunctional requirements: From elicitation to conceptual models*. IEEE Transactions on Software Engineering, 30(5), 328–350.
- [Danezis14] Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Metayer, D. Le, Tirtea, R., & Schiffner, S. (2014). *Privacy and Data Protection by Design - from policy to engineering*. European Union Agency for Cybersecurity. <https://doi.org/10.2824/38623>
- [DPPC10] Data Protection and Privacy Commissioners. (2010). *Resolution on privacy by design*. In *32nd International Conference of Data Protection and Privacy Commissioners*.
https://edps.europa.eu/sites/edp/files/publication/10-10-27_jerusalem_resolutionon_privacybydesign_en.pdf
- [De16] De, S. J., & Métayer, D. Le. (2016). *Privacy Risk Analysis*. Synthesis Lectures on Information Security, Privacy, and Trust, 81(3), 1–133. <https://doi.org/10.2200/s00724ed1v01y201607spt017>

- [DelAlamo18] Del Alamo, J. M., Martín, Y. S., & Caiza, J. C. (2018). *Towards organizing the growing knowledge on privacy engineering*. In IFIP Advances in Information and Communication Technology (Vol. 526, pp. 15–24). https://doi.org/10.1007/978-3-319-92925-5_2
- [Deng11] Deng, M., Wuyts, K., Scandariato, R., Preneel, B., & Joosen, W. (2011). *A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements*. Requirements Engineering, 16(1), 3–32.
- [Doty11a] Doty, N., & Gupta, M. (2011). *Privacy Patterns: Collaborative development of privacy software design patterns*.
<https://github.com/privacypatterns>
- [Doty11b] Doty, N., Gupta, M., Hoepman, J.-H., Colesky, M., Caiza, J. C., Del Alamo, J. M., Martín, Y. S., Le Métayer, D., Drozd, O., Boesch, C., Kargl, F., Kopp, H., & Mosby, P. (2011). *Privacy Patterns*. <https://privacypatterns.org/>
- [Drozd16a] Drozd, O. (2016a). *Catalog of Privacy Patterns*. <http://privacypatterns.wu.ac.at:8080/catalog/>
- [Drozd16b] Drozd, O. (2016b). *Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process*. In D. Aspinall, J. Camenisch, M. Hansen, S. Fischer-Hübner, & C. Raab (Eds.), Privacy and Identity Management. Time for a Revolution?. Privacy and Identity 2015. IFIP Advances in Information and Communication Technology (Vol. 476, pp. 129–140). Springer, Cham.
- [Easterbrook08] Easterbrook, S., Singer, J., Storey, M. A., & Damian, D. (2008). *Selecting empirical methods for software engineering research*. In Guide to Advanced Empirical Software Engineering (pp. 285–311). Springer London. https://doi.org/10.1007/978-1-84800-044-5_11
- [Elsevier20] Elsevier. (2020). *Scopus content coverage guide*. <https://www.elsevier.com/?a=69451>
- [EuroC02] European Commission. (2002). DIRECTIVE 2002/58/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). In Official Journal of the European Communities.
- [EuroC10] European Commission. (2010). COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS - A Digital Agenda for Europe.

- [EuroC15] European Commission. (2015). COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS - A Digital Single Market Strategy for Europe.
- [EuroC17] European Commission. (2017). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications).
- [EuroC18] European Commission. (2018). Horizon 2020 - Work Programme 2018-2020: 14. Secure societies-Protecting freedom and security of Europe and its citizens.
- [EuroC19] European Comission. (2019). PROTECTING YOUR DATA WITHIN THE EU.
https://ec.europa.eu/info/sites/info/files/data-protection-overview-citizens_en_0.pdf
- [EuroC20] European Comission. (2020). European data strategy.
https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en
- [EuroDPS18] European Data Protection Supervisor. (2018). Preliminary Opinion on privacy by design.
- [EuroPC16] European Parliament & Council of the European Union. (2016). 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)
- [EuroPC95] European Parliament & Council of the European Union. (1995). 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.
- [EuroPLOP04] EuroPLOP. (2004). *European Conference on Pattern Languages of Programs*. <http://www.europlop.net/>
- [Falessi18] Falessi, D., Juristo, N., Wohlin, C., Turhan, B., Münch, J., Jedlitschka, A., & Oivo, M. (2018). *Empirical software engineering experts on the use of students and professionals in experiments*. Empirical Software Engineering, 23(1), 452–489.

- [Fhom11] Fhom, H. S., & Bayarou, K. M. (2011). *Towards a holistic privacy engineering approach for smart grid systems*. IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, 234–241.
<https://doi.org/10.1109/TrustCom.2011.32>
- [Finn13] Finn, R. L., Wright, D., & Friedewald, M. (2013). *Seven Types of Privacy*. In *European Data Protection: Coming of Age* (pp. 3–32). Springer Netherlands.
https://doi.org/10.1007/978-94-007-5170-5_1
- [Fischer-Hübner10] Fischer-Hübner, Simone, Köffel, C., Pettersson, J.-S., Wolkerstorfer, P., Graf, C., Holtz, L. E., König, U., Hedbom, H., & Kellermann, B. (2010). *HCI Pattern Collection – Version 2*. Privacy and Identity Management in Europe for Life.
- [Gamma95] Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley.
- [García18] García González, B. (2018). *A Systematic Literature Review on the Applicability of Security Patterns*.
https://www.dit.upm.es/~posgrado/doc/TFM/TFMs2017-18/tfm_GARCÍA_GONZÁLEZ_BEATRIZ_2017-18.pdf
- [Glaser17] Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. In Routledge. Taylor and Francis. <https://doi.org/10.4324/9780203793206>
- [GobEsp17] Gobierno de España. Ministerio de Economía Industria y Competitividad. (2017). *Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020*.
- [GobEsp13] Gobierno de España. Ministerio de Industria Energía y Turismo. (2013). *Agenda Digital para España*.
https://avancedigital.gob.es/planes-TIC/agenda-digital/DescargasAgendaDigital/Plan-ADpE_Agenda_Digital_para_Espana.pdf
- [Goldberg97] Goldberg, I., Wagner, D., & Brewer, E. (1997). *Privacy-enhancing technologies for the Internet*. IEEE Computer Society International Conference, 103–109. <https://doi.org/10.1109/cmpcon.1997.584680>
- [Gope16] Gope, P., & Amin, R. (2016). *A novel reference security model with the situation based access policy for accessing EPHR data*. Journal of Medical Systems, 40, 242.
<https://doi.org/10.1007/s10916-016-0620-4>

- [Guaman20] Guaman, D. S., Del Alamo, J. M., & Caiza, J. C. (2020). *A Systematic Mapping Study on Software Quality Control Techniques for Assessing Privacy in Information Systems*. IEEE Access, 8, 74808–74833. <https://doi.org/10.1109/access.2020.2988408>
- [Gürses16] Gürses, S., & Del Alamo, J. M. (2016). *Privacy engineering: shaping an emerging field of research and practice*. IEEE Security and Privacy, 14(2), 40–46. <https://doi.org/10.1109/MSP.2016.37>
- [Gürses08] Gürses, S., Rizk, R., & Günther, O. (2008). *Privacy design in online social networks: Learning from privacy breaches and community feedback*. International Conference on Information Systems – ICIS.
- [Guzman20] Guzman, D. (2020). *Desarrollo de un sistema de soporte en el proceso de experimentación con patrones de privacidad*.
- [Hadar18] Hadar, I., Hasson, T., Ayalon, O., Toch, E., Birnhack, M., Sherman, S., & Balissa, A. (2018). *Privacy by designers: software developers' privacy mindset*. Empirical Software Engineering, 23(1), 259–289. <https://doi.org/10.1007/s10664-017-9517-1>
- [Hafiz13] Hafiz, M. (2013). *A pattern language for developing privacy enhancing technologies*. Software - Practice and Experience, 43(7), 769–787. <https://doi.org/10.1002/spe.1131>
- [Hafiz06] Hafiz, M. (2006). *A collection of privacy design patterns*. Proceedings of the 2006 Conference on Pattern Languages of Programs - PLoP '06, 1–13.
- [Hannay07] Hannay, J. E., Sjøberg, D. I. K., & Dybå, T. (2007). *A systematic review of theory use in software engineering experiments*. In IEEE Transactions on Software Engineering (Vol. 33, Issue 2, pp. 87–107). <https://doi.org/10.1109/TSE.2007.12>
- [Hazeyama17] Hazeyama, A., Washizaki, H., Yoshioka, N., Kaiya, H., & Okubo, T. (2017). *Literature survey on technologies for developing privacy-aware software*. 2016 IEEE 24th International Requirements Engineering Conference Workshops (REW), 86–91. <https://doi.org/10.1109/REW.2016.27>
- [Heurix15] Heurix, J., Zimmermann, P., Neubauer, T., & Fenz, S. (2015). *A taxonomy for privacy enhancing technologies*. Computers and Security, 53, 1–17. <https://doi.org/10.1016/j.cose.2015.05.002>
- [Hoepman14] Hoepman, J.-H. (2014). *Privacy design strategies*. In N. Cuppens-Boulahia, F. Cuppens, S. Jajodia, A. Abou El Kalam, & T. Sans (Eds.), ICT Systems Security and Privacy Protection. SEC 2014. IFIP Advances in Information and Communication Technology (Vol. 428, pp. 446–459). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-55415-5_38

- [Hoepman18] Hoepman, J.-H. (2018). *Privacy Design Strategies (The Little Blue Book)* (Nijmegen : Radboud University (ed.)). Nijmegen : Radboud University. <http://www.cs.ru.nl/~jhh/publications/pds-booklet.pdf>
- [Iachello05] Iachello, G., & Abowd, G. D. (2005). *Privacy and proportionality: Adapting legal evaluation techniques to inform design in ubiquitous computing*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05, 91–100. <https://doi.org/10.1145/1054972.1054986>
- [IEEE14] IEEE Computer Society. (2014). *Guide to the Software Engineering Body of Knowledge SWEBOK*® A Project of the IEEE Computer Society.
- [IEEE17] Institute of Electrical and Electronics Engineers (IEEE). (2017). *IEEE Thesaurus, Version 1.0*.
https://www.ieee.org/content/dam/ieee-org/ieee/web/org/pubs/ieee_thesaurus.pdf
- [Inventado16] Inventado, P. S., & Scupelli, P. (2016). *Data-driven design pattern production: a case study on the ASSISTments online learning system*. EuroPLoP '15: Proceedings of the 20th European Conference on Pattern Languages of Program, 13.
- [Islam10] Islam, S., Mouratidis, H., & Wagner, S. (2010). *Towards a framework to elicit and manage security and privacy requirements from laws and regulations*. In R. Wieringa & A. Persson (Eds.), *Requirements Engineering: Foundation for Software Quality* (Vol. 6182, pp. 255–261). Springer, Berlin, Heidelberg.
https://doi.org/10.1007/978-3-642-14192-8_23
- [ISO/IEC29100] ISO/IEC. (2011). ISO/IEC 29100:2011 Information technology — Security techniques — Privacy framework.
<https://www.iso.org/standard/45123.html>
- [ISO/IEC24744] ISO/IEC. (2014). 24744:2014, Software engineering – Metamodel for development methodologies. In Switzerland, Geneva (p. 96). International Organization for Standardization.
<https://www.iso.org/standard/62644.html>
- [ISO/IEC29134] ISO/IEC. (2017a). ISO/IEC 29134:2017 Information technology — Security techniques — Guidelines for privacy impact assessment.
<https://www.iso.org/obp/ui/#iso:std:iso-iec:29134:ed-1:v1:en>
- [ISO/IEC29151] ISO/IEC. (2017b). ISO/IEC 29151:2017 Information technology — Security techniques — Code of practice for personally identifiable information protection.
<https://www.iso.org/obp/ui/#iso:std:iso-iec:29151:ed-1:v1:en>

- [ISO/IEC27550] ISO/IEC. (2019). ISO/IEC TR 27550:2019 Information technology — Security techniques — Privacy engineering for system life cycle processes.
<https://www.iso.org/standard/72024.html>
- [ISO/IEC/IEEE24765] ISO/IEC/IEEE. (2014). ISO/IEC/IEEE 24765:2017 - Systems and software engineering -- Vocabulary.
<https://www.iso.org/standard/71952.html>
- [Ito15] Ito, Y., Washizaki, H., Yoshizawa, M., Fukazawa, Y., Okubo, T., Kaiya, H., Hazeyama, A., Yoshioka, N., & Fernandez, E. B. (2015). *Systematic Mapping of Security Patterns Research*. 22nd Conference on Pattern Languages of Programs, 3–4.
- [JefEstEsp99] Jefatura del Estado. (1999). *Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal* (pp. 43088–43099). BOE 298.
<https://www.boe.es/boe/dias/1999/12/14/pdfs/A43088-43099.pdf>
- [Kalloniatis08] Kalloniatis, C., Kavakli, E., & Gritzalis, S. (2008). *Addressing privacy requirements in system design: The PriS method*. Requirements Engineering, 13(3), 241–255.
<https://doi.org/10.1007/s00766-008-0067-3>
- [Kalloniatis09] Kalloniatis, C., Kavakli, E., & Gritzalis, S. (2009). *Methods for designing privacy aware information systems: A review*. 2009 - 13th Panhellenic Conference on Informatics, 185–194.
<https://doi.org/10.1109/PCI.2009.45>
- [Karjoth02] Karjoth, G., Schunter, M., & Waidner, M. (2002). *Privacy-enabled services for enterprises. Proceedings*. 13th International Workshop on Database and Expert Systems Applications, 483–487.
<https://doi.org/10.1109/DEXA.2002.1045944>
- [Kitchenham 04] Kitchenham, B. A., Dybå, T., Jørgensen, M., Dyba, T., & Jorgensen, M. (2004). *Evidence-Based Software Engineering*. Proceedings of the 26th International Conference on Software Engineering, 273–281.
- [Krippendorff09] Krippendorff, K. (2009). *Testing the reliability of content analysis data: What is involved and why*. In The Content Analysis Reader (pp. 350–357). Sage Publications.
- [Kuneva09] Kuneva, M. (2009). *Roundtable on Online Data Collection, Targeting and Profiling*. European Consumer Commissioner.
https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_09_156

- [Kung14] Kung, A. (2014). *PEARs: Privacy enhancing architectures*. In B. Preneel & D. Ikononou (Eds.), *Privacy Technologies and Policy*. APF 2014. Lecture Notes in Computer Science (Vol. 8450, pp. 18–29). Springer, Cham. https://doi.org/10.1007/978-3-319-06749-0_2
- [Lenhard17] Lenhard, J., Fritsch, L., & Herold, S. (2017). *A literature study on privacy patterns research*. Proceedings of the 43rd Euromicro Conference on Software Engineering and Advanced Applications, 194–201. <https://doi.org/10.1109/SEAA.2017.28>
- [Lobato09] Lobato, L. L., Fernandez, E. B., & Zorzo, S. D. (2009). *Patterns to support the development of privacy policies*. Proceedings - International Conference on Availability, Reliability and Security, ARES 2009, 744–749. <https://doi.org/10.1109/ARES.2009.114>
- [Lombard10] Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2010). *Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects*.
- [Mackenzie06] Mackenzie, C. M., McCabe, F., Brown, P. F., Net, P., Metz, R., & Hamilton, A. (2006). *Reference Model for Service Oriented Architecture 1.0*.
- [Mangalaraj14] Mangalaraj, G., Nerur, S., Mahapatra, R., & Price, K. H. (2014). *Distributed Cognition in Software Design: An Experimental Investigation of the Role of Design Patterns and Collaboration*. MIS Quarterly, 38(1), 249–274. <https://doi.org/10.2307/26554877>
- [Marin03] Marin, J. M. (2003). *Análisis no paramétrico*. Análisis Descriptivo Con SPSS.
<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/GuiaSPSS/19nparam.pdf>
- [Martín17] Martín, Y.-S., & Del Álamo, J. M. (2017). *A metamodel for privacy engineering methods*. 3rd International Workshop on Privacy Engineeringco-Located with 38th IEEE Symposium on Security and Privacy (S&P 2017), 1873, 41–48.
- [Meszaros97] Meszaros, G., & Doble, J. (1997). *A pattern language for pattern writing*. Proceedings of International Conference on Pattern Languages of Program Design, 164.
- [Moncrieff09] Moncrieff, S., Venkatesh, S., & West, G. (2009). *A Framework for the design of privacy preserving pervasive healthcare*. IEEE International Conference on Multimedia and Expo, 1696–1699. <https://doi.org/10.1109/ICME.2009.5202847>
- [Morales-Trujillo19] Morales-Trujillo, M. E., García-Mireles, G. A., Matla-Cruz, E. O., & Piattini, M. (2019). *A Systematic Mapping Study of Privacy by Design in Software Engineering*. CLEI Electronic Journal, 22(1).

- [Ng06] Ng, T. H., Cheung, S. C., Chan, W. K., & Yu, Y. T. (2006). *Toward effective deployment of design patterns for software extension: A case study*. Proceedings of the International Conference on Software Engineering, 51–56.
- [Nissenbaum10] Nissenbaum, H. (2010). *Privacy in context: technology, policy, and the integrity of social life*. In Stanford university Press.
<https://doi.org/10.5860/choice.47-6940>
- [Noble98] Noble, J. (1998). *Classifying relationships between object-oriented design patterns*. Proceedings 1998 Australian Software Engineering Conference (Cat. No.98EX233), 98–107.
<https://doi.org/10.1109/ASWEC.1998.730917>
- [Notario15] Notario, N., Crespo, A., Martin, Y. S., Del Alamo, J. M., Metayer, D. Le, Antignac, T., Kung, A., Kroener, I., & Wright, D. (2015). *PRIPARE: Integrating privacy best practices into a privacy engineering methodology*. 2015 IEEE Security and Privacy Workshops, 151–158. <https://doi.org/10.1109/SPW.2015.22>
- [OECD80] Organisation for Economic Co-operation and Development. (1980). *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*.
- [OECD14] Organisation for Economic Co-Operation and Development. (2014). SUMMARY OF THE OECD PRIVACY EXPERT ROUNDTABLE Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking.
<https://doi.org/10.1787/9789264193307-en>
- [Pandit19] Pandit, H. J., Polleres, A., Bos, B., Brennan, R., Bruegger, B., Ekaputra, F. J., Fernández, J. D., Hamed, R. G., Kiesling, E., Lizar, M., Schlehn, E., Steyskal, S., & Wenning, R. (2019). *Creating a vocabulary for data privacy*. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11877 LNCS, 714–730.
- [Pearson11] Pearson, S., & Benameur, A. (2011). *A decision support system for design for privacy*. In S Fischer-Hübner, P. Duquenoy, M. Hansen, R. Leenes, & G. Zhang (Eds.), Privacy and Identity Management for Life. Privacy and Identity 2010. IFIP Advances in Information and Communication Technology (Vol. 352). Springer, Berlin, Heidelberg.
- [Pearson10] Pearson, S., & Shen, Y. (2010). *Context-aware privacy design pattern selection*. In S. Katsikas, J. Lopez, & M. Soriano (Eds.), Trust, Privacy and Security in Digital Business. TrustBus 2010. Lecture Notes in Computer Science (Vol. 6264, pp. 69–80). Springer, Berlin, Heidelberg.

- [Petersen08] Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). *Systematic Mapping Studies in Software Engineering*. 12Th International Conference on Evaluation and Assessment in Software Engineering, 68–77.
<https://doi.org/10.1142/S0218194007003112>
- [Petersen15] Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). *Guidelines for conducting systematic mapping studies in software engineering: An update*. Information and Software Technology, 64, 1–18.
<https://doi.org/10.1016/j.infsof.2015.03.007>
- [Prechelt02] Prechelt, L., Unger-Lamprecht, B., Philippsen, M., & Tichy, W. F. (2002). *Two controlled experiments assessing the usefulness of design pattern documentation in program maintenance*. IEEE Transactions on Software Engineering, 28(6), 595–606.
<https://doi.org/10.1109/TSE.2002.1010061>
- [Preschern15] Preschern, C., Kajtazovic, N., & Kreiner, C. (2015). *Building a safety architecture pattern system*. Proceedings of the 18th.
- [PRIPARE15] PRIPARE Project. (2015). *privacypatterns.eu - Collecting patterns for better privacy*. <https://privacypatterns.eu/>
- [Riaz15] Riaz, M., Breau, T., & Williams, L. (2015). *How have we evaluated software pattern application? A systematic mapping study of research design practices*. In Information and Software Technology (Vol. 65, pp. 14–38). Elsevier.
<https://doi.org/10.1016/j.infsof.2015.04.002>
- [Romanosky06] Romanosky, S., Acquisti, A., Hong, J., Cranor, L. F., & Friedman, B. (2006). *Privacy patterns for online interactions*. Proceedings of the 2006 Conference on Pattern Languages of Programs - PLoP '06, 1–9. <https://doi.org/10.1145/1415472.1415486>
- [Sadicoff05] Sadicoff, M., Larrondo-Petrie, M. M., & Fernández, E. B. (2005). *Privacy-aware network client pattern*. Proceedings of the Conference on Pattern Languages of Programs - PLoP.
- [Schmidt06] Schmidt, D. C. (2006). *Model-Driven Engineering*. In IEEE Computer (Vol. 39, Issue 2).
- [Schumacher03] Schumacher, M. (2003). *Security Engineering with Patterns: Origins, Theoretical Model, and New Applications*. Springer.
<https://doi.org/10.1007/b11930>
- [Schümmer04] Schümmer, T. (2004). *The public privacy - patterns for filtering personal information in collaborative systems*. Proceedings of CHI Workshop on Human-Computer-Human-Interaction Patterns, 1–35.

- [Senarath18] Senarath, A., & Arachchilage, N. A. G. (2018). *Why developers cannot embed privacy into software systems? An empirical investigation*. Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018 - EASE'18, 211–216.
<https://doi.org/10.1145/3210459.3210484>
- [Senarath19] Senarath, A., Grobler, M., & Arachchilage, N. A. G. (2019). *Will They Use It or Not? Investigating Software Developers' Intention to Follow Privacy Engineering Methodologies*. ACM Transactions on Privacy and Security, 22(4), 1–30. <https://doi.org/10.1145/3364224>
- [Shen11] Shen, Y., & Pearson, S. (2011). *Privacy enhancing technologies: A review*. In HP Laboratories Technical Report (Issue 113). Springer, Berlin, Heidelberg. https://doi.org/10.1007/10929179_51
- [Shostack14] Shostack, A. (2014). *Threat Modeling: Designing for Security*. John Wiley & Sons.
- [Sjøberg05] Sjøberg, D. I. K., Hannay, J. E., Hansen, O., Kampenes, V. B., Karahasanović, A., Liborg, N. K., & Rekdal, A. C. (2005). *A survey of controlled experiments in software engineering*. IEEE Transactions on Software Engineering, 31(9), 733–753.
<https://doi.org/10.1109/TSE.2005.97>
- [Solove06] Solove, D. J. (2006). *A taxonomy of privacy*. In University of Pennsylvania Law Review (Vol. 154, Issue 3, pp. 477–564). <https://doi.org/10.2307/40041279>
- [Sommerville11] Sommerville, I. (2011). *Software engineering*. Pearson.
- [Such18] Such, J. M., & Criado, N. (2018). *Multiparty privacy in social media*. Communications of the ACM, 61(8), 74–81. <https://doi.org/10.1145/3208039>
- [Suphakul17] Suphakul, T., & Senivongse, T. (2017). *Development of privacy design patterns based on privacy principles and UML*. 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 369–375. <https://doi.org/10.1109/SNPD.2017.8022748>
- [ThomsonR18] Thomson Reuters. (2018). *InCites essential science indicators: Field baselines*.
<http://esi.incites.thomsonreuters.com/BaselineAction.action>
- [TNSPS16] TNS Political & Social. (2016). *Flash Eurobarometer 443: Report e-Privacy*.

- [USDHEW73] U.S. Dep't. of Health, Education and Welfare, S. A. C. (1973). *On Automated Personal Data Systems, Records, computers, and the Rights of Citizens* viii.
https://www.epic.org/privacy/consumer/code_fair_info.html
- [UnitedN48] United Nations. (1948). *Universal Declaration of Human Rights*. Retrieved April 22, 2020, from <https://www.un.org/en/universal-declaration-human-rights/index.html>
- [VanRest14] van Rest, J., Boonstra, D., Everts, M., van Rijn, M., & van Paassen, R. (2014). *Designing privacy-by-design*. In B. Preneel & D. Ikonou (Eds.), *Privacy Technologies and Policy*. APF 2012. Lecture Notes in Computer Science (Vol. 8319, pp. 55–72). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-54069-1_4
- [Viola18] Viola, R. (2018). *European Dimension of the Digital Economy / Shaping Europe's digital future*. European Commission.
<https://ec.europa.eu/digital-single-market/en/blogposts/european-dimension-digital-economy>
- [Wagner18] Wagner, I., & Eckhoff, D. (2018). *Technical privacy metrics: A systematic survey*. In *ACM Computing Surveys* (Vol. 51, Issue 3). Association for Computing Machinery.
<https://doi.org/10.1145/3168389>
- [Washizaki17] Washizaki, H. (2017). *Security patterns: Research direction, metamodel, application and verification*. 2017 International Workshop on Big Data and Information Security (IWBIS), 1–4.
<https://doi.org/10.1109/IWBIS.2017.8275094>
- [Wieringa06] Wieringa, Roel, Maiden, N., Mead, N., & Rolland, C. (2006). *Requirements engineering paper classification and evaluation criteria: A proposal and a discussion*. *Requirements Engineering*, 11(1), 102–107. <https://doi.org/10.1007/s00766-005-0021-6>
- [Wohlin12] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. In *Experimentation in Software Engineering*.
<https://doi.org/10.1007/978-3-642-29044-2>
- [Xia18] Xia, T., Washizaki, H., Kato, T., Kaiya, H., Ogata, S., Fernandez, E. B., Kanuka, H., Yoshino, M., Yamamoto, D., Okubo, T., Yoshioka, N., & Hazeyama, A. (2018). *Cloud Security and Privacy Metamodel - Metamodel for Security and Privacy Knowledge in Cloud Services*. *Proceedings of the 6th International Conference on Model-Driven Engineering and Software Development*, 379–386.
<https://doi.org/10.5220/0006600403790386>

- [Yoder98] Yoder, J., & Barcalow, J. (1998). *Architectural patterns for enabling application security*. Proceedings of PLoP 1997, 51, 31.
- [Yoshioka08] Yoshioka, N., Washizaki, H., & Maruyama, K. (2008). *A survey on security patterns*. In Progress in Informatics (Issue 5, pp. 35–48). <https://doi.org/10.2201/NiiPi.2008.5.5>
- [Young10] Young, J. D., & Antón, A. I. (2010). *A method for identifying software requirements based on policy commitments*. 18th IEEE International Requirements Engineering Conference, 47–56. <https://doi.org/10.1109/RE.2010.17>
- [Yskout15] Yskout, K., Scandariato, R., & Joosen, W. (2015). *Do security patterns really help designers?* Proceedings - International Conference on Software Engineering, 1, 292–302. <https://doi.org/10.1109/ICSE.2015.49>
- [Yskout12] Yskout, K., Scandariato, R., & Joosen, W. (2012). *Does organizing security patterns focus architectural choices?* Proceedings - International Conference on Software Engineering, 617–627.
- [Yskout18] Yskout, K., Wuyts, K., Van Landuyt, D., Scandariato, R., & Joosen, W. (2018). *Empirical Research on Security and Privacy by Design*. In Empirical Research for Software Security (pp. 1–46). CRC Press. <https://doi.org/10.1201/9781315154855-1>
- [Yuan10] Yuan, X., Jiang, K., Murthy, S., Jones, J., & Yu, H. (2010). *Teaching security management with case studies: Experiences and evaluation*. Journal on Education, Informatics and Cybernetics, 2, 25–30.
- [Zhang12] Zhang, C., & Budgen, D. (2012). *What do we know about the effectiveness of software design patterns?* In IEEE Transactions on Software Engineering (Vol. 38, Issue 5, pp. 1213–1231). <https://doi.org/10.1109/TSE.2011.79>

Datos bibliográficos de artículos incluidos en el estudio sistemático (Capítulo 3)

Tabla A.1: Datos bibliográficos de artículos incluidos en el estudio sistemático

Id	Artículo
2	M. Hansen, “Top 10 mistakes in system design from a privacy perspective and privacy protection goals” 2012.
9	A. Squicciarini et al., “Preventing information leakage from indexing in the cloud” 2010.
12	G. Iachello and G. D. Abowd, “Privacy and proportionality: Adapting legal evaluation techniques to inform design in ubiquitous computing” 2005.
100	M. Kost, et al, “Privacy verification using ontologies” 2011.
138	G. Karjoth et al., “Privacy-enabled services for enterprises” 2002.
162	S. Moncrieff et al., “A Framework for the design of privacy preserving pervasive healthcare” 2009.
167	M. Hafiz, “A collection of privacy design patterns”, 2006.
182	H. S. Fhom et al., “Towards a holistic privacy engineering approach for smart grid systems”, 2011.
219	C. Kalloniatis et al., “Addressing privacy requirements in system design: The PriS method”, 2008.
221	C. Kalloniatis et al., “Using privacy process patterns for incorporating privacy requirements into the system design process”, 2007.
255	D. Kotz et al., “A privacy framework for mobile health and home-care systems,”, 2009.
263	J. I. Hong et al., “Privacy risk models for designing privacy-sensitive ubiquitous computing systems” 2004.
289	S. Pearson and Y. Shen, “Context-aware privacy design pattern selection,”, 2010.
313	D. Le Métayer and S. Monteleone, “Automated consent through privacy agents: Legal requirements and technical architecture”, 2009.
372	J. I. Hong and J. A. Landay, “An architecture for privacy-sensitive ubiquitous computing”, 2004.
383	C. Bier and E. Krempel, “Common privacy patterns in video surveillance and smart energy”, 2012.
384	C. Kalloniatis et al., “Dealing with privacy issues during the system design process,”, 2005.
418	S. Pearson and A. Benameur, “A decision support system for design for privacy”, 2011.
423	B. Friedman et al., “Development of a privacy addendum for open source licenses: Value sensitive design in industry”, 2006.
577	E. Kavakli et al., “Protecting privacy in system design: the electronic voting case”, 2007.
701	S. Wicker and R. Thomas, “A privacy-aware architecture for demand response systems”, 2011.
725	J.-S. Lee and U. Chandra, “Mobile phone-to-phone personal context sharing”, 2009.
746	S. Romanosky et al., “Privacy patterns for online interactions”, 2006.
958	A. Molina-Markham et al., “Private memoirs of a smart meter,”, 2010.
983	P. W. L. Fong et al., “A privacy preservation model for facebook-style social net. systems”, 2009.
996	A. Lampinen <i>et al.</i> , “We’re in it together: interpersonal management of disclosure in social network services”, 2011.
1023	I. Ciuciu <i>et al.</i> , “Semantic support for security-annotated business process models,”, 2011.
1726	N. Notario <i>et al.</i> , “Pripare: A new vision on engineering privacy and security by design,”, 2014.
1728	M. Colesky <i>et al.</i> , “A critical analysis of privacy design strategies”, 2016.
1736	T. Antignac <i>et al.</i> , “A privacy-aware conceptual model for handling personal data,”, 2016.
1740	D. Feth <i>et al.</i> , “A user-centered model for usable security and privacy,”, 2017.
1751	M. Degeling <i>et al.</i> , “Privacy by socio-technical design: A collaborative approach for privacy friendly system design,”, 2016.

1762	E. Luger <i>et al.</i> , “Playing the legal card: Using ideation cards to raise data protection issues within the design process,” 2015.
1781	B. Berendt and S. Preibusch, “Toward accountable discrimination-aware data mining: The importance of keeping the human in the loop—and under the looking glass,” 2017.
1794	M. Aljohani <i>et al.</i> , “Proposed privacy patterns for privacy preserving healthcare systems in accord with Nova Scotia’s Personal Health Information Act,” 2016.
1798	J. van Rest <i>et al.</i> , “Designing privacy-by-design,” 2014.
1804	T. Suphakul and T. Senivongse, “Development of privacy design patterns based on privacy principles and UML,” 2017.
1807	O. Drozd, “Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process,” 2015.
1808	V. Diamantopoulou <i>et al.</i> , “Supporting privacy by design using privacy process patterns,” 2017.
1825	N. Notario <i>et al.</i> , “PRIPARE: Integrating privacy best practices into a privacy engineering methodology,” 2015.
1837	M. H. Davis <i>et al.</i> , “A cybermodel for privacy by design: Building privacy protection into consumer electronics,” 2015.
1847	J.-H. Hoepman, “Privacy design strategies,” 2014.
1869	M. Ataei and C. Kray, “Ephemerality is the new black: A novel perspective on location data management and location privacy in LBS,” 2017.
1873	S. Thiebes <i>et al.</i> , “One size does not fit all: Information security and information privacy for genomic cloud services,” 2016.
1882	M. Aljohani <i>et al.</i> , “Participatory design research to understand the legal and technological perspectives in designing health information technology,” 2017.
1926	T. Länger <i>et al.</i> , “Selected cloud security patterns to improve end user security and privacy in public clouds,” 2016.
1935	A. S. Ahmadian, S. Peldszus, Q. Ramadan, and J. Jürjens, “Model-based privacy and security analysis with CARiSMA,” 2017.
1996	N. Argyropoulos <i>et al.</i> , “Incorporating privacy patterns into semi-automatic business process derivation,” 2016.
2031	M. Hafiz, “A pattern language for developing privacy enhancing technologies,” 2013.
2126	V. Diamantopoulou, N. Argyropoulos, C. Kalloniatis, and S. Gritzalis, “Supporting the design of privacy-aware business processes via privacy process patterns,” 2017.
2198	A. Kung, “PEARs: Privacy enhancing architectures,” 2014.
2266	D. Le Métayer, “Privacy by design: a formal framework for the analysis of architectural choices,” 2013.
2272	J. C. Caiza, <i>et al.</i> , “Organizing design patterns for privacy: A taxonomy of types of relations,” 2017.
2283	K. Katkalov <i>et al.</i> , “Modeling information flow properties with UML,” 2015.
2298	T. Basso <i>et al.</i> , “Requirements, design and evaluation of a privacy reference architecture for web applications and services,” 2015.
2309	Z. Ma <i>et al.</i> , “Towards a multidisciplinary framework to include privacy in the design of video surveillance systems,” 2014.
2537	A. Mohaien <i>et al.</i> , “Secure encounter-based mobile social networks: Requirements, designs, and tradeoffs,” 2013.
2566	P. Maret <i>et al.</i> , “Data model for health telemonitoring and persuasive system design,” 2017.
2709	A. Tsirtsis <i>et al.</i> , “Cyber security risks for minors: A taxonomy and a software architecture,” 2016.
2731	J. Wolff, “Application-layer design patterns for accountable–anonymous online identities,” 2013.
2794	A. Senarath <i>et al.</i> , “Designing privacy for you: A practical approach for user-centric privacy,” 2017.
2867	L. Diver and B. Schafer, “Opening the black box: Petri nets and privacy by design,” 2017.
2899	V. Beltran <i>et al.</i> , “An ARM-compliant architecture for user privacy in smart cities: SMARTIE—quality by design in the IoT,” 2017.
2932	E. S. Lara <i>et al.</i> , “Enhancing privacy notice applications through interaction design,” 2016.

-
- 3014 M. Nitzlnader *et al.*, “Architecture for an advanced biomedical collaboration domain for the European paediatric cancer research community (ABCD-4-E)”, 2015.
-
- 3235 P. Gope and R. Amin, “A novel reference security model with the situation based access policy for accessing EPHR data,”, 2016.
-

Artículos sobre la aplicación de patrones de seguridad (Capítulo 5)

Tabla A.2: Artículos sobre la aplicación de patrones de seguridad

Artículo
N. Yoshioka et. al, "A survey on security patterns", 2008.
A. Wiesauer and J. Sametinger, "A security design pattern taxonomy based on attack patterns: Findings of a systematic literature review", 2009.
E. Fernandez, "Security patterns and a methodology to apply them", 2009.
H. Washizaki et al., "Improving the classification of security patterns", 2009.
A. Mourad et al., "A novel approach for the development and deployment of security patterns", 2010.
R. Ortiz et al., "Applicability of security patterns", 2010.
Y. Shiroma et al., "Model-driven security patterns application based on dependences among patterns", 2010.
A. Mourad et al., "New Approach Targeting Security Patterns Development and Deployment", 2011.
R. Bouaziz and B. Coulette, "Applying security patterns for component based applications using UML profile", 2012.
A. Uzunov et al., "Securing distributed systems using patterns: A survey", 2012.
K. Yskout et al., "Does organizing security patterns focus architectural choices?", 2012.
R. Bouaziz et al., "An engineering process for security patterns application in component based models", 2013.
T. Kobashi et al., "Validating security design patterns application using model testing", 2013.
M. Bunke, "On the description of software security patterns", 2014.
M. Yoshizawa et al., "Verifying implementation of security design patterns using a test template", 2014.
R. Bouaziz et al., "An approach for security patterns application in component based models", 2014.
T. Li and J. Mylopoulos, "Modeling and applying security patterns using contextual goal models", 2014.
E. Fernandez, "Building secure web systems architectures using security patterns", 2015.
M. Bunke, "Software-security patterns: Degree of maturity", 2015.
R. Bouaziz and S. Kammoun, "A decision support map for security patterns application", 2015.
Y. Ito et al., "Systematic Mapping of Security Patterns Research", 2015.
H. Guan et al., "An ontology-based approach to security pattern selection", 2016.
M. Yoshizawa et al., "Implementation support of security design patterns using test templates", 2016.
P. Ponde et al., "An analytical study of security patterns", 2016.
R. Bouaziz and S. Kammoun, "SCRISTUDIO: A security pattern integration tool", 2016.
X. He and Y. Fu, "Modeling and analyzing security patterns using high level petri nets", 2016.
B. B. Mayvan et al., "The state of the art on design patterns: A systematic mapping of the literature", 2017.
H. Washizaki, "Security patterns: Research direction, metamodel, application and verification", 2017.
P.S. Ponde et al., "Knowledge representation of security design pattern landscape using formal concept analysis", 2017.

Patrones de la estrategia Informar

Tabla A.3: Patrones de privacidad de la táctica *explicar*

Patrón	Resumen
<i>Cliente de red respetuoso con la privacidad</i>	Convierta automáticamente las políticas de privacidad a un formato estandarizado y de fácil lectura sobre un canal seguro.
<i>Código de colores de privacidad</i>	Proporcione elementos visuales con colores estandarizados sobre las preferencias y políticas de privacidad para ayudar a transmitir más rápidamente la información a los usuarios.
<i>Contraseñas seguras informadas</i>	Asegúrese de que los usuarios mantengan hábitos de autenticación buenos fomentando su concienciación y comprensión.
<i>Diseño de políticas por capas</i>	Fomente la comprensión de las políticas de privacidad proveyendo información en capas, sucesivamente más concisa y resumida.
<i>Etiquetas de privacidad</i>	Estandarice etiquetas de advertencia similares a las de información nutricional para informar rápidamente sobre las preferencias y políticas de privacidad.
<i>Iconos de privacidad apropiados</i>	Utilice iconos consistentes que transmitir su significado de manera fiable y sin dejar espacio para malentendidos.
<i>Iconos para políticas de privacidad</i>	Los iconos son capaces de transmitir información más rápidamente que un documento y, por lo tanto, son una forma útil de mejorar las políticas.
<i>Información y retroalimentación de impacto</i>	Proporcionar información acerca de a quiénes los usuarios revelarán su información para configurar su privacidad antes de publicar realmente esa información.
<i>Redacción consciente sobre privacidad</i>	Asegúrese de que el contenido de la información relacionado a la privacidad que se proporciona al usuario se redacte cuidadosamente, manteniendo su atención y comprensión.
<i>Selección informada de credenciales</i>	Asegúrese que los usuarios estén informados de las consecuencias potenciales a la privacidad al compartir diversos datos de autenticación.

Tabla A.4: Patrones de privacidad de la táctica *notificar*

Patrón	Resumen
<i>Actividades inusuales</i>	Prevenga accesos sospechosos a los datos de los usuarios a través de alertas y autenticación con múltiples factores.
<i>Aviso ambiental</i>	Provee avisos no intrusivos, no modales y continuos cuando se acceden a los datos personales para aumentar la conciencia de seguimiento en tiempo real.
<i>Aviso asíncrono</i>	Proporciona proactivamente avisos continuos de accesos recurrentes a los datos personales, incluyendo el rastreo, el almacenamiento o la redistribución de ellos.
<i>Consentimiento implícito informado</i>	Los controladores deben proveer avisos inevitables a los usuarios sobre el consentimiento implícito al usar un servicio.
<i>Notificación de violación de datos</i>	Garantice que el acceso y el procesamiento no autorizados de datos personales se detecten y se comuniquen a la autoridad de supervisión y a los usuarios afectados sin ninguna demora.
<i>Prevenir errores o reducir su impacto</i>	Prevenga la divulgación automática accidental de información personal.
<i>Visualización dinámica de política de privacidad</i>	Provea información normativa contextual sobre la naturaleza y los riesgos de la divulgación de los datos a través de globos emergentes.

Tabla A.5: Patrones de privacidad de la táctica *proveer*

Patrón	Resumen
<i>Asimetría de información mínima</i>	Evite que los usuarios se vean privados de sus derechos por su falta de familiaridad con las políticas, los riesgos potenciales y su actuación en el proceso.
<i>Aumento de la conciencia de la agregación de información</i>	Informe a los usuarios sobre eventos potencialmente identificables en la agregación de información para evitar que pongan en peligro su privacidad sin saberlo.
<i>Concienciación</i>	Informe a los usuarios sobre el grado de visibilidad de sus datos y lo que puede derivarse de ellos. Esto les permite reconsiderar si se sienten cómodos al compartir, y tomar medidas si lo desean.
<i>Espejos de privacidad</i>	Proporcione al usuario de un sistema información de alto nivel sobre los datos personales que el sistema conoce, el acceso que tienen los demás y sobre los datos personales que se pueden deducir.
<i>Evaluación de la confianza de los Servicios</i>	Use realces visuales proporcionado por una autoridad que indique el grado en que se cumplen determinados criterios de privacidad. Debe colocarse claramente y ser fácil de encontrar, con enlaces a información adicional.
<i>Panel de Concienciación de la Privacidad</i>	Establecer la conciencia de los usuarios sobre los riesgos inherentes a la divulgación de sus datos, ya sea al propio controlador o a otros usuarios.
<i>Panel de privacidad</i>	Un panel de información sobre la privacidad puede proporcionar resúmenes de los datos personales recogidos o procesados de un usuario determinado.
<i>Plataforma de preferencias de privacidad</i>	Utilice políticas de privacidad que usen elementos de datos y un vocabulario estandarizados y extensibles, los cuales deben ser conocidos por los agentes de usuario a fin de facilitar su lectura y revisión.
<i>Retroalimentación apropiada de privacidad</i>	Brinde retroalimentación sobre privacidad al usuario, especialmente en lo que concierne a lo que es monitoreado, accedido, y por quién.
<i>Panel de datos personales</i>	Para que los usuarios vean la información que un controlador tiene sobre ellos, se les puede proveer, previa solicitud, un resumen tabular detallado de esos datos.
<i>Términos y condiciones abreviados</i>	Permite al usuario entender mejor los Términos y Condiciones de un sistema a través de resúmenes. Los elementos más importantes se condensan en un resumen más conciso.
<i>Visualización de coincidencias de políticas</i>	Permita que los usuarios especifiquen sus preferencias de privacidad y muéstreles de forma no intrusiva las divergencias de las políticas con ellas.
<i>Visualización de política de privacidad</i>	Proporcione al usuario información sobre qué información, por qué y quién la solicita. Debe utilizarse siempre que se requieran datos personales del usuario.
<i>¿Quién está escuchando?</i>	Informe a los usuarios de los contenidos a los que otros usuarios o personas no autenticadas también pueden acceder.

Patrones de la estrategia Controlar

Tabla A.6: Patrones de privacidad de la táctica *actualizar*

Patrón	Resumen
<i>Difusión activa de presencia</i>	Los usuarios pueden elegir de manera dinámica cuándo quieren proporcionar actualizaciones sobre su presencia, de tal manera que puedan incrementar la relevancia y el control sobre lo compartido.
<i>Nivel razonable de control</i>	Deje que los usuarios compartan selectivamente y pongan a disposición información específica a grupos o usuarios predefinidos.
<i>Reciprocidad</i>	Permita a los usuarios beneficiarse de acuerdo con las contribuciones que hagan.

Tabla A.7: Patrones de privacidad de la táctica *consentir*

Patrón	Resumen
<i>Consentimiento informado para transacciones web</i>	Este patrón describe la forma en que los controladores pueden informar a los usuarios cuando tienen la intención de recopilar o utilizar sus datos personales.
<i>Consentimiento legal</i>	Un elemento crucial en la protección de la privacidad es garantizar que todo procesamiento delicado vaya precedido de la obtención de un consentimiento libremente otorgado, informado, específico y explícito.
<i>Externalización [con consentimiento]</i>	El controlador debe lograr un consentimiento adicional, específico, informado, explícito y libremente otorgado antes de externalizar datos a terceros.
<i>Firma de un acuerdo para resolver la falta de confianza en un contexto de uso de datos privados</i>	Los servicios que ofrece un controlador, previo a ser usados, pueden recurrir a la firma de contratos que estipulen las obligaciones y los fines del procesamiento para obtener el consentimiento del usuario. Esto garantiza que los usuarios puedan confiar en el controlador, ya que éste está obligado a cumplir el contrato que firma.
<i>Obtención de consentimiento explícito</i>	Los controladores deben obtener el consentimiento de forma voluntaria y específica cuando procesen de alguna manera los datos personales de sus usuarios.
<i>Punto único de contacto</i>	El punto único de contacto es una autoridad de seguridad que protege la privacidad y seguridad de datos sensibles almacenados online validando la autoría de peticiones y asegurando los canales de comunicación.

Tabla A.8: Patrones de privacidad de la táctica *retirar*

Patrón	Resumen
<i>Disociación de contenido e información de localización</i>	Permita a los usuarios configurar de manera retroactiva los ajustes de privacidad para la información de localización considerando el contexto.
<i>Enmascarar</i>	Deje a los usuarios filtrar toda o parte de la información personal que podrían proporcionar a un servicio.

Tabla A.9: Patrones de privacidad de la táctica consentimiento *elegir*

Patrón	Resumen
<i>Apoyo a la divulgación selectiva</i>	Este patrón recomienda que los servicios apoyen la divulgación selectiva de datos, adaptando la funcionalidad de los servicios o productos para que funcionen con el nivel de datos que el usuario se sienta cómodo compartiendo.
<i>Control de acceso selectivo</i>	Permita a los usuarios especificar quién puede acceder al contenido que ellos generan, durante y posteriormente al envío.
<i>Desalentar estrategias generales</i>	Brinde a los usuarios la posibilidad de definir un determinado nivel de privacidad, a partir de una serie de opciones, cada vez que comparten un contenido.
<i>Enlace privado</i>	Habilite la compartición y re-compartición de un enlace privado que permita acceder a datos personales evitando un proceso de autenticación complicado.
<i>Habilitar/deshabilitar funciones</i>	Permita que los usuarios decidan de manera granular las funciones a las que dan su consentimiento antes de usarlas.
<i>Lista de amigos</i>	Por defecto, de entre todas las conexiones sociales, establezca un círculo de confianza definido por el usuario. Permita expandir este círculo o crear otros nuevos, basados en los miembros existentes.
<i>Negociación de la política de privacidad</i>	Permita a los usuarios construir sus preferencias de privacidad desde una configuración muy restrictiva por defecto, a través de dejarles optar por entrar o por abandonar las opciones.
<i>Participación incentivada</i>	Los usuarios están más dispuestos a hacer contribuciones cuando pueden hacerlo sin filtrar datos personales, o si perciben un intercambio igual o mayor en valor, ya sea monetario o social.
<i>Restituir</i>	Brinde algunos beneficios a los usuarios a cambio de proporcionar información.