Universidad
Rey Juan Carlos

TESIS DOCTORAL

Estrategias de privacidad en esquemas de gestión de identidades y accesos

Autor:
Carlos Alberto Villarán Núñez

Directora:
Dra. Marta Beltrán Pardo

Programa de doctorado en Tecnologías de la Información y las Comunicaciones

Escuela Internacional de Doctorado

2022

DEPARTAMENTO DE CIENCIAS
DE LA COMPUTACIÓN,
ARQUITECTURA DE
COMPUTADORES, LENGUAJES Y
SISTEMAS INFORMÁTICOS Y
ESTADÍSTICA E INVESTIGACIÓN
OPERATIVA

Estrategias de privacidad en esquemas de gestión de identidades y accesos

TESIS DOCTORAL

Autor: Carlos Alberto Villarán Núñez
Graduado en Ingeniería Telemática

Directora: Dra. Marta Beltrán Pardo
Doctora en Informática

2022

Gracias Marta por tu dedicación, tiempo y paciencia.
Sin tu ayuda no estaría donde estoy.

Gracias a mi padre, madre y Mané por vuestro apoyo y motivación para seguir estudiando.

Gracias Almudena por escuchar mis acordes y darle un toque funky a mi blues del delta del Guadalquivir.

Gracias a mis amigos, sin vosotros el camino hubiera sido mucho más difícil.

Gracias Música, pensadero y quebradero.

ACTA EST FABULA

Índice general

Índice de figuras

Índice de tablas

Capítulo 1

Introducción

1.1. Contexto de la investigación

La gestión de identidades y accesos (IAM o *Identity and Access Management* permite saber en el mundo digital quién es el usuario y qué es lo que puede hacer en un sistema o servicio. La gestión de identidad se encarga de verificar la identidad del individuo y la gestión de accesos de las acciones que tiene permitidas. Para lograr este objetivo se han utilizado a lo largo del tiempo diferentes metodologías, algoritmos, protocolos y herramientas.

Cuando un individuo interactúa con un sistema o servicio digital que necesita conocer qué usuario es, éste tiene que tener una identidad digital. En este momento comienza el ciclo de vida de la identidad digital. La creación de la identidad digital se corresponde con una traslación total o parcial de la identidad de la persona (entendida como un conjunto de atributos que permiten identificarla) al mundo digital. Además, se le asignan los accesos que le han sido autorizados. Durante el tiempo que el individuo hace uso del sistema o servicio digital, es posible que el usuario necesite actualizar su identidad, lo que se corresponde con la modificación de la identidad (de los atributos asociados) o con que se concedan más o menos privilegios de los que ya tenía, la modificación de los accesos. La última parte del ciclo de vida de la identidad se corresponde con el borrado de ésta una vez que ya no es necesaria, junto con aquellas autorizaciones que tenga concedidas.

En la gestión de la identidad, para que el usuario pueda demostrar que es quien dice ser, se necesita que lo haga mediante algo que conoce (como una contraseña), algo que tiene (como un teléfono móvil) y/o algo que es (como el uso de la huella digital). Cuantos más factores se necesiten para demostrar la identidad del usuario más podremos asegurar que, efectivamente, es quien dice ser. Además, se pueden modular para que no siempre se requieran los

mismos factores dependiendo de la acción que quiera realizar. Por ejemplo, para acceder a una aplicación bancaria puede que sólo sea necesario solicitar la contraseña (algo que sabe) pero para hacer una transferencia se necesite, además de la contraseña, la huella digital (algo que es).

Por otra parte, en la gestión de accesos, se autoriza a los usuarios a realizar determinadas acciones que tienen permitidas una vez han demostrado que son quiénes dicen ser. Por ejemplo, lectura, modificación o borrado.

En resumen, la gestión de identidades y accesos debe garantizar la capacidad de resolver los problemas de Identificación, Autenticación, Autorización y Auditoría (IAAA) de manera que se ofrezcan los niveles adecuados de confidencialidad, disponibilidad, integridad y no repudio en el uso de recursos, aplicaciones o servicios digitales.

De hecho, se trata de una de las disciplinas de la ciberseguridad que más ha evolucionado con los años y que más importancia ha cobrado, dados los requisitos tan exigentes que plantean en este sentido el Internet, no sólo de los contenidos, sino de las personas (redes sociales), de los servicios (*cloud computing*) o de las cosas (*Internet of Things* o entornos inteligentes). Esta importancia ha provocado que surjan multitud de modelos de negocio alrededor de la gestión de estas identidades digitales y que se pueda hablar de una economía de la identidad relacionada con la economía de los datos o incluso de la vigilancia. Y es que la gestión de las identidades y de los accesos de todos los usuarios actuales de servicios digitales genera una cantidad ingente de datos que pueden llegar a ser muy valiosos para diferentes empresas y organizaciones. Y es imprescindible exigir que los recopilen, gestionen y utilicen de manera ética.

Estos datos sobre los usuarios pueden emplearse con fines muy diversos, desde ofrecer recomendaciones, soporte a la decisión o personalización de los servicios, hasta la detección de enfermedades o la prevención del fraude. Sin embargo, en algunos casos puede que tengan fines que el usuario desconozca o con los que el usuario no esté de acuerdo. Como el rastreo de las páginas por las que navega o el seguimiento de sus actividades cotidianas.

Hay que recordar que la privacidad se recoge en la Declaración Universal de Derechos Humanos, artículo 12, o en la Constitución Española, en su artículo 18. En el ámbito digital, la privacidad de los usuarios ha ido evolucionando con el paso del tiempo intentándose adaptar al cambio continuo en el uso de la tecnología en su esfera personal y profesional. Los riesgos para la privacidad en el mundo digital son muy diversos teniendo en cuenta la cantidad de datos que se generan y almacenan cuando se accede a distintos tipos de recursos, servicios y aplicaciones. Y todos estos datos, incluidos los datos personales y sensibles, como pueden ser los datos médicos o bancarios, están expuestos a posibles robos, modificaciones no deseadas, tratamientos o

transferencias no autorizados, etcétera.

Para ayudar a proteger los datos personales de sus ciudadanos, la Unión Europea se ha dotado de un Reglamento General de Protección de Datos (RGPD), que entró en vigor en el año 2018. En este reglamento se persigue la protección de los datos personales de los ciudadanos europeos y, para ello, se expresan de manera explícita sus derechos, así como las obligaciones de los responsables de tratar estos datos. Este reglamento establece por primera vez el concepto de responsabilidad proactiva, así como sanciones que implican altas cuantías para aquellas empresas u organismos que no lo cumplan. Está siendo una referencia para otras regiones que desean regular la protección de los datos personales de sus ciudadanos y, por tanto, su privacidad, derechos y libertades.

La privacidad en el mundo digital ha ido tomando una mayor relevancia con el paso del tiempo y, cada vez de forma más habitual, los servicios digitales incorporan capacidades relacionadas con esta privacidad para ofrecer una mayor calidad de servicio y experiencia para sus usuarios. Lo ideal es que esta privacidad se incorpore desde el diseño de los servicios, considerando los principios de la protección de datos personales desde las fases iniciales de su concepción. Lo ideal sería también que fuera una privacidad centrada en el usuario, como sujeto activo que pueda tener el control, y que pueda tomar decisiones como propietario de sus datos.

En relación con la gestión de identidades y accesos, el problema es que muchos de los modelos de negocio antes mencionados son, como mínimo, opacos. En el caso de los modelos federados para la gestión de identidades como SAML, OAuth, OpenID Connect o Mobile Connect, objeto de la investigación realizada en esta tesis doctoral, existe un proveedor de identidades que permite que el usuario se autentique en un recurso, aplicación o servicio sin necesidad de tener una cuenta en él.

Este tipo de modelos permiten el uso del *Social Login* o del *Single Sign-On*, tan extendidos en la actualidad. En el primer caso, un usuario puede acceder a la facturación de un vuelo, a un servicio de noticias o al área de cliente en la página web de su operador de telefonía sin necesidad de tener cuentas específicas en los sitios web de cada uno de estos proveedores (la aerolínea, el proveedor del servicio de noticias o el operador de telefonía), sino usando para ello a un proveedor de identidades como Facebook, Google, LinkedIn, Apple o GitHub, por mencionar sólo alguno de estos ejemplos. En el segundo caso, un empleado de una empresa puede acceder a todas las aplicaciones o servicios que necesita para trabajar en su día a día, ofrecidas también por diferentes proveedores, sin necesidad de iniciar sesión en cada una de ellas por separado, sino a través de un único servicio que hace de pasarela de entrada al resto y en el que se autentica una única vez.

¿Por qué las grandes empresas tecnológicas se ofrecen de manera aparentemente gratuita como proveedores de identidad en Internet? En el caso de las soluciones *Single Sign-On*, ¿la única fuente de ingresos que tienen sus proveedores es la suscripción que cobran por sus servicios? ¿o hay otros modelos de negocio que alimentan su cuenta de resultados? Obviamente la respuesta a estas dos preguntas está relacionada con los datos que genera la gestión de identidades y accesos de los usuarios, en muchos casos personales, y que permiten su perfilado, trazado, geolocalización, etc.

En esta tesis doctoral, la investigación se centrará en el modelo *honest but curious* en el que todos los participantes en las federaciones de identidades son legítimos y no se desvían de lo que las especificaciones y protocolos tecnológicos proponen. Pero intentarán obtener todos los datos posibles de los mensajes y comunicaciones legítimamente recibidos dentro de los esquemas IAAA.

1.2. Hipótesis de partida

'Es posible modelar de manera sistemática y exhaustiva las amenazas para la privacidad que supone para los usuarios de Internet el uso del modelo federado para la gestión de identidades y accesos, considerando el modelo "honest but curious" para todos los proveedores de la federación, tanto los de identidades como los de recursos, aplicaciones y servicios. El modelo de amenazas producido servirá como guía para proponer estrategias de privacidad desde el diseño y centradas en los usuarios que se puedan aplicar con las especificaciones federadas actuales sin necesidad de modificarlas significativamente, de manera que se facilite su adopción.'

1.3. Objetivos

Los objetivos generales de esta tesis doctoral surgen de su hipótesis de partida y son dos:

- Modelar las amenazas para la privacidad que implica para los usuarios la utilización de un modelo federado para la gestión de identidades y accesos.
- Proponer y validar estrategias de privacidad desde el diseño y centradas en el usuario que permitan evitar o mitigar un conjunto suficiente de las amenazas identificadas.

Para conseguir estos objetivos generales, que se centran en permitir un tratamiento ético de los datos dentro de la federaciones de identidades y por lo tanto, en contribuir a la sostenibilidad de este modelo en el tiempo, se han planteado los siguientes objetivos específicos:

1. Encontrar una metodología de modelado de amenazas que sea adecuada para identificar de manera exhaustiva y sistemática las amenazas para la privacidad, teniendo en cuenta los impactos para las personas y el ciclo de vida de los datos.

2. Analizar el ciclo de vida de los datos y los flujos que se producen en cualquier modelo federado de gestión de identidades y accesos. Es necesario trabajar de manera genérica para que el modelo que se produzca posteriormente sea válido en cualquier federación de identidades, independientemente de la especificación tecnológica en la que se base.

3. Aplicar la metodología escogida al ciclo de vida analizado para producir el modelo de amenazas requerido.

4. Identificar el subconjunto más amplio posible de amenazas para la privacidad modeladas que puedan evitarse o mitigarse con estrategias de privacidad desde el diseño y centradas en el usuario. Es decir, que se basen de alguna manera en proporcionarle el máximo control posible sobre sus propios datos y en facilitarle el ejercicio de sus derechos relacionados con la protección de datos, así como la toma de decisiones.

5. Proponer estrategias de privacidad desde el diseño y centradas en el usuario que eviten o mitiguen este subconjunto de amenazas identificadas. Estas estrategias deben ser, de nuevo, aplicables con las diferentes especificaciones federadas existentes en la actualidad, y su uso no debe exigir apenas cambios para que sean fáciles de incorporar a infraestructuras en producción.

6. Implementar prototipos que permitan validar y evaluar las estrategias propuestas.

1.4. Metodología

La metodología propuesta para la demostración de la hipótesis de partida planteada en la presente tesis y la consecución de los objetivos descritos se describe a continuación:

- Investigación y análisis en profundidad de los modelos de gestión de identidades y accesos, centrando dicho análisis en el modelo federado mediante las especificaciones e implementaciones más extendidas en la actualidad.

- Investigación y análisis de los mecanismos, protocolos, algoritmos y estrategias de privacidad relevantes para esta investigación. Así como de metodologías para el modelado de amenazas para la privacidad.

- Determinación de posibles amenazas para la privacidad en los modelos federados, priorizando el estudio en el protocolo OpenID Connect/Mobile Connect (por ser el más extendido), mediante una técnica mixta de estudio de investigaciones previas y de experimentación en entornos controlados.

- Diseño de estrategias de privacidad desde el diseño y centradas en los usuarios intentando emplear las capacidades ofrecidas por los proveedores que participan en las federaciones de identidades actuales o incorporando proveedores nuevos, si es necesario, pero sin modificar sustancialmente las especificaciones ya existentes.

- Implementación de las estrategias propuestas mediante el desarrollo de la funcionalidad en código fuente, APIs (*Application Programming Interfaces*) y pruebas de concepto.

- Validación y evaluación de los prototipos producidos, tanto desde el punto de vista de rendimiento como de usabilidad, utilidad, etc.

1.5. Estructura del documento

Además de este capítulo de Introducción, este documento contiene los siguiente capítulos:

- En el capítulo 2 se analiza el estado del arte.

- En el capítulo 3 se presentan los resultados del proceso de modelado de amenazas.

- En el capítulo 4 se propone mitigar las amenazas identificadas en el capítulo anterior con una solución que garantice que los usuarios europeos pueden ejercer los derechos recogidos en el RGPD.

- En el capítulo 5 también se propone la mitigación de amenazas identificadas en el capítulo 3 con un sistema de recomendación de privacidad que ayude a los usuarios en su toma de decisiones.
- En el capítulo 6 se finaliza este documento con la exposición de las conclusiones obtenidas tras la realización de la presente tesis doctoral. Además se identifican las líneas más prometedoras de trabajo futuro.

Capítulo 2

Estado del Arte

En este capítulo se realiza una introducción a los distintos modelos para la gestión de la identidad, en particular, el modelo federado y el centrado en el usuario. Además, se presenta en detalle la especificación OpenID Connect en la que se centra la presente tesis doctoral. A continuación, se analiza en profundidad el concepto de privacidad y se particulariza para los escenarios de gestión de la identidad mediante el modelo federado. Por último, se profundiza en los trabajos previos que se han realizado en este campo y se discuten los motivos que justifican la realización de esta investigación.

2.1. Modelos de control de accesos

Los modelos de control de accesos tienen como finalidad el gobierno del acceso a los recursos de un sistema, es decir, conseguir que únicamente puedan acceder a los recursos los usuarios con autorización para ello. Algunos ejemplos de recursos pueden ser ficheros, servicios, programas o procesos que se ejecutan en el sistema. Los modelos de control de accesos más extendidos son:

- *Discretionary Access Control* (DAC) [1]: este modelo se basa en la definición de políticas de acceso a los recursos. Los usuarios con permisos para definir la política de un recurso (por ejemplo, el dueño del recurso) definen quién tiene acceso al recurso y qué pueden hacer con él. Un ejemplo de DAC es la definición de permisos que se realiza en los sistemas de ficheros UNIX, donde se pueden conceder permisos a usuarios (o grupos de usuarios) de lectura, escritura y ejecución.

- *Mandatory Access Control* (MAC) [2]: en este modelo se etiquetan los recursos a dos niveles. El primero se refiere al nivel de clasificación

del recurso en base a la importancia que tenga para la organización. Algunos ejemplos de este tipo de etiquetas son *secret*, *restricted* o *unclassified*. El segundo se refiere a la categoría en el que se definen los grupos de personas que podrían tener acceso al recurso. Las categorías pueden basarse en el departamento, proyecto, etc. La definición de estas etiquetas están a cargo de un grupo de personas autorizadas para ello, como los administradores del sistema. Para que un usuario pueda acceder a un recurso debe cumplir dos condiciones. Primero tiene que pertenecer a la categoría que tiene permisos de acceso al recurso y, además, tiene que tener la autorización (*clearance)* para poder acceder a recursos del nivel de clasificación del recurso. Este modelo de control de accesos se utiliza normalmente en sistemas militares o críticos y está disponible, por ejemplo, en el sistema operativo SELinux [3].

- *Role-Based Access Control* (RBAC) [4]: en este modelo se crean una serie de roles y los usuarios pertenecen a éstos. Los roles llevan asociados privilegios para el acceso a los recursos, y se pueden definir de forma jerárquica permitiendo tener una mayor flexibilidad en la gestión. Este modelo permite tener una mayor granularidad que DAC, lo que facilita la gestión del control de accesos a los recursos. Un ejemplo representativo del modelo RBAC es el utilizado en Windows para el control de accesos de usuarios, máquinas y procesos [5].

- *Attribute-Based Access Control* (ABAC) [6]: este modelo se caracteriza por utilizar atributos en el control del accesos. Los atributos definen al usuario que quiere acceder a un recurso, el entorno y el contexto en el que lo hace. Por tanto, permite tener una mayor granularidad que RBAC, ya que, para acceder a un recurso, se puede tener en cuenta más de un atributo (se traduce en el modelo anterior en la pertenencia del usuario a un rol). Algunos ejemplos de atributos que se pueden definir para limitar el acceso a un recurso son el departamento al que pertenece un usuario, la geolocalización desde donde se puede acceder o el horario en el que se accede. Para la implementación de ABAC se ha definido el estándar XACML (*eXtensible Access Control Markup Language*) [7].

- *Role-Centric Attribute-Based Access Control* (RABAC) [8]: este modelo pretender solventar problemáticas identificadas en RBAC y ABAC. El modelo RBAC tiene problemas de escalabilidad en cuanto a que la gestión de un gran número de roles es muy compleja. Además, el modelo ABAC también presenta problemas de gestión si se definen un gran número de reglas de acceso a los recursos. Sobre el modelo RBAC se añaden el uso de atributos a usuarios, recursos y una política de filtrado

de permisos (PFP - *Permission Filtering Policy*). La PFP, en base a funciones de filtrado, limita los permisos de roles de usuarios teniendo en cuenta los atributos de los usuarios y recursos. Las funciones de filtrado devuelven un valor booleano ("Verdadero" o "Falso"), y únicamente se tiene acceso al recurso si todas ellas devuelven "Verdadero".

- Otros modelos de control de acceso: además de los modelos definidos, existen otros modelos basados en distintos principios básicos para el control del acceso a los recursos. Algunos ejemplos son *Context Aware Access Control)*, basado en el contexto del acceso al recurso, *Organization-Based Access Control* (OrBAC), basado en la definición de políticas para usuarios, acciones y recursos, o *Risk-Based Access Control*, que evalúa el riesgo que implica el acceso de un usuario a un recurso determinado en un momento determinado para definir las medidas que se deben tomar (dar acceso, limitar acceso, solicitar verificaciones, denegar acceso, etc.).

2.2. Modelos de gestión de identidad

Los modelos de control de accesos resuelven el problema de la autorización, pero antes de llegar a ellos suele ser necesario saber quién es el usuario y verificar que es quien dice ser. El objetivo principal de los modelos de gestión de identidad es la correcta realización del ciclo de vida de las cuentas de usuario, que permiten estas funciones de identificación y autenticación.

Este ciclo de vida se compone de tres fases o pasos principales. La primera fase es la creación de la cuenta (paso 1 de la figura 2.1), donde el usuario introduce los datos necesarios para poder crear un registro en el sistema que gestiona las identidades. La cantidad de datos que se introducen depende en gran medida del modelo concreto del que se trate. Además, durante el tiempo en el que la cuenta existe en el sistema, es susceptible a cambios relativos a la identidad del usuario o a los permisos que éste tiene en el sistema. Esto se corresponde con la segunda fase, la de modificación (paso 2 de la figura 2.1). Una vez que el usuario no necesita la cuenta en el sistema, se procede la última fase, la baja en el mismo y, por tanto, al borrado de la cuenta y de los datos asociados a ella (paso 3 de la figura 2.1).

Cada uno de los modelos de gestión de la identidad, para realizar estas fases, tienen una serie de ventajas e inconvenientes. Dependiendo de la situación, condiciones, requisitos y necesidades, será mejor emplear uno u otro. Según dónde se almacena la identidad del usuario y cómo se gestiona, los modelos de gestión de la identidad se dividen en cuatro grandes gru-

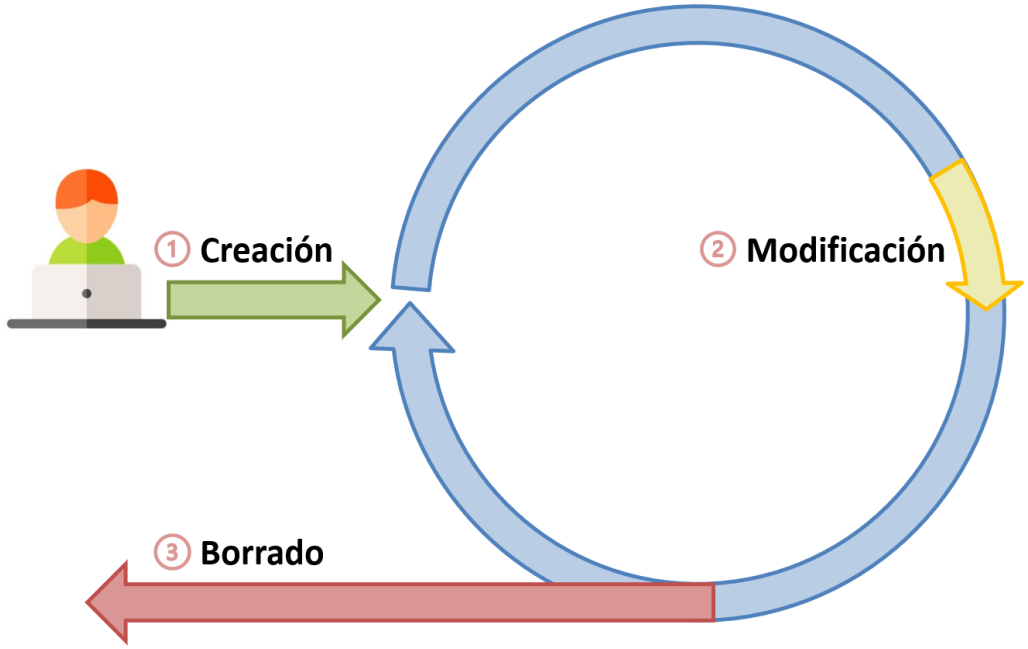

Figura 2.1: Pasos necesarios para la gestión de la identidad de los usuarios

pos [9], [10], [11] (figura 2.2): aislados (o distribuidos o silo), centralizados, federados y centrados en el usuario.

En los modelos aislados cada sistema resuelve por su cuenta la identificación y la autenticación, por lo que la identidad del usuario se gestiona de manera autónoma en cada sistema, sin que se establezca ningún tipo de colaboración o comunicación. Tampoco existe ningún mecanismo que permita compartir esta identidad con otras organizaciones. Por lo tanto, los usuarios tienen que ser los provisionadores de la identidad para cada recurso que quieran utilizar. La gestión de estas identidades es completamente independiente entre recursos y organizaciones, el usuario necesita una identidad diferente en cada uno de ellos.

En los modelos centralizados, por el contrario, la identidad se almacena en un único lugar desde donde se ofrece el servicio de identificación y autenticación. Todos los sistemas delegan la autenticación en esta entidad centralizada. Los usuarios únicamente tienen que provisionar su identidad en el sistema central para poder utilizar los recursos, evitando inconsistencias y la redundancia de las identidades. Sin embargo, en este tipo de modelos existe un punto único de fallo, el sistema central que gestiona las identidades. Además, no tiene por qué ser confiable para todos los recursos que el

usuario quiera utilizar, que pueden pertenecer a organizaciones diferentes con requisitos y necesidades de seguridad diferentes o incluso incompatibles.

En el modelo federado, la identidad del usuario se gestiona desde varios proveedores de identidad (IdP o *Identity Providers*), y los recursos y usuarios hacen uso del que más le conviene en cada caso. De esta forma, se evita el punto único de fallo y se mitigan los problemas de confiabilidad de los recursos en los proveedores de identidad, ya que tienen diferentes opciones entre las que elegir y con las que integrarse.

Por último, en el modelo centrado en el usuario, la identidad la posee el usuario, manteniéndola bajo su control, y pudiéndola utilizar directamente en sus accesos a los recursos. Hay que tener en cuenta que en otros modelos de gestión de identidad el usuario tiene que almacenar los atributos de su identidad (nombre, apellidos, correo electrónico, etc.) en una entidad externa (recursos distribuidos, sistema central, proveedores de identidades). En este caso lo hace en un dispositivo bajo su control, por ejemplo, con el uso de una tarjeta inteligente o de un TPM (*Trusted Platform Module*) de un ordenador o teléfono móvil. Estos atributos se pueden verificar, ya que son emitidos por entidades terceras como organismos públicos, bancos y comercios, de confianza para la entidad verificadora. Además, las soluciones basadas en este modelo son capaces de adaptarse a las necesidades de cada usuario mediante técnicas como la personalización.

Dentro de estos modelos centrados en el usuario destaca la identidad soberana, *self-sovereign identity* (SSI). Las entidades verificadoras de atributos puede que no necesiten el valor concreto del atributo, sino saber si se cumple cierta condición. Por ejemplo, se responderá verdadero o falso si la persona es mayor de edad, pero no se responderá el valor concreto si no es imprescindible.

En la figura 2.2 se pueden observar las diferencias entre cada uno de los modelos. Además, se puede observar dónde se almacena la identidad del usuario para cada caso.

Las siguientes secciones se centran en los modelos federados y los centrados en el usuario, objeto de la investigación realizada en esta tesis doctoral.

2.2.1. Modelo federado para la gestión de identidades

Como ya se ha explicado, los modelos federados, también conocidos en inglés como *Federated Identity Management* (FIM), permiten descentralizar la identidad en varios proveedores de identidad. Además, los recursos accedidos por los usuarios (*Relying Party*, RP) tienen la capacidad de utilizar uno o más IdP estableciendo con ellos una relación de confianza. Cuando se realiza el proceso de autenticación, no solo el usuario se autentica, sino

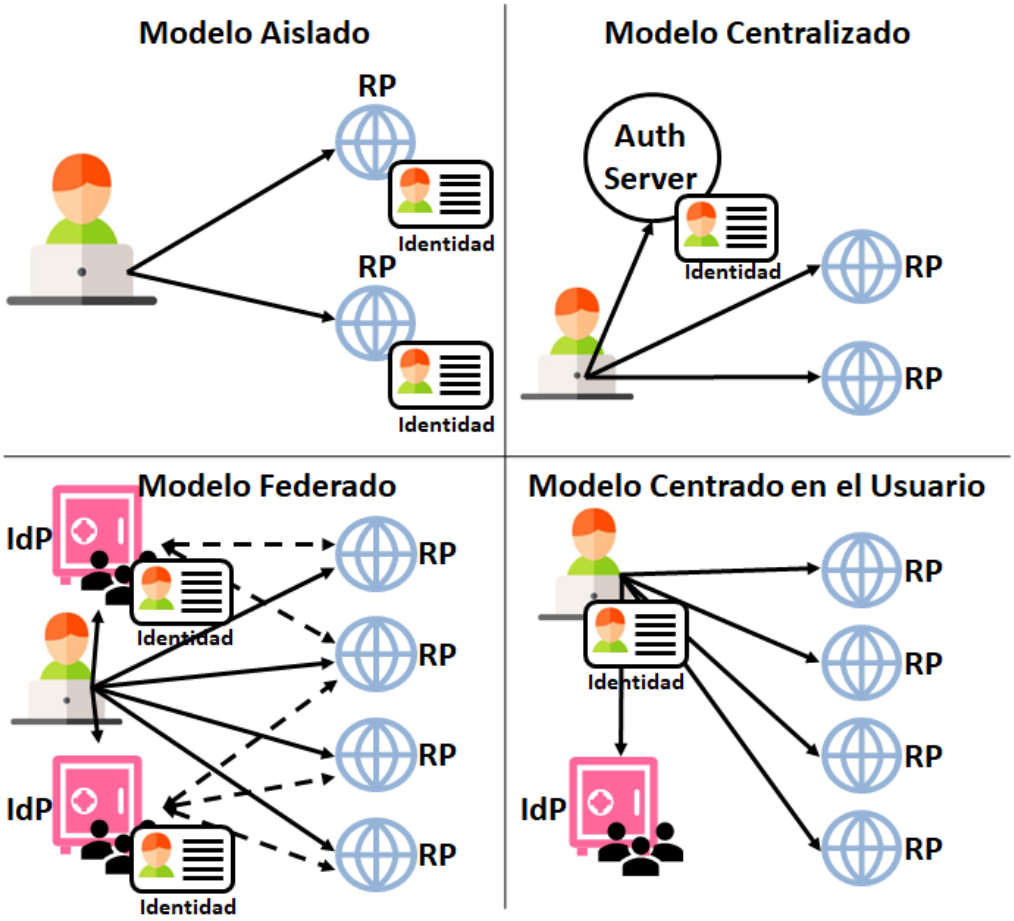

Figura 2.2: Comparativa entre los cuatro modelos de gestión de identidad

que también lo hace el recurso accedido. Además, la segregación de funciones entre la autenticación y autorización con el uso de proveedores, permite la implementación del *Single Sign-On* (SSO). Algunos ejemplos de protocolos de autenticación o autorización federados son SAML 2.0 (*Security Assertion Markup Language* 2.0) [12], OAuth 2.0 [13] y OpenID Connect [14], explicado en mayor detalle a continuación, por ser fundamental para la presente investigación.

Especificación OpenID Connect

OpenID Connect (OIDC) [14] es una especificación que permite la autenticación de un usuario de forma independiente al recurso, aplicación o servicio al que va a acceder. Es decir, se trata de un modelo de autenticación federada

en el que el encargado de realizar esta función es el proveedor de identidad (IdP). En este tipo de modelos, un usuario que quiere acceder a un recurso autenticado de una entidad que ofrece un servicio (RP), es capaz de hacerlo sin crearse una cuenta específica para ello, simplemente autenticándose a través de un IdP. La RP es capaz de comprobar que la autenticación de ese usuario es correcta mediante un *token* de identidad (*ID Token*). También se incluye en la especificación un *token* de acceso (*Access Token*), que gestiona las autorizaciones para el uso de un servicio de una RP. Es decir, OpenID Connect no resuelve solo la identificación y la autenticación, sino también la autorización (el control de accesos).

Con esta especificación se obtienen una serie de ventajas con respecto a los modelos de gestión de identidad tradicionales, en los que cada sistema autentica al usuario de forma independiente. Por un lado, una vez se ha autenticado un usuario en el IdP, puede utilizar múltiples servicios de distintas RP únicamente aceptando las condiciones de la autorización. Esto permite que la experiencia de usuario en el uso de los recursos sea mucho mejor, ya que no tiene que autenticarse en cada una de ellas de manera independiente. Además, permite que el usuario tenga un menor número de cuentas, lo que facilita su manejo, y que la gestión de contraseñas sea más sencilla. Únicamente tiene que conocer la contraseña para autenticarse en el IdP. Como regla general, las contraseñas solamente se envían al IdP, aumentando notablemente la seguridad.

En [14], se identifican las formas de obtención de *ID Tokens* y *Access Tokens* así como sus formatos, con los atributos que son obligatorios y los que son opcionales. En concreto, se definen tres tipos de flujos distintos, todos ellos basados en HTTP (*Hypertext Transfer Protocol*): autenticación utilizando el flujo *Authorization Code*, autenticación utilizando *Implicit Flow* y autenticación utilizando *Hybrid Flow*.

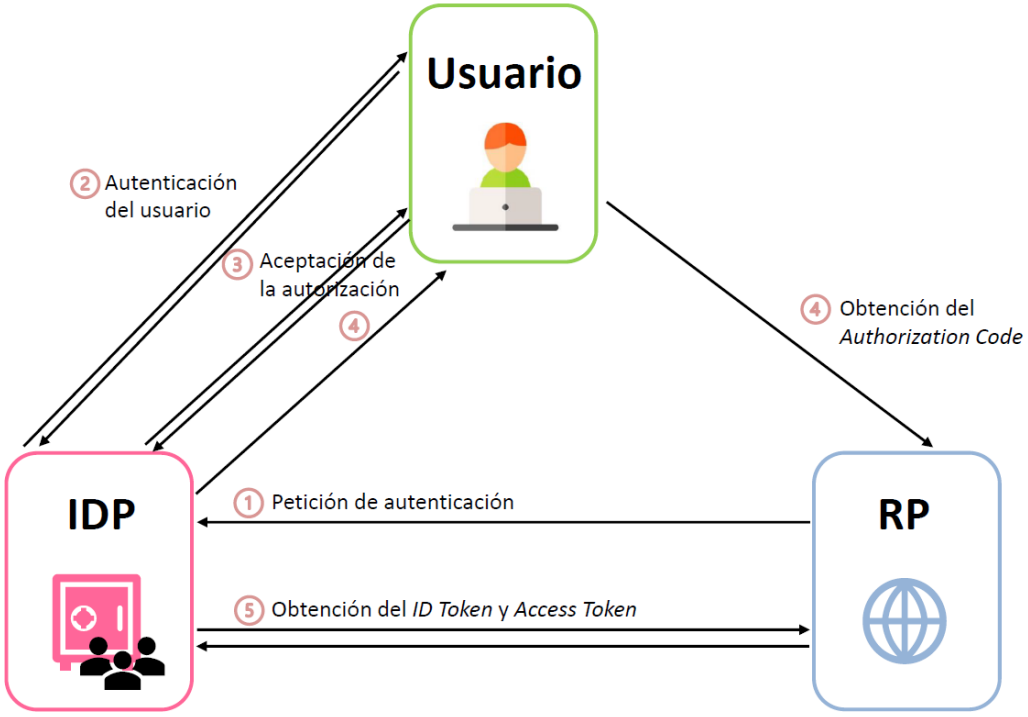

Figura 2.3: Flujo de OpenID Connect utilizando el flujo *Authorization Code*

La autenticación utilizando el flujo *Authorization Code* (figura 2.3) realiza los siguientes pasos:

1. La RP conforma y envía una petición de autenticación al servidor de autorizaciones del IdP.
2. El usuario final se autentica en el IdP.
3. El servidor de autorizaciones solicita al usuario la aceptación de las condiciones de la autorización.
4. El servidor de autorizaciones envía a la RP, a través del usuario, un código de autorización (*Authorization Code*).
5. La RP solicita, con el código de autorización, el *ID Token* y el *Access Token* y lo valida.

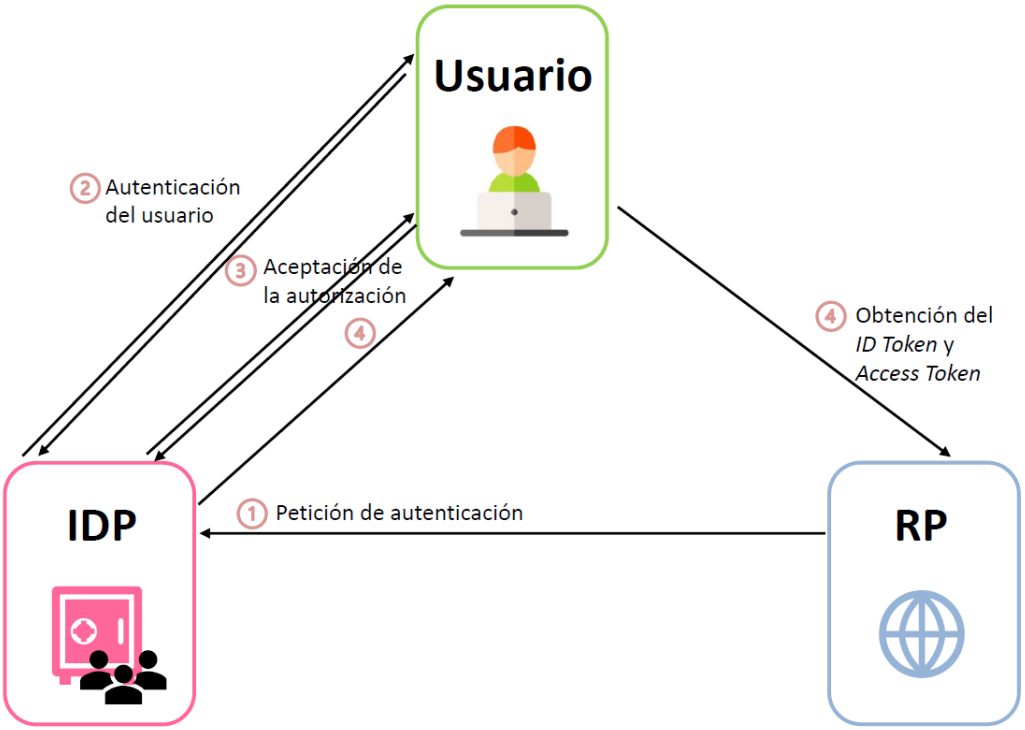

Figura 2.4: Flujo de OpenID Connect utilizando *Implicit Flow*

La autenticación utilizando *Implicit Flow* (figura 2.4) sigue los siguientes pasos:

1. La RP conforma y envía una petición de autenticación al servidor de autorizaciones del IdP.
2. El usuario final se autentica en el IdP.
3. El servidor de autorizaciones solicita al usuario la aceptación de las condiciones de la autorización.
4. El servidor de autorizaciones envía a la RP, a través del usuario, directamente el *ID Token* y/o el *Access Token*.

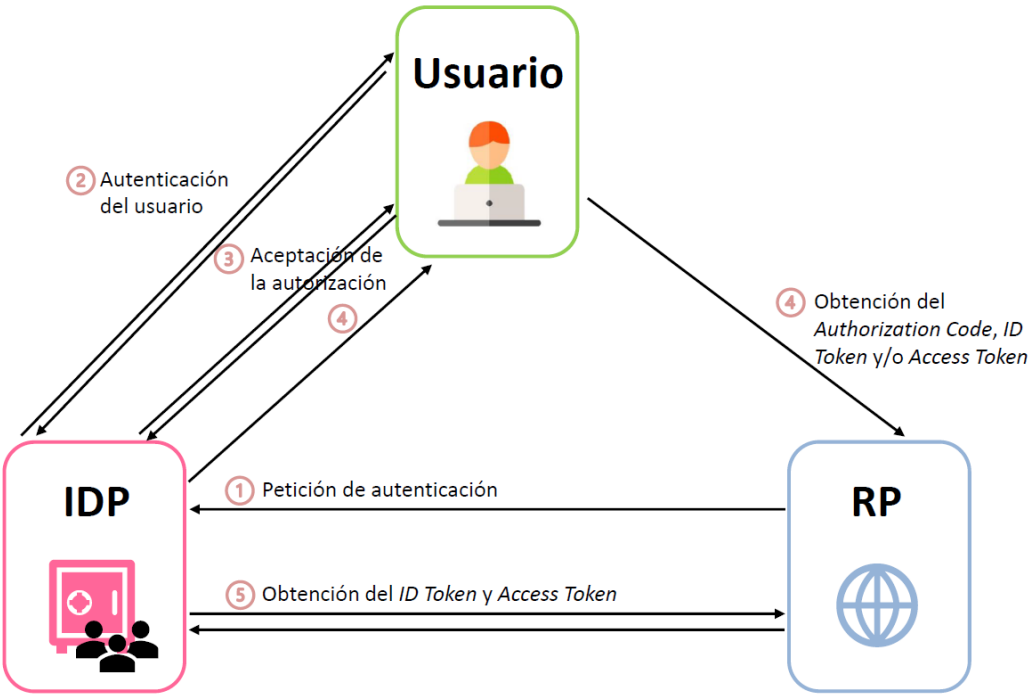

Figura 2.5: Flujo de OpenID Connect utilizando *Hybrid Flow*

Por último, la autenticación utilizando *Hybrid Flow* (figura 2.5) sigue los siguientes pasos:

1. La RP conforma y envía una petición de autenticación al servidor de autorizaciones del IdP.
2. El usuario final se autentica en el IdP.
3. El servidor de autorizaciones solicita al usuario la aceptación de las condiciones de la autorización.
4. El servidor de autorizaciones envía a la RP, a través del usuario, un código de autorización, *ID Token* y/o *Access Token*.
5. La RP solicita con el código de autorización el *ID Token* y el *Access Token* y lo valida.

Además de los distintos métodos de autenticación definidos, existe un mecanismo para forzar el cierre de sesión de usuarios en una RP con una comunicación desde el IdP a la RP. Es decir, una comunicación servidor a servidor sin necesidad de intervención del usuario en el proceso. Este flujo de comunicación es el *Back-channel logout* [15]. Como requisito inicial para la utilización

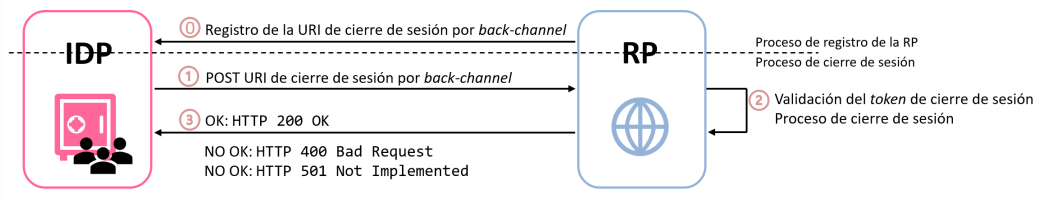

Figura 2.6: Flujo de cierre de sesión por *back-channel*

de este flujo, la RP tiene que estar registrada en el IdP, incluyendo una URI (*Uniform Resource Identifier*) para este tipo de cierre de sesión (paso 0 de la figura 2.6). Para proceder al cierre de la sesión, el IdP realiza una petición HTTP tipo POST a la URI de cierre de sesión de *back-channel* con un *token* JWT, *JSON* (*JavaScript Object Notation*) *Web Token*, de cierre de sesión (paso 1 de la figura 2.6). Tanto la llamada HTTP, como el *token* JWT, tienen que seguir el formato descrito en la figura 2.7, donde se incluirá el evento de cierre de sesión (*"events": {"http://schemas.openid.net/event/backchannel-logout":* {}}) y una referencia al usuario afectado. El usuario se identifica directamente mediante el campo sub, utilizando la sesión del usuario (campo sid) o usando ambos. Tras la recepción de esta petición, la RP tiene que validar el *token* recibido y, si es correcto, proceder con el cierre de sesión del usuario (paso 2 de la figura 2.6) siguiendo las recomendaciones de OpenID Connect [14]. Cada RP contestará al IdP (paso 3 de la figura 2.6) con un `HTTP 200 OK` si la petición se ha procesado correctamente, con un `HTTP 400 Bad request` si la petición no es correcta, o con un `HTTP 501 Not Implemented` si el proceso de cierre de sesión falla.

El problema de la privacidad

Con este tipo de modelos federados y, en concreto, con la especificación OpenID Connect, los proveedores de redes sociales, diferentes compañías tecnológicas y operadores de telefonía han encontrado un mecanismo que provea un nuevo servicio a sus usuarios: la autenticación unificada (*Single Sign-On*) para recursos, aplicaciones y servicios en Internet. Como ejemplo, Facebook, Google o Apple permiten a sus usuarios autenticarse en una entidad tercera, sin necesidad de crear una nueva cuenta específica para esa aplicación o servicio. Por tanto, se simplifican los procesos de registro e inicios de sesión y se mejora la experiencia del usuario. Esto es lo que se ha denominado *Social Login*. Para las RP, este tipo de inicio de sesión tiene beneficios significativos, ya que delegan la gestión y mantenimiento de los usuarios en otra entidad tercera, y evitan molestar a los usuarios con la creación de nuevas cuentas para evitar perder usuarios en los procesos de registro o procesos de reseteo,

Petición de cierre de sesión por *back-channel*

```
POST /logout HTTP/1.1
  Host: rp1.com
  Content-Type: application/x-www-form-urlencoded
  Cache-Control: no-cache
  no-store Pragma: no-cache
```

***Token* JWT para el cierre de sesión**

```
{
  "iss": "https://idp.com",
  "sub": "248289761001",
  "aud": “pqR9hmn4rEw",
  "iat": 1611229023,
  "jti": "RftV",
"events": { "http://schemas.openid.net/event/backchannel-logout":{} }
}
```

Figura 2.7: Ejemplo de petición HTTP de cierre de sesión (parte superior) y de *token* JWT de cierre de sesión (parte inferior)

lo que incrementa el número de visitas.

Sin embargo, los proveedores de identidad ofrecen normalmente los servicios de forma gratuita porque los datos de los usuarios son un bien muy preciado [16]. Los datos recogidos permiten a los proveedores mejorar el contenido, servicios (a través de la personalización), su capacidad de mostrar anuncios personalizados, etc. En [17] se proponen experimentos para explorar la concienciación de los usuarios en privacidad cuando utilizan métodos de inicio de sesión desarrollando tutoriales y campañas de concienciación para informar de las ventajas e inconvenientes. Los resultados de los experimentos sugieren que la concienciación en privacidad de los usuarios aumenta cuando se ofrecen los mecanismos en los proveedores de identidad.

En trabajos como [18] y [19], se han identificado una serie de amenazas a la privacidad relacionadas con los modelos federados:

[Amenaza A1] **Fuga de datos personales**: El escenario de la fuga de datos personales tiene lugar cuando existe una brecha de seguridad en algún punto del modelo federado, de tal forma que se revelen, de forma no intencionada, datos personales de los usuarios que utilizan este modelo. La filtración intencionada no se tiene en cuenta como riesgo, ya que se considera a IdP

y RP como confiables (si no, el propio concepto de federación de identidades no tendría sentido, ya que se basa en la relación de confianza entre los agentes que la conforman).

[Amenaza A2] **Falta de control sobre los datos personales en el proveedor de identidad**: La falta de control sobre los datos personales en el proveedor de identidad se manifiesta cuando el usuario facilita sus datos personales al IdP. Una vez que éstos están en el IdP, el usuario no puede tomar decisiones acerca de cómo se almacenan, se limita su uso, o se borran sus datos personales.

[Amenaza A3] **Falta de transparencia al compartir los datos personales**: El riesgo de la falta de transparencia al compartir los datos personales se manifiesta cuando el IdP y RP intercambian datos personales de un usuario sin informarle. Tampoco existen formas de evitar el uso de los datos personales en la RP una vez que se le entregan. Esta falta de transparencia se puede deber a diferentes causas, como el desarrollo de modelos de negocio opacos en IdP o RP o la asimetría en la compartición de información entre los agentes de la federación de identidades (en la que el usuario final suele ser el menos informado).

[Amenaza A4] **Perfilado de los usuarios**: El riesgo de perfilado de usuarios se basa en que las entidades del modelo federado podrían ser capaces de rastrear las actividades de un usuario, así como conseguir información muy valiosa como sus gustos o hábitos. Por ejemplo, el IdP puede conocer los hábitos de consumo de un usuario analizando las RP a las que accede a lo largo del tiempo. Además, si se enriquece la información con otras fuentes, como redes sociales, se puede obtener un perfilado del usuario mucho más completo.

[Amenaza A5] **Localización de los usuarios**: El riego de la localización de los usuarios se manifiesta en la capacidad que tendrían IdP y RP de obtener la posición de un usuario, analizando las RP que utiliza u otros mecanismos de localización disponibles a través de los dispositivos que utiliza el usuario como el GPS (*Global Positioning System*), dirección IP (*Internet Protocol*), repetidores de telefonía, etc.

Estas amenazas se basan en considerar al IdP en el modelo federado como honesta pero curiosa (*Honest but curious*), es decir, el IdP cumple el protocolo del modelo federado pero es posible que analice los datos para otros propósitos más allá de la pura gestión de la identidad.

2.3. Los puntos de vista de la privacidad

La privacidad no es un término trivial de definir. El ámbito de aplicación, la subjetividad y momento histórico influyen en gran medida en la definición. Para ello, en [20] se propone abarcar el término bajo tres perspectivas distintas: la histórica, la legal y la de uso. Además, se analizan opiniones contrarias a la privacidad en las que se argumenta que, si no se protegiera tanto, se mejoraría en seguridad y a nivel sanitario. La conclusión de esta referencia, así como de otras similares de la literatura, es que hay que conseguir un equilibrio entre la privacidad de los usuarios y los beneficios que aporta la ausencia de ésta en ciertos contextos.

En [21] se trata el concepto de la privacidad desde el diseño (*Privacy-by-Design*) y de los patrones de diseño de privacidad. También se identifican cinco categorías de privacidad que se corresponden a los siguientes:

- Privacidad en el comportamiento.
- Privacidad territorial.
- Privacidad de las personas.
- Privacidad de las comunicaciones.
- Privacidad de los datos personales.

Todos los tipos de privacidad que se identifican se basan en el concepto de la revelación de *Personally Identifiable Information* (PII). En el NIST (*National Institute of Standards and Technology*) [22] se identifica como PII cualquier dato que permita identificar a una persona. Incluye algunos ejemplos como nombre, datos biométricos, número de pasaporte o de la seguridad social. Se contemplan los tipos de datos que se componen de identificadores, cuasi-identificadores, atributos confidenciales y no confidenciales, definidos en [23].

- Identificadores: son atributos que permiten identificar a un individuo directamente. Por ejemplo, el número del documento nacional de identidad.

- Cuasi-identificadores: son atributos que por sí mismos no identifican a un individuo, pero en combinación con otros cuasi-identificadores permiten reidentificarlo. Algunos ejemplos son la fecha de nacimiento, el género o la profesión.

- Atributos confidenciales: son atributos que contienen información sensible del individuo. Por ejemplo, la raza o sus datos médicos.

- Atributos no confidenciales: son atributos que no contienen información sensible y, por tanto, su exposición no acarrea un daño para el individuo. Por ejemplo, su color favorito.

2.4. Privacidad en los datos personales

Realizadas estas consideraciones iniciales acerca de un área tan extensa como la privacidad, hay que tener en cuenta que el foco de la investigación realizada en esta tesis doctoral está en la mejora de la privacidad de los datos de los usuarios que confían en el modelo federado para la gestión de su identidad.

Estos usuarios deberían poder conseguir el equilibrio ya mencionado entre su seguridad, comodidad y la protección de sus datos personales. Al mismo tiempo, las especificaciones en las que se basan los modelos federados deberían incluir aspectos de privacidad desde el diseño para ofrecer garantías suficientes, sobre todo en aquellos casos en los que se maneja PII.

El problema se puede abordar desde distintos puntos de vista, teniendo en cuenta que durante todo el proceso los datos están expuestos en distintos elementos de la cadena y a distintos niveles. Se han identificado en el estado del arte propuestas en las categorías descritas en la figura 2.8.

2.4.1. Privacidad en la publicación de datos

Los IdP almacenan una gran cantidad de datos personales de todos los usuarios a los que ofrecen el servicio de autenticación. Estos datos son susceptibles de ser compartidos con terceros o tratados con tecnologías como *Big data* o para realizar estadísticas. Por tanto, la protección de la privacidad de estos datos en este punto de la cadena es fundamental. En [23] se definen dos tipos de aproximaciones, la sintáctica y la semántica.

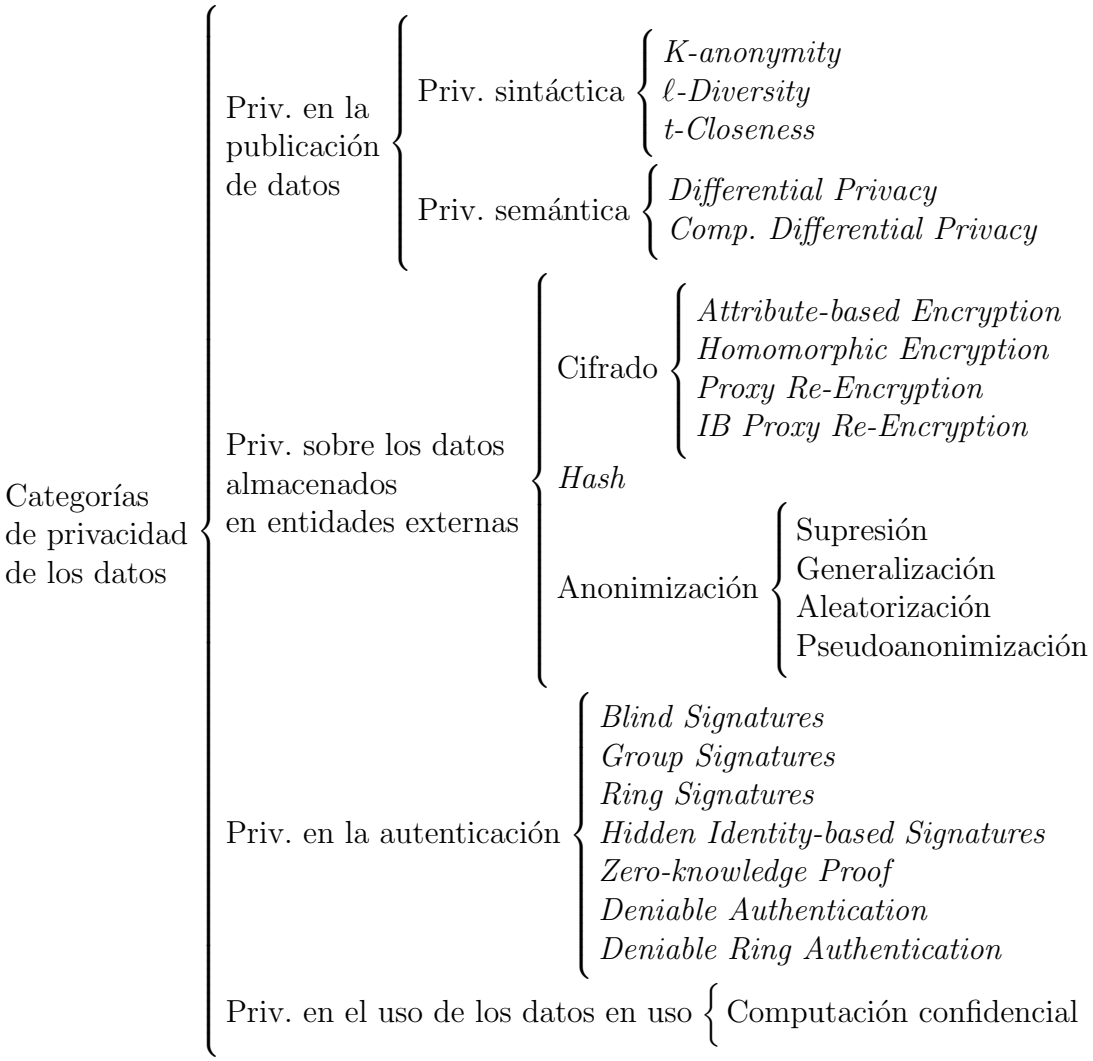

Figura 2.8: Categorías de privacidad de los datos

Privacidad sintáctica

La privacidad sintáctica se ocupa de proteger aquella información publicada, de tal forma que, un observador de dicha información publicada sea incapaz de identificar a un individuo concreto. A su vez, ésta se divide en proteger aquellos datos que identifican a una persona directa o indirectamente, es decir, la PII, y en la protección de atributos publicados, entendiendo atributo como aquella información que no identifica a un usuario, como su color favorito o los artículos comprados en una tienda.

Para la protección de los datos se utiliza *K-Anonymity*. Esta técnica consiste, en que cada una de las entradas de datos publicadas se debe poder asociar a K personas distintas dentro de la publicación. Para ello, se generaliza o suprime PII. Por ejemplo, sustituyendo el número de teléfono por $*$ exceptuando los dos últimos números.

Existen otras técnicas complementarias a *K-Anonymity* como *ℓ-Diversity* y *t-Closeness*, que permiten mitigar algunos tipos de ataques. *ℓ-Diversity* se utiliza para evitar ataques de homogeneidad (*Homogeneity attacks*). Por ejemplo, las K personas distintas han comprado el mismo artículo. También se puede utilizar para evitar revelar información en aquellos casos en los que el atacante tiene información externa (*External knowledge attacks*). Para ello, en los datos publicados se incluye que tiene que haber una diversidad ℓ en los atributos a proteger, en combinación con la protección de la identidad con *K-Anonymity*.

Por su parte, en [24] se define *t-Closeness* para evitar ataques de asimetría (*Skewness attacks*). Por ejemplo, el conjunto de K personas ha comprado libros exceptuando únicamente una persona que ha comprado alcohol. Se puede estimar con una elevada probabilidad, que una persona perteneciente a ese grupo ha comprado un libro. Además, evita ataques de similitud (*Similarity attacks*). Por ejemplo, el conjunto de K personas ha comprado «La Odisea», «El Quijote» y «Los Miserables». Esto nos indica que ese conjunto de K personas ha comprado libros. Para ello, en *t-Closeness* se establece que tiene que haber una distancia t entre las distribuciones de los atributos.

Privacidad semántica

La privacidad semántica protege a todos los individuos como conjunto, tanto aquellos incluidos en la publicación de datos, como aquellos que no lo están. Se define en [25] la técnica *Differential Privacy*. En este modelo se tiene en cuenta que la capacidad de conocimiento de datos de fuentes externas y de la capacidad de cómputo no están limitadas (*computationally-unbounded adversaries*), de modo que se introduce ruido en los datos para

que no sea posible identificar a una persona en concreto. En [26], se categoriza en escenarios interactivos y no interactivos. En los escenarios interactivos, se protege cada una de las consultas que se hacen a los datos publicados. En los escenarios no interactivos, se protegen los datos antes de su publicación y no las consultas a los mismos. Estos escenarios tienen limitaciones en la cantidad de consultas que se realizan a los datos publicados, ya que se pueden hacer correlaciones entre consultas para eliminar el ruido. Existen estudios, como [27] y [28], donde se intenta alargar la vida útil de los datos publicados aumentando el número de consultas que se pueden realizar sobre éstos.

Con *Differential Privacy* los recursos y conocimientos de un adversario son ilimitados. En el mundo real esto no se cumple, y en [29] se detalla la *Computational Differential Privacy* donde se considera que tienen un acceso a recursos y conocimientos limitado. Es una técnica más relajada en lo que a privacidad se refiere.

2.4.2. Privacidad sobre los datos almacenados en entidades externas

Además de los datos que el IdP tiene almacenados en su propio sistema, éste comparte con las RP los datos necesarios para que puedan dar el servicio. Por tanto, es importante establecer unos límites en la propagación de estos datos a terceros de tal forma que se garantice la privacidad de los usuarios. Con este tipo de privacidad se abarca todos aquellos datos que se almacenan en un tercero. Estos datos se pueden proteger de varias formas distintas con técnicas de cifrado, *hashing* y anonimización.

El cifrado de los datos se puede realizar al almacenar los datos en el servidor donde se alojan. También se puede realizar antes del envío al servidor, en aquellos casos en los que el usuario sea el portador del dato. En el cifrado en reposo se protege la privacidad con respecto a quien intente acceder, una vez se escriben los datos en disco. Sin embargo, cuando el servicio quiere disponer de algún dato, tiene la capacidad de descifrarlo y utilizarlo. En el caso del cifrado previo al envío, el servicio desconoce el contenido de lo que está almacenando y, además, es incapaz de utilizarlo sin la autorización del usuario que ha cifrado dicho dato. En [30] el NIST recomienda criptosistemas específicos para esto.

Attribute-based Encryption (ABE) permite el descifrado de datos, siempre y cuando se satisfagan las condiciones de descifrado, basadas en atributos. Existen dos tipos de cifrado basado en atributos: *Ciphertext Policy Attribute Based Encryption* (CP-ABE) [31] y *Key Policy Attribute Based Encryption* (KP-ABE). En CP-ABE, toda persona posee una serie de atributos que con-

forma su identidad en el sistema. Los datos cifrados tienen asociado un árbol, formado con puertas lógicas y valores de atributos, que representa que aquellas personas que cumplan las condiciones de dicho árbol según sus atributos, podrán descifrar dicho dato. En KP-ABE, los usuarios tienen una clave, basada también en un árbol con puertas lógicas y atributos, que forma su identidad en el sistema. Los usuarios únicamente podrán descifrar de forma satisfactoria aquellos datos que les permita su clave.

El cifrado homomórfico, *Homomorphic encryption*, es un tipo de cifrado, que puede ser tanto simétrico como asimétrico, con la posibilidad de operar sobre los datos cifrados. Es decir, se obtiene el mismo resultado realizando el descifrado y después realizando las operaciones, que realizando las operaciones y descifrando. Esta posibilidad de operar sobre los datos cifrados, permite que se pueda delegar a terceras partes no confiables la computación, sin riesgo de que conozca los datos en claro. Existen tres tipos de cifrado homomórfico: *Partially homomorphic encryption* (PHE), *Somewhat homomorphic encryption* (SHE) y *Fully Homomorphic Encryption* (FHE). *Partially homomorphic encryption* (PHE) permite realizar operaciones de suma o multiplicación sin ningún tipo de límite, pero no ambas. *Somewhat homomorphic encryption* (SHE) permite la realización de multiplicaciones y sumas de forma limitada. *Fully Homomorphic Encryption* (FHE) permite realizar ambas operaciones sin ningún tipo de límite.

Una variante sobre las técnicas de cifrado para el almacenamiento de datos es el *Proxy Re-Encryption*. Con esta herramienta se permite cambiar el cifrado que tiene un mensaje por otro sin que éste llegue a estar en claro. En primer lugar, un usuario cifra con su clave pública un mensaje que almacena el *Proxy Re-Encryption*. Además, genera una clave de recifrado que también le envía con la clave pública y privada del usuario y la del destino. Cuando el destino reclama el mensaje, el *Proxy Re-Encryption* haciendo uso de la clave de cifrado, es capaz de obtener un mensaje cifrado con la clave pública del destino. Este nuevo mensaje cifrado se envía al destino que es capaz de descifrar con su clave privada. Existen muchas variante de estos *Proxy Re-Encryption* según las características que ofrecen. Existen esquemas unidireccionales y bidireccionales. En los unidirecionales, únicamente el usuario puede generar las claves de recifrado, por lo que únicamente el usuario puede enviar mensajes al destino. En los bidireccionales, ambos extremos de la comunicación pueden generar claves de recifrado y, por tanto, ambos pueden enviar mensajes. Otros esquemas se centran en el número de recifrados que se puede realizar. Algunos esquemas solo permiten un recifrado del mensaje, otros permiten un número limitado de recifrados del mensaje y otros no tienen límites en el número de recifrados que se pueden hacer. También existen esquemas interactivos y no interactivos. Los esquemas interactivos

necesitan la clave privada del destino para poder generar la clave de recifrado, mientras que para los esquemas no interactivos no es necesaria. En [32], se hace un análisis en profundidad de los *Proxy Re-Encryption*, así como de sus características y tipologías.

En [33] se ha definido *Identity-based Proxy Re-encryption*, es decir, *Proxy Re-Encryption* basado en la criptografía *Identity Based Encryption*. *Identity Based Encryption* (IBE) [34] es un método de cifrado asimétrico para el envío de mensajes, donde la distribución de las claves es mucho más sencilla que en las infraestructuras de clave pública (PKI - *Public Key Infraestructure*). Ambos usuarios, emisor y receptor, comparten parámetros sobre su identidad, como su correo electrónico, con un *Private Key Generator* (PKG). El PKG es un sistema de confianza que posee una clave maestra pública y una clave maestra privada. En primer lugar, el emisor obtiene la clave maestra pública del PKG, y cifra el mensaje con la clave pública y la dirección de correo del destino. Una vez el destino obtiene el mensaje cifrado, se autentica en el PKG para validar su identidad y obtiene su clave privada. Con ella, es capaz de descifrar el mensaje. El procedimiento para la firma de mensajes sería análogo, pero utilizando la clave privada del emisor, que se obtiene previamente autenticándose en el PKG. Para *Identity-based Proxy Re-encryption*, un usuario cifra con un atributo de su propia identidad, que corresponde con la clave pública, los datos. Después, envía estos datos cifrados junto con la clave de recifrado del destinatario. Cuando el destinatario solicita el envío de uno o varios datos del usuario, el *Proxy Re-encryption* los recifra con la clave pública del destino. El destino solicita al PKG su clave privada autenticándose. Una vez la obtiene puede descifrar los datos.

El *hashing* de datos permite que no se pueda obtener una salida de tamaño fijo (*hash*) de entradas de datos variables. Además, un *hash* tiene la particularidad de que no es posible obtener los datos originales, es decir, no es reversible. Adicionalmente a los datos de entrada, se puede añadir sal criptográfica para evitar aquellos casos en los que los datos de entrada puedan ser iguales. De esta forma, se generan valores *hash* distintos para datos iguales.

En la anonimización de los datos, la PII se modifica para que no se pueda identificar a una persona. De esta forma, la identidad de una persona se mantiene a salvo, pero los atributos asociados a esa persona se mantienen legibles. En [35] se definen los métodos para anonimizar los datos: supresión, generalización, aleatorización y pseudonanonimización. La supresión elimina la PII asociada a los atributos, mientras que la generalización suprime cierta parte de cada dato identificativo. La aleatorización añade ruido o permuta la PII y la pseudoanonimización se sustituye por un identificador, que dependiendo del identificador elegido, puede ser reversible o no.

2.4.3. Privacidad en la autenticación

Para mejorar la privacidad de los usuarios que interactúan con el IdP y la RP, existe la posibilidad de que estas entidades desconozcan qué usuarios se han autenticado o están utilizando un servicio. De esta forma, se evita que el perfilado. Una vez se ha autenticado el usuario, los proveedores de servicios no son capaces de saber si están sirviendo al mismo usuario o a uno distinto.

Se pueden utilizar técnicas de *Blind Signatures*, *Group Signatures*, *Ring Signatures*, *Hidden Identity-based Signatures*, *Zero-knowledge Proof*, *Deniable Authentication* y *Deniable Ring Authentication*.

Blind Signatures permite que un origen obtenga un mensaje firmado por el destino, sin que éste conozca el contenido del mensaje. El funcionamiento es el siguiente:

1. El origen envía al destino un mensaje m, ofuscado con un número aleatorio r cifrado con su clave pública.

$$x = r^e m \bmod n$$

2. El destino recibe x, que lo firma con su clave privada, obteniendo t.

$$t = x^d \bmod n = (r^e m)^d \bmod n = r m^d \bmod n$$

3. El origen, que conoce el valor r, obtiene el mensaje original firmado por el destinatario.

$$s = r^{-1} t \bmod n = m^d \bmod n$$

Group Signatures [36] define un grupo de miembros que pueden realizar firmas. De esta forma, una persona que verifica dicha firma sabe que es válida y que pertenece a ese grupo, pero no puede saber qué miembro del grupo ha realizado la firma. Únicamente el coordinador del grupo (*Group Manager*) puede saber qué miembro ha realizado la firma. A este termino se le conoce como abrir una firma (*signature open*). Desde su definición en 1991, se ha mejorado el proceso y añadido nuevas funcionalidades como [37], [38], [39], [40].

Ring Signatures [41] tiene un funcionamiento semejante al de *Group Signatures*, con ciertas salvedades. En este tipo de firmas anónimas no existe el rol de coordinador del grupo, por lo que no existe la posibilidad de realizar la apertura de firmas. Además, no es necesario realizar una etapa previa de

configuración más allá de que el usuario que va a realizar la firma conozca las claves públicas de otros usuarios. En [42] se define que, a diferencia de *Group Signatures*, *Ring Signatures* no necesita la colaboración de otros usuarios que cooperan para su elaboración.

Hidden Identity-based Signatures [43] tiene un funcionamiento similar a *Group Signatures*, con algunas salvedades. Primero, se sustituye la función del *Group Manager* por la del *Identity Manager*, que se encarga de coordinar las identidades de los usuarios. Al no existir un coordinador por grupo, hay una autoridad externa e independiente al resto, llamada *Opening Authority*, que puede abrir las firmas que se hacen. También, se elimina la necesidad de tener una lista con los usuarios del grupo. Con estas cualidades, se resuelve el «*Anonymity Catch-22 issue*» identificado en [43], ya que, el *Opening Authority* no necesita consultar ningún otra entidad para poder abrir la firma. Se han mejorado las cualidades de esta técnica como en [44] y [45].

Zero-knowledge Proof es otra técnica que se suele utilizar de forma reiterada para mantener la privacidad durante el proceso de autenticación. Se basa en que el origen puede demostrar a un destino el conocimiento de un mensaje x sin que el destino conozca dicho mensaje. Para ello, se sigue un protocolo de tres rondas donde $p = 2q + 1$, siendo p y q números primos y g el generador:

1. El origen envía al destino un número aleatorio t.

$$y = g^x \bmod p$$

$$r = g^t \bmod p$$

2. El destino le devuelve un número aleatorio c (el reto).

3. El origen calcula s y se lo envía al destino.

$$s = t - cx \bmod q$$

4. El destino puede comprobar que el origen conoce x si se cumple la siguiente condición.

$$g^s y^c = r \bmod p$$

Deniable Authentication [46] se basa en la característica de que un observador externo al proceso de autenticación no puede constatar de manera unívoca el autenticado ni el autenticador una vez ha ocurrido el proceso. Sin embargo, el autenticador es capaz de realizar este proceso y autenticar al usuario. El proceso es el siguiente:

1. El sistema autenticador envía una clave k cifrada con la clave pública del usuario.

2. El usuario, que es el único que puede descifrar el mensaje, obtiene la clave k y genera un mensaje m un *tag*, mediante una MAC (*Message Authentication Code*) con la clave k.

De esta forma, el autenticador sabe que únicamente ese usuario ha podido descifrar la clave k y si es capaz de mandar el mensaje significa que es el usuario. Sin embargo, este autenticador no puede convencer a un tercero de que dicho mensaje lo ha generado el usuario, ya que como conoce también la clave k, él mismo podría haberlo fabricado.

Deniable Ring Authentication [47] añade tanto las cualidades expuestas para *Deniable Authentication* como para *Ring Signature*. Esto es que, el sistema que autentica sabe con certeza que un mensaje proviene de un grupo de firma, pero no puede demostrar a un tercero que esto es así. Por lo tanto, provee de un grado mayor de privacidad con respecto a *Deniable Authentication* y *Ring Signature*.

2.4.4. Privacidad de los datos en uso

Durante la operativa normal de los IdP y RP se necesita el procesamiento de ciertos datos de los usuarios, y para esto se tienen que cargar los datos en memoria no persistente. Los datos cargados en este tipo de memoria se tienen que descifrar para poder hacer uso de ellos (a excepción de los casos en los que se utiliza cifrado homomórfico) y, por tanto, están expuestos al robo o manipulación de los mismos. Para mitigar este riego existe la computación confidencial (*Confidential computing*) [48].

Este tipo de técnica permite separar en un entorno hardware de ejecución confiable (TEE, *Trusted Execution Environment*) aquellos datos que se consideren sensibles, pudiendo únicamente acceder a estos datos aquellos programas o rutinas autorizados. Los TEE tienen que cumplir tres principios básicos: la confidencialidad de los datos para que únicamente las entidades autorizadas puedan verlos. La integridad de los datos para que únicamente las entidades autorizadas puedan modificar los datos. Y la integridad del código para que únicamente las entidades autorizadas puedan modificar el código que ejecuta en el TEE. Además de estas propiedades básicas, puede tener otras complementarias como la confidencialidad del código, la autenticación de procesos que ejecutan el código, la capacidad de ejecutar únicamente código definido en tiempo de fabricación, ofrecer evidencias verificables por terceros y la capacidad de recuperarse de un estado comprometido.

2.5. Privacidad en los modelos federados

Una vez expuesto el funcionamiento de los modelos federados para la gestión de la identidad (y en concreto, OpenID Connect como base de los más extendidos), las amenazas que suponen para la privacidad y los aspectos esenciales acerca de la privacidad de datos personales, se puede pasar a analizar el estado del arte en la investigación relacionada con la mejora de la privacidad en estos modelos.

Para la mitigación de las amenazas identificadas, la mayor parte de trabajos aplican una combinación de aproximaciones y técnicas de las diferentes categorías identificadas en la sección 2.4.

Las soluciones propuestas hasta ahora en el estado del arte tienen algunos elementos o propiedades en común, que pueden ayudar a comprender mejor su naturaleza y alcance:

[Propiedad P1] **Escenario de aplicación**: Para la mitigación de las amenazas se especifica un escenario concreto de resolución. Es posible que la solución que se proponga busque la solución de una serie de amenazas en un entorno concreto donde el problema esté todavía sin resolver total o parcialmente. Algunos ejemplos de escenarios comunes son el Internet de las cosas (IoT - *Internet of Things*), dispositivos móviles o computación en nube.

[Propiedad P2] **Protocolo federado**: Las propuestas se pueden realizar para protocolos o especificaciones concretos dentro del modelo federado,o incluso crear un nuevo protocolo que se ajuste a las necesidades del escenario para la mitigación de las amenazas para la privacidad escogidas.

[Propiedad P3] **Inmutabilidad del protocolo federado**: Las propuestas pueden necesitar la modificación de un protocolo estándar para que la solución sea efectiva. Dentro de las modificaciones a la especificación estándar contempladas están tanto los intercambios de mensajes, como los propios mensajes. Algunos ejemplos son el intercambio de mensajes adicionales entre las entidades del modelo federado, el uso de nuevos mensajes o cabeceras no contempladas en el estándar originalmente. Las implementaciones del protocolo tendrían que ser modificadas para que funcionara la solución propuesta en este tipo de investigaciones.

[Propiedad P4] **Entidades complementarias**: El uso de entidades complementarias (nuevos agentes en las federaciones de identidades) es habitual para solventar amenazas que el protocolo no puede solucionar desde el propio diseño. El uso de estas nuevas entidades puede mantener el protocolo del modelo federado intacto o puede que sea necesario modificarlo para integrarlas.

En la tabla 2.1 se hace una recopilación de los trabajos de investigación relativos a la mitigación de la amenazas para la privacidad en modelos federados. En la parte izquierda de la tabla se indican las amenazas mitigadas (utilizando las listadas al final de la sección 2.2) y en la parte derecha se muestran las propiedades de la solución propuesta. En aquellos casos en los que se diseñe un nuevo protocolo federado, la propiedad P3 se ha marcado como no evaluable (con un guión), ya que, para este tipo de casos, dicha columna no aporta nada a la comparativa.

La primera amenaza se puede mitigar o incluso evitar completamente mejorando la seguridad de la especificaciones federadas, así como sus implementaciones en el proveedor de identidad. En este aspecto se centran los trabajos [49], [50], [51], [52], [53], [54] y [55].

En [49] se propone el uso de una nueva entidad, PrOfESSOS, que permite el análisis del protocolo OpenID Connect utilizado en el IdP y la RP. Esta herramienta tiene en cuenta los distintos tipos de flujo de autenticación (*Authorization Code*, *Implicit* y *Hybrid*) y una serie de comprobaciones automáticas para determinar si existe una fuga potencial de los datos personales en base a amenazas ya conocidas u otras nuevas como *Identity Provider Confusion* o *Malicious Endpoints*. El proceso de comprobación se divide en tres fases: en la primera, el usuario accede a PrOfESSOS y se generan de forma automática los elementos necesarios para las comprobaciones. En la segunda, el usuario introduce la RP que quiere examinar y PrOfESSOS intenta obtener la URL (*Uniform Resource Locator*) para iniciar el proceso de inicio de sesión (si no es capaz el usuario deberá proporcionarla manualmente). Y en la tercera, se realizan las comprobaciones para determinar si existe la amenaza de fuga de datos personales.

En [50] se propone un modelo que identifica una serie de amenazas relativas a OpenID Connect. Se tratan de solventar mediante la aplicación de una guía de uso seguro para este protocolo.

En [51] se ha desarrollado un plugin web que analiza las comunicaciones entre las entidades de la federación para detectar posibles fugas de datos personales. Se basa en revelaciones de los *token* de identidad (*ID Token*) o

Tabla 2.1: Comparativa de trabajos previos en al área de la mejora de privacidad del modelo federado

Trabajo	Amenaza mitigada					Propiedades			
	A1	A2	A3	A4	A5	P1	P2	P3	P4
[49]	X					Todos	OIDC	X	X
[50]	X	X				Todos	OIDC		
[51]	X					Todos	OIDC	X	X
[52]	X	X				Todos	Liberty Alliance		X
[53]	X	X				*Smartphone*	Liberty Alliance		X
[54]	X					Nube	OIDC	X	
[55]	X					Nube	OpenID 2.0		X
[56]		X	X			Todos	IdMRep	-	
[57]	X	X	X			Nube	OIDC	X	X
[58]		X	X			Nube	OIDC		X
[59]		X	X			Nube	OIDC		X
[19]	X	X	X			Todos	OIDC	X	X
[60]		X	X			Todos	Facebook Login		X
[61]			X			Nube	SAML 2.0	X	X
[62]			X			Todos	OIDC		X
[63]	X		X			Nube	OIDC		X
[64]	X	X	X			Nube	SAML 2.0		X
[65]	X	X	X			Nube	OIDC		X
[66]				X		Todos	OIDC		X
[67]				X		Todos	OIDC	X	X
[68]				X		Todos	SPRESSO	-	X
[18]				X		Todos	OAuth 2.0		
[69]				X		Todos	OIDC		X
[70]				X		*Smartphone*	Passphone	-	X
[71]					X	Todos	PPLBS	-	X
[72]					X	*Smartphone*	Grid-File Index	-	X
[73]	X			X		Todos	PRIMA	-	
[74]				X		Nube	NEXTLEAP	-	
[75]				X		IoT	OAuthing	-	
[76]			X	X		*Smartphone*	OIDC/SAML 2.0		X
[77]	X	X		X		*Smartphone*	Mobile eID	-	
[78]				X		Todos	Nuevo protocolo	-	
[79]			X	X		Todos	OIDC		
[80]	X	X	X			Todos	reclaimID	-	X
[81]	X	X	X			*Smartphone*	Crypto-Book	-	X
[82]	X					*Smartphone*	OFELIA	-	X

token de acceso (*Access Token*). Este plugin tiene la capacidad de avisar al usuario de la fuga de datos o incluso de abortar el flujo de autenticación si es necesario.

En [52] y [53], basados en la arquitectura *Liberty Alliance project* [83], el usuario puede controlar sus datos personales utilizando políticas personalizadas. La entidad *Privacy Controller* (PC) permite a los usuarios en [52] comprobar o modificar sus datos personales, así como controlar su procesamiento y con quién se comparten. Este mecanismo permite a los usuarios saber cómo se produce la revelación de estos datos. En [53], la solución propuesta es muy similar pero adaptada al entorno móvil utilizando un nuevo elemento, el *Mobile Information Service Broker*.

En [54], se propone una revelación de datos personales controlada utilizando *Brokers* y *Proxy Re-Encryption*. Los *Brokers* centralizan la gestión de IdP y RP, y lo *Proxy Re-Encryption* permiten cambiar la clave de cifrado sobre un elemento cifrado sin llegar nunca a descifrarlo. Estos nuevos componentes permiten a los usuarios cifrar los datos personales que necesita la RP utilizando una clave que únicamente conoce esa RP, evitando revelación de datos no intencionadas.

En [55] se busca una solución concreta al cifrado de los datos personales en el IdP. Para ello, se propone la utilización de *Proxy Re-Encryption*. El usuario envía sus datos personales cifrados con una clave asimétrica, cuya clave privada es custodiada por él. Cuando quiere enviar parte de sus datos a una RP, el *Proxy Re-Encryption* es capaz de cambiar el cifrado realizado por el usuario, por uno en el que solamente esa RP puede descifrar(en ningún momento los datos personales está sin cifrar). Para el intercambio de atributos se utiliza OpenID Attribute Exchange 1.0, actualmente obsoleto, complementándolo con nuevas etapas de cifrado y descifrado por parte del *Proxy Re-Encryption*.

Otros trabajos, como [56], [57], [58], [59], [19], [60], [61], [62], [63], [64] o [65] proponen soluciones para la segunda y tercera amenaza, intentando minimizar los datos personales necesarios en el proceso de registro en el proveedor de identidad y mejorando la transparencia del proceso de compartición de dichos datos personales, así como el control que se le proporciona al usuario.

En [56] se trata de resolver estas amenazas mediante el diseño de un nuevo protocolo federado basado en sistemas de reputación, IdMRep. Este protocolo se fundamenta en el uso de listas de confianza dinámica utilizándolas para obtener la información de reputación de IdPs y RPs.

En [57] se utiliza una capa intermedia donde se identifican métricas de riesgo y de reputación. Mediante modelos de diseminación de datos, el uso de

políticas y cifrado de datos personales se consigue mitigar estas amenazas.

En [58] se propone una extensión en los *scopes* de OpenID Connect que permiten añadir niveles de privacidad. Permite mitigar las amenazas dos y tres, y pseudoanonimizar o anonimizar los datos para mitigar la amenaza de fuga de datos personales.

En [59] se modifica el protocolo OpenID Connect para que haga uso de un *Privacy token*. Este *token*, al igual que en OIDC, es un JWT, pero está especialmente diseñado para definir las preferencias de privacidad del usuario en el uso de los servicios. También se utiliza un agente, *Security manager*, que intermedia en la comunicación con la RP en lo relativo a las preferencias de privacidad. Con el uso del *Privacy token* y este agente se mitigan las amenazas de falta de control y transparencia al compartir los datos personales de los usuarios.

En [19] se propone el uso de un árbitro de privacidad que permita controlar la propagación de los datos personales de los usuarios. También se añade la posibilidad del cifrado de los datos personales de los usuarios, utilizando este nuevo agente como intermediario.

En [60] se asiste a los usuarios en la toma de decisiones mediante el consentimiento informado. Para lograr este objetivo, se ha desarrollado un asistente para mitigar las amenazas relativas a la falta de control y transparencia de los datos personales, y un tutorial explicativo sobre las ventajas e inconvenientes del modelo federado.

En [61] se trata de mejorar la privacidad en el proceso de autenticación en SAML 2.0 para el contexto de la computación en la nube. Se define una nueva entidad, *Privacy Engine Module*, responsable de monitorizar el uso de los datos de los usuarios. También se propone una capa de confianza (*Trust Layer*) que utiliza un nuevo sistema para obtener información sobre la reputación y analizar los riesgos para la privacidad.

En [62] para que el IdP pueda validar que la información es correcta, utiliza un *Validation Service* (VS). Un VS puede ser un registro civil del estado, banco, etc. Se centra en la ayuda de toma de decisiones basándose en sistemas de reputación, políticas y peligros potenciales para mantener bajo control la propagación de los datos personales.

Por su parte, en [63] se centran en un modelo de IdP en nube, donde el usuario también tiene la capacidad de subir sus datos cifrados. Como medidas adicionales al control de la propagación de los datos personales, el usuario tiene la capacidad de registrar políticas de privacidad (*sticky policies*) y de propagación en dicha nube.

En [64] y [65] se pretende mejorar la privacidad usando también un *Proxy Re-Encryption* para SAML 2.0 y OpenID Connect. Se explota el concepto de privacidad y control de la identidad en el ámbito de IDaaS, *Identity Mana-*

gement as a Service. En este modelo, la nube no es totalmente de confianza, ya que, el proveedor podría acceder a los datos personales de los usuarios (*honest but curious*). Para ello, utilizan el modelo BlindIdM [64]. En BlindIdM, se ofrece el servicio de gestión de la identidad del usuario, delegando la autenticación a la organización con SAML 2.0. En primer lugar, la organización conforma los atributos de los usuarios en formato clave-valor, y cifra con su clave pública el valor de cada atributo. A continuación, envía a la nube los atributos indicando el usuario al que pertenece y las claves de recifrado para las RP que quiere utilizar. Cuando una RP necesita ciertos atributos del usuario, el IDaaS, haciendo uso de *Proxy Re-Encryption*, envía los atributos cifrados con la clave pública de la RP a través del usuario. En [65] se propone la adaptación de este modelo para el protocolo OpenID Connect.

Por otro lado, [66], [67], [68], [18], [69], [70] y [71] y [72] proponen soluciones para las amenazas cuarta y quinta, que se centran en evitar el perfilado y localización de los usuarios.

En [66] se utilizan *blind signatures* para probar la identidad en el IdP para una aplicación o servicio específico sin revelarla. Para ello, se utiliza un servicio de firmas ciegas (BSS, *Blind Signatures Service*) que genera *tokens* de acceso usando este algoritmo. Por tanto, se mitiga la amenaza de perfilado del usuarios ya que no puede conocer qué servicios utiliza un usuario concreto.

En [67] el IdP puede conocer la información sobre las RP que utiliza un usuario y el IdP puede conocer qué RP está accediendo a la información del usuario que tiene almacenada el IdP. Se proponen realizar modificaciones en el protocolo y el uso de un intermediario, como un plugin web, entre el usuario y el IdP que permita mitigar estas amenazas.

En [68] se propone un nuevo modelo de autenticación federado, SPRESSO (*Secure Privacy-REspecting Single Sign-On*), modificando la raíz de confianza del proveedor de identidad al navegador del usuario. El IdP no es capaz de diferenciar las RP a las que se conecta un usuario, por lo que se mitiga la amenaza de perfilado de los usuarios utilizando una nueva entidad intermedia, los *Forwarders*.

Por su parte, en [18] se propone el uso de pseudo-identidades, credenciales anónimas basadas en atributos que usan aMACs (*algebraic Message Authentication Codes*).

En [69] se sustituye el *client_id* enviado al IdP con un *hash* del identificador y dos *nonces* (uno de ellos para el uso de un servicio o aplicación concreto y otro para el *user-agent*). Se oculta, por tanto, la identidad de la RP a la que el usuario quiere acceder, mitigando la amenaza de perfilado de usuarios.

En [70] se diseña todo un nuevo protocolo con comunicaciones cifradas

usando TLS (*Transport Layer Security*), en el que el usuario tiene que tener disponible un teléfono móvil con una aplicación instalada de un tercero confiable (IdP) y un buzón de correo electrónico. Con estos elementos se establece el doble factor de autenticación. En este diseño el IdP desconoce el usuario que debe autenticar y evita que los proveedores de servicios puedan relacionarlo con una identidad única.

En [71] y [72], se tiene como objetivo mantener la privacidad de los usuarios dando la exactitud de la respuesta, donde el proveedor del servicio desconoce la identidad y localización del usuario. [71] se basa en el esquema de autenticación *Deniable Ring Authentication* y en [72] se diseñan nuevos modelos de federados para autenticación en *Grid-file index* y *R-tree index*.

En los trabajos [73], [74], [75], [76], [77], [78], [79], [80], [81] y [82] se propone un cambio completo del concepto de federación permitiendo que los proveedores de identidad emitan credenciales y *tokens* que el usuario almacena de forma local. Estos usuarios utilizan estas credenciales y *tokens* para el acceso a los servicios y aplicación sin la necesidad de interacción del IdP. Además, los usuarios mantienen el control de los datos personales.

En PRIMA [73] se propone un flujo de autenticación que no requiere de interacción entre las RP y los IdP, evitando el perfilado de los usuarios. Además se permite controlar la revelación de los datos de los usuarios.

En NEXTLEAP [74] se busca mitigar la amenaza de perfilado asociando la identidad de los usuarios a un conjunto de ellos, por lo que no se puede identificar las acciones de un usuario en particular dentro del conjunto.

En [75] se centran en la autenticación en el escenario del Internet de las cosas (IoT), donde la capacidad de estos dispositivos es muy limitada y se gobiernan a través de un sistema central en la nube. Se define un nuevo modelo de autenticación federada, OAuthing, que hace uso de protocolos de autenticación y autorización como SAML 2.0, OpenID Connect y OAuth 2.0 para que se identifiquen tanto usuarios (UIdP) como dispositivos (DIdP). La generación de la identidad de los dispositivos la realiza la empresa proveedora, que además registra un *ClientID*, una clave secreta y una URL única por dispositivo, que se imprime en el mismo, por ejemplo, en *QR Code* (*Quick Response Code*). Una vez el usuario dispone de él, escanea dicho código para acceder al DIdP, que permite la autenticación del usuario a través del UIdP. Tras la autenticación utilizando OpenID Connect o SAML 2.0, se autoriza el dispositivo y se genera un *token* de OAuth 2.0 vinculando la cuenta del usuario al dispositivo. El objetivo de este modelo es que las entidades únicamente tengan acceso parcial a los datos y acciones de los usuarios, mitigando las amenaza de fuga de datos personales.

En [76], se diseña un nuevo modelo con un tipo de IdP distinto, el *Porta-*

ble Personal Identity Provider o PPIdP. Debido a las capacidades de computación crecientes en los teléfonos móviles, el PPIdP se instala en el teléfono móvil donde el usuario almacena sus datos personales. Esto permite que el grado de exposición de los datos se reduzca y que el control del usuario aumente, ya que, puede operar sobre el propio proveedor de identidad. Además, ofrece capacidades de uso de protocolos estándares, como SAML 2.0 y OpenID Connect, lo que facilita en gran medida las integraciones con las RP. También permite controlar qué datos personales de distribuyen a las RP ya que se le presentan al usuario antes de compartirse y este puede seleccionar cuáles compartir.

En [77] se esbozan una serie de ideas para que la verificación de la identidad sea lo más sencilla posible, pero manteniendo la privacidad de la identidad y evitando el perfilado. Se propone un sistema que emite las identidades electrónicas a los usuarios en dispositivos de fácil lectura para las personas, como en tarjetas inteligentes o en el teléfono móvil. Cuando un usuario quiere comprobar algún dato de la identidad de otro, realizará consultas a sus datos a través del dispositivo, siendo el resultado de esa consulta verdadero o falso. Así, se limita la exposición de datos a terceros, salvo en aquellos casos en los que sea imprescindible. Por motivos operativos, las consultas se podrían hacer tanto de forma *online*, como *offline*, e incluso la persona que verifica no necesitaría un dispositivo conectado a corriente eléctrica o batería. Se establecen también requisitos de revocación de atributos y mecanismos para el control de la propagación de los datos personales. Además, se utiliza *K-Anonymity* con respecto al global de todos los usuarios para evitar el perfilado de las acciones que realiza un usuario específico.

En [78] se diseña un nuevo flujo de autenticación, en el cual, el propio usuario hace de intermediario en las comunicaciones entre la RP y el IdP. La RP genera un *token* que se forma realizando una función *hash* del *EndPoint* de la RP, *Nonce* y *TimeStamp*. Este *token* se envía de la RP al IdP, pasando por el usuario, que verifica que se haya generado de forma correcta. Tras la autenticación del usuario en el IdP, éste genera una respuesta firmada que retorna a la RP (también a través del usuario), y que la RP puede verificar a través de su firma.

En [79] se utiliza este mismo flujo con modificaciones en la validación de los *token* para mejorar la seguridad y se añade la posibilidad del envío de atributos desde el IdP a la RP para el control de la propagación de los datos personales de los usuarios.

En [80] se diseña una nueva arquitectura (reclaimID) que permite que los usuarios compartan atributos haciendo uso de *Attribute-based Encryption* (ABE). Los usuarios registran sus atributos en su *namespace*. Para esos atributos registrados, el usuario solicita al IdP que genere tickets, que permiten

a una RP consumir dichos atributos. La RP solicita atributos haciendo uso de uno o varios tickets que se le hayan emitido.

En [81] y [82] también se diseña una nueva arquitectura. En [81] se centran en mantener la identidad bajo el control de usuario y evitar el seguimiento de las RP. Para ello, diseñan una nueva arquitectura (*Crypto-Book*), basándose en sistemas de autenticación y autorización federada. En primer lugar, el usuario se autentica en el IdP y obtiene un *token* OAuth. Después, se conecta y envía el *token* a cada uno de los servidores de claves en nube utilizando la red Tor y el canal cifrado, donde al menos uno de los servidores de la nube es honesto. El usuario recibe cada una de las partes de la clave DSA (*Digital Signature Algorithm*) privada que le envían los servidores de claves y la recompone. La clave DSA pública la compone simplemente enviando el usuario con el que está registrado en el IdP a la nube de servidores de claves. Con su clave pública y la de otros usuarios, compone una *ring signature* que utiliza para mantener el anonimato en los servicios que consume. En [82] también se describe una nueva arquitectura, OFELIA (*Open Federated Environment for Leveraging of Identity and Authorization*), que se compone de una nube de *Attribute Authorities* (AA), donde se guarda los datos personales, y un *Identity Broker* (IdB), que se gestiona a través de una aplicación instalada en el teléfono móvil. Todas las comunicaciones se realizan cifradas. El IdB se encarga de recolectar la información de los AA y enviársela a la RP que el usuario quiere utilizar, haciendo de capa interfaz entre las RP y los AA. Por su parte, el teléfono móvil es una pieza fundamental dentro de este esquema, ya que, es donde reside la función del control del IdB.

2.6. Privacidad centrada en los usuarios

Como se ha analizado en la sección anterior, existen algunas propuestas en el estado del arte orientadas a mitigar algunas de las amenazas para la privacidad que supone el uso de OpenID Connect en escenarios genéricos (propiedad 1). El problema es que muchas de ellas obligarían a modificar la especificación estándar, lo que no es muy realista ya que todos los esquemas de *Social Login* y Mobile Connect se basan en esta especificación.

Para evitar tener que modificarla, algunas de las soluciones propuestas recurren a añadir un nuevo agente a las federaciones de identidades, un *proxy*, intermediario, árbitro o *broker* (la nomenclatura cambia según el trabajo de investigación). Estos agentes suelen tener funciones criptográficas, de control de propagación de datos y de desacoplamiento entre el IdP y la RP.

En la presente tesis doctoral se propondrá también el uso de un nuevo agente, pero con una función diferente, que no está incluida en las categorías

clásicas o tradicionales discutidas en la sección 2.4.

Muchos de los servicios y aplicaciones actuales permiten a los usuarios definir una configuración personalizada de seguridad y privacidad, a priori, esto permite un enfoque centrado en el usuario basado en la personalización, en el respeto a sus preferencias y necesidades. En cualquier caso, entender estas configuraciones y sus implicaciones no es fácil, incluso para los usuarios más expertos.

Por este motivo, diferentes autores han propuesto en el pasado sistemas de recomendación en forma de asistentes o asesores de privacidad para ayudar a los usuarios a encontrar la configuración de privacidad más apropiada para cada uno de ellos. En algunos trabajos de investigación previos se ha demostrado que los sistemas de recomendación tienen efectos significativos en el comportamiento de los usuarios y que los usuarios ven de utilidad una lista corta de recomendaciones [84], [85], [86].

En la tabla 2.2 se muestra una comparación entre sistemas de recomendación para privacidad, que permiten este enfoque de concesión de poder al usuario basado en proporcionarle información para que optimice sus decisiones. Todos los sistemas de recomendación se han categorizado en esta tabla considerando su dominio de aplicación. La columna "Personalización" indica la capacidad de los usuarios para seleccionar y modificar algunos criterios de evaluación del sistema de recomendación. La columna "Facilidad de adopción" se refiere al volumen y complejidad de las acciones que tiene que realizar el usuario para poder usar el sistema de recomendación en su proceso de toma de decisiones. La columna "Facilidad de uso" expresa si es fácil de utilizar y de entender la recomendación dada por el sistema de recomendación. Y por último, la columna "Facilidad de integración" indica si el sistema de recomendación es fácil de integrar en el dominio de aplicación para el que está diseñado, sin tener que realizar cambios sustanciales en las especificaciones, protocolos o implementaciones con los que debe integrarse.

El primer grupo de trabajos analizados se centra en proporcionar recomendaciones de privacidad en sitios web, servicios *online* y redes sociales. En [87] se propone el uso un sistema de recomendación que utiliza técnicas de *machine learning* para recomendar a los usuarios configuraciones de privacidad en la red social *Facebook*. Tiene como principal objetivo evitar que los usuarios dejen la configuración de privacidad por defecto.

En [88] se propone una métrica de privacidad para obtener mecanismos *soft-patternalism*, que aconseja a los jóvenes cómo mejorar su privacidad antes de realizar las publicaciones de datos personales en redes sociales.

En [89] se propone el uso del sistema de recomendación y monitorización *YourPrivacyProtection*. Muestra cómo obtener recomendaciones de privaci-

Tabla 2.2: Comparación de trabajos previos en el área de sistemas de recomendación centrados en la toma de decisiones acerca de privacidad

Trabajo	Dominio	Personalización	Fac. Adopción	Fac. Uso	Fac. Integ.
[87]	Redes sociales		X	X	X
[88]	Redes sociales		X		X
[89]	Redes sociales		X	X	X
[90]	Redes sociales	X	X	X	X
[91]	Redes sociales	X	X		
[92]	Portales Web		X	X	
[93]	Portales Web	X	X		
[94]	Portales Web		X		
[95]	Aplicaciones móvil	X			X
[96]	Aplicaciones móvil	X			X
[97]	Aplicaciones móvil		X		
[98]	Aplicaciones móvil	X			X
[99]	Big Data	X			X
[100]	IoT	X	X		

dad utilizando técnicas de *Collaborative filtering*. Esta técnica consiste en preparar recomendaciones de preferencias de privacidad a usuarios teniendo en cuenta perfiles de otros usuarios parecidos.

También se ha diseñado el asistente Tagvisor [90], que genera etiquetas cuando los usuarios comparten contenido en redes sociales. Tagvisor utiliza técnicas de generación, sustitución y borrado de etiquetas, y se centra en que no se revele la geolocalización de los usuarios.

Para la mejora de la privacidad cuando se utilizan imágenes en redes sociales se propone VISPR (*Visual Privacy Advisor*) [91]. VISPR es capaz de analizar estas imágenes para encontrar riesgos de privacidad utilizando etiquetas y da una recomendación según las preferencias que se hayan definido.

En [92] se diseña un *plugin* web que calcula una puntuación de privacidad en función de los *Third-party trackers* (TPT) que tenga dicha web.

En [93] se propone un asistente de privacidad que compara las políticas de privacidad P3P según su similitud. Trata de dar una recomendación utilizando el razonamiento basado en casos (CBR, *Case Based Reasoning*).

En [94] se propone una aproximación distinta, ya que el asistente de privacidad resalta a los usuarios los aspectos más relevantes de la política de privacidad de los sitios web.

El segundo grupo de trabajos analizados se centra en aplicaciones móviles. En [95] se propone un sistema de recomendación de colaboración abierta para el manejo de los permisos en el teléfono móvil y genera recomendaciones de privacidad sobre éstos.

En [96] se diseña un nuevo asistente personalizado de privacidad para el manejo de los permisos de las aplicaciones móviles.

Otro trabajo interesante es POLICHECK [97], que realiza un análisis

flow-to-policy en aplicaciones móviles para encontrar flujos de datos personales inconsistentes con las políticas de privacidad de las aplicaciones.

El proyecto *Personalised Privacy Assistant* [98] permite realizar configuraciones semi-automáticas en base a las preferencias de los usuarios, evitando interacciones con ellos en tiempo real. También permite generar alertas cuando se detectan problemas de privacidad y ofrece recomendaciones sobre cambios en configuraciones.

El último grupo de trabajos se centra en recomendaciones para IoT y *Big Data*. En [99] se propone el uso de un gestor de datos personales (PDM, *Personal Data Manager*). El PDM intercepta el tráfico entre el usuario y las aplicaciones de *fitness* para controlar la configuración de privacidad. También da recomendaciones de privacidad que evitan el perfilado de usuarios.

En [100] el usuario etiqueta parte de sus datos personales en base a sus preferencias y mediante técnicas de *machine learning* se completa la parte no etiquetada. La propagación de estos datos entre dispositivos IoT se realiza según las etiquetas configuradas, pudiendo tener un mayor control sobre la propagación.

Cabe destacar que no hay ningún trabajo centrado en los dominios de aplicación del control de accesos, gestión de identidades, modelos federados, *Social Login* o similar. No existe estado del arte en la aplicación de sistemas de recomendación en estos contextos.

Capítulo 3

Modelo de amenazas para la privacidad en el modelo federado

Como ya se ha explicado y analizado en los capítulos anteriores, cuando se emplea el modelo federado para la identificación y autenticación, el proceso es transparente y la experiencia de los usuarios es muy cómoda (su cometido se suele reducir a hacer *click* en un botón). Pero los usuarios proporcionan información sensible a los proveedores de identidad (datos personales estáticos, datos contextuales dinámicos), que podría implicar amenazas para su privacidad y todos los impactos potenciales asociados a ellas.

Este capítulo se centra en realizar un modelado exhaustivo de las amenazas para la privacidad que implica este tipo de federación de identidades. El objetivo es identificarlas y comprenderlas en toda su extensión, para pasar a proponer en los siguientes capítulos, estrategias de privacidad centradas en el usuario que puedan ayudar a evitarlas o mitigarlas.

3.1. Metodología LINDDUN

LINDDUN [101] es una metodología de modelado de amenazas para la privacidad. Permite analizar sistemas de forma estructurada en busca de amenazas específicas para este aspecto. Se ha seleccionado en esta tesis doctoral porque se ha convertido en una de las metodologías de análisis de amenazas para la privacidad más maduras y, por lo tanto, más utilizadas [102], [103], [104].

En el contexto de la ciberseguridad existen otras metodologías mucho más extendidas como STRIDE [105], DREAD [106], OCTAVE [107] y TARA [108]. También existen catálogos de tácticas, técnicas y procedimientos o

patrones de ataque como MITRE ATT&CK [109] o CAPEC [110]. Pero son específicas para las amenazas relacionadas con la ciberseguridad y, por lo tanto, se centran en la confidencialidad, la integridad y la disponibilidad. En el caso de la privacidad, es necesario tener en cuenta otros aspectos como el anonimato, el control o el cumplimiento de la regulación o de códigos éticos, sólo por poner algunos ejemplos. Además, es necesario tener en cuenta los impactos que las amenazas tienen para las personas, no sólo los impactos técnicos. Y que los activos que se ven amenazados, datos principalmente, no se mantienen estáticos sino que evolucionan y tiene un ciclo de vida.

Por estos motivos LINDDUN resulta el método más adecuado para esta tesis doctoral, en el que se identifican una serie de categorías a las que pertenecen las amenazas para la privacidad y que le dan nombre:

- *Linkability* permite identificar la relación entre al menos dos datos (registros, ficheros, etc.) o acciones, que corresponden con un usuario o grupo de usuarios.
- *Identifiability* permite identificar a un usuario a partir de un conjunto de datos o acciones.
- *Non-repudiation* impide a un usuario negar que ha realizado una acción, ya que existen evidencias de que un dato o acción pertenecen a ese usuario en concreto. También se incluye en esta categoría de amenazas el caso contrario, cuando existe una evidencia de que un dato o acción no pertenece, con toda seguridad, a un usuario específico.
- *Detectability* hace imposible ocultar la actividad de los usuarios.
- *Disclosure of information* cuando existe una exposición o fuga de datos que permite que un usuario no autorizado pueda verlos. Esta categoría de amenazas es la brecha de datos tradicional en los modelos de amenazas para la ciberseguridad, cuando lo que preocupa es la confidencialidad.
- *Unawareness*, cuando el usuario desconoce qué información ha expuesto a terceros, así como las consecuencias de dicha exposición.
- *Non-compliance* se corresponde con la falta de cumplimiento de leyes, normativas, políticas y consentimientos.

La metodología LINDDUN define tres fases distintas [101] cuando se realiza modelado de amenazas para la privacidad, que permiten identificar de una manera sistemática y ordenada todas las amenazas que en un sistema, producto o proyecto se producen en estas siete categorías ya mencionadas:

1. Modelado del sistema, consiste en representar de forma gráfica el sistema, producto o proyecto objeto de estudio utilizando para ello un Diagrama de Flujo de Datos (DFD, *Data Flow Diagram*). El DFD se compone de entidades, procesos, almacenamiento y flujos de datos. Como se ha comentado antes, es muy importante, en el caso de amenazas para la privacidad, comprender cómo los datos evolucionan y se transforman a lo largo del tiempo (se capturan, se almacenan, se procesan, se transfieren, etc.).

2. Identificación de amenazas para la privacidad, basándose en el DFD definido en la fase anterior. En [111] se define, para cada categoría de amenazas y tipo de elemento, un árbol de amenazas que se puede utilizar. Para ello se recomienda usar una tabla en la que se establece la correspondencia entre los diferentes elementos del DFD y las amenazas para la privacidad.

3. Mitigación de las amenazas, buscando soluciones para cada amenaza identificada en la tabla de la fase anterior. Es habitual el uso de PETs (*Privacy Enhancing Technologies*) como solución, así como la definición de prioridades, ya que no todas las amenazas tienen los mismos impactos potenciales para las personas y se debe comenzar por mitigar aquellas con impactos críticos o más significativos.

3.2. LINDDUN aplicado al modelo federado

Existen trabajos previos que han identificado una serie de amenazas para la privacidad en los esquemas federados para la gestión de identidades y accesos, como [18] ó [19]. Pero no lo hacen de manera exhaustiva y sistemática con una metodología que sea estándar *de facto* como LINDDUN, sino de una manera informal, mediante razonamientos lógicos o pruebas técnicas, y sin explicitar las asunciones o condiciones en las que el modelo de amenazas generado es o no válido. En [102] sí que se utiliza LINDDUN para realizar modelado de amenazas en un proceso de autenticación, pero el análisis que se realiza se centra en situaciones en las que el IdP es confiable y es el usuario el que intenta vulnerar la privacidad de otros usuarios. Por lo tanto, no es un punto de partida adecuado para la investigación que se pretende realizar en esta tesis doctoral.

En la presente investigación el usuario se considera confiable, ya que es el principal interesado en proteger su privacidad. En cuanto al IdP y RP, se considera que siguen el modelo *honest but curious*, tal y como se ha indicado anteriormente. Es, por tanto, necesaria la aplicación de LINDDUN para el análisis del modelo federado desde este punto de vista, con este tipo de asunciones.

3.3. Fase 1: Modelado del sistema

En la figura 3.1 se muestra el DFD propuesto como punto de partida para comprender el ciclo de vida de los datos en cualquier federación de identidades. Se identifican en color verde los elementos relativos al usuario, en color rosa los relativos al IdP y en azul los de la RP. La idea es que el modelo producido en esta tesis sea aplicable a cualquier especificación o implementación que siga del modelo federado, teniendo en cuenta estos tres roles tradicionales.

Como ya se ha mencionado el DFD se compone de entidades, procesos, almacenamiento y flujo de datos. La entidad fundamental es el usuario en este caso. El usuario es un elemento que utiliza una serie de procesos de forma directa o indirecta (desencadenados por alguna acción suya). Además, para el modelo federado se ha definido el proceso que da el servicio de identificación y autenticación al usuario, el servicio de datos de usuario, del cual se pueden obtener datos de los usuarios si se está autorizado para ello, y el servicio al usuario, que corresponde con el proceso de la RP que el usuario quiere utilizar (recurso, aplicación o servicio) y para el cual realiza el proceso de autenticación.

El almacenamiento está representado en el DFD mediante dos líneas, superior e inferior, y cada proceso y entidad hace uso del suyo propio. Por último, el flujo de datos está representado por flechas negras que identifican las comunicaciones entre la entidad, los procesos y almacenamientos.

Además, como parte la definición del DFD se indica el límite de confianza en el modelo (*Trust boundary*) representado con línea discontinua en gris. Se considera que no existen amenazas para la privacidad entre aquellos elementos que se encuentren dentro del límite de confianza. Dado el objetivo del modelo de amenazas en esta tesis doctoral, este límite de confianza incluye exclusivamente al usuario final.

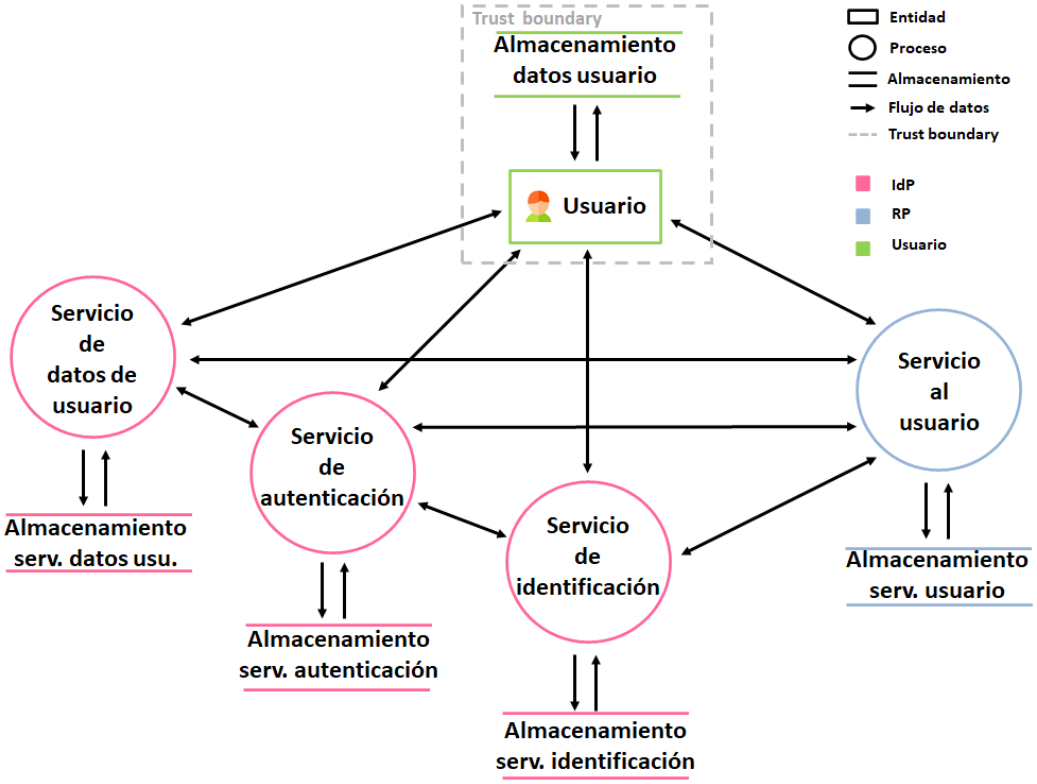

Figura 3.1: Definición del DFD en el modelo federado

3.4. Fase 2: Identificación de amenazas a la privacidad

Antes de pasar a identificar las amenazas asociadas al DFD representado para el modelo federado en la primera fase del proceso de modelado de amenazas, es necesario resumir las condiciones o asunciones que se establecen en la presente tesis doctoral dados lo objetivos que se persiguen:

- El usuario se considera confiable, por lo que está dentro del *trust boundary* del DFD. El usuario es el principal interesado en que no se vulnere la protección de sus datos o su privacidad.
- El IdP y la RP siguen el modelo *honest but curious*. Es decir, los agentes que participan en la federación de identidades no se desvían del protocolo utilizado en el modelo federado (de la especificación), pero pueden utilizar los datos que obtienen para otros fines que no son estrictamente los necesarios para el funcionamiento del protocolo.

- No se tienen en consideración los compromisos de los dispositivos del usuario, los ataques entre usuarios o los ataques desde el IdP o la RP (por ejemplo, debidos a una amenaza interna) debido a las dos asunciones anteriores.

- Las amenazas relativas a los flujos de datos no se tienen en cuenta, ya que las especificaciones federadas no determinan cómo se realizan los intercambios de mensajes entre los agentes que forman parte de la federación. De esta forma, se dejan fuera del modelo amenazas específicas relativas a implementaciones concretas del IdP o de las RP.

Una vez aclaradas estas asunciones, la tabla 3.1 muestra las amenazas para la privacidad a las que se expone cada elemento definido en el DFD para el modelo federado. Estas amenazas están agrupadas utilizando las categorías que propone LINDDUN. Los elementos que tienen como valor en la tabla un guión (-) no forman parte del análisis. En el caso de Almacenamiento de usuario (AU) y el flujo entre el usuario y el almacenamiento (U - AU), se encuentra dentro del *trust boundary*, por lo que se considera de confianza y no existen amenazas para la privacidad. Para el resto de los flujos de datos, no se analizan las amenazas para la privacidad debido a la asunción número 4, por eso también aparecen guiones en la tabla en las filas correspondientes.

Como se puede observar en la tabla 3.1, se identifican ocho amenazas diferentes para la privacidad siguiendo la metodología LINDDUN. Se pasa a continuación a explicar en más detalle cada una de ellas.

[A1] **Perfilado del usuario en el modelo federado**
Categoría: *Linkability*.
Elemento: Entidad (U) y Almacenamiento (ASA, ASI, ASD, ASU).
Resumen: Se pueden relacionar las cuentas que utiliza un usuario (entidad) entre sí y enriquecer la información que el usuario había consentido a proporcionar inicialmente (no implica necesariamente la revelación de la identidad física o real).
Objetivo: Enriquecimiento de los datos del usuario utilizando la correlación establecida entre las diferentes cuentas (dentro de la federación o fuera de ella).
Actores: IdP ó RP.
Precondiciones: El usuario tiene que tener varias cuentas registradas en IdPs, RPs o proveedores fuera de la federación. Existen datos como la combinación usuario-contraseña o metadatos (como la dirección IP o cualquier huella del dispositivo que usa el usuario o de su conexión) que permiten identificar a un usuario específico.
Pasos:

Tabla 3.1: Relación de amenazas para la privacidad sobre los elementos del DFD

Tipo	Elemento	L	I	N	D	D	U	N
Entidad	Usuario (U)	A1 A2 A3	A4				A6 A7	
Procesos	Servicio autenticación (SA)					A5		A8
	Servicio identificación (SI)					A5		A8
	Servicio datos de usu. (SD)					A5		A8
	Servicio al usuario (SU)					A5		A8
Almacenamiento	Alm. usuario (AU)	-	-	-	-	-	-	-
	Alm. serv. autent. (ASA)	A1	A4	A7	A5	A5	A7	
	Alm serv. identif. (AIS)	A1	A4	A7	A5	A5	A7	
	Alm. serv. datos usu. (ASD)	A1	A4	A7	A5	A5	A7	
	Alm. serv. al usuario (ASU)	A1	A4	A7	A5	A5	A7	
Flujo datos	U - SI	-	-	-	-	-	-	-
	U - SA	-	-	-	-	-	-	-
	U - SD	-	-	-	-	-	-	-
	U - SU	-	-	-	-	-	-	-
	SA - SI	-	-	-	-	-	-	-
	SA - SD	-	-	-	-	-	-	-
	SA - SU	-	-	-	-	-	-	-
	SI - SU	-	-	-	-	-	-	-
	SD - SU	-	-	-	-	-	-	-
	U - AU	-	-	-	-	-	-	-
	SA - ASA	-	-	-	-	-	-	-
	SI - ASI	-	-	-	-	-	-	-
	SD - ASD	-	-	-	-	-	-	-
	SU - ASU	-	-	-	-	-	-	-

1. El usuario se registra en el IdP ó RP con una o varias cuentas.
2. El IdP ó RP es capaz de relacionar estas cuentas diferentes y de enriquecer la información de la que dispone acerca de ese usuario con información disponible internamente o en fuentes externas (por ejemplo, redes sociales), integrándola en un único registro completo.

Desencadenante: Los pasos pueden ocurrir en cualquier momento debido a que el IdP y la RP tienen siempre la información disponible (de las cuentas del usuario y de su actividad) en su almacenamiento. Además, en muchos casos no se sigue el principio de minimización y los datos se

retienen demasiado tiempo.
Caso de ejemplo: Un usuario accede a un servicio de comercio electrónico (RP) autenticándose con una cuenta en Google (IdP). Posteriormente, el usuario accede al mismo servicio con otra cuenta diferente, pero también en Google. En ambas cuentas el usuario tiene dado de alta el mismo método de pago y dirección de envío de sus pedidos, por lo que tanto el IdP como la RP pueden determinar que ambos accesos pertenecen al mismo usuario (sin implicar necesariamente la revelación de la identidad física o real del usuario). Ambos proveedores pueden, a partir de ese momento, enriquecer los datos que conocen sobre ese usuario y aumentar la cantidad de información que tienen sobre él, ya que no tiene por qué haber la misma información en ambas cuentas de Google. El usuario probablemente utilice estas cuentas con diferentes fines, niveles de seguridad, etc.

[A2] **Perfilado del usuario en el uso de servicios**
Categoría: *Linkability*.
Elemento: Entidad (U).
Resumen: Se pueden relacionar las RP que utiliza un usuario (no implica necesariamente la revelación de la identidad física o real).
Objetivo: Perfilado del usuario desde el punto de vista de sus atributos, hábitos, preferencias, costumbres, etc.
Actores: IdP.
Precondiciones: El usuario tiene que tener una cuenta en el IdP y haber accedido a, al menos, una RP.
Pasos:

1. El usuario accede a una o varias RP utilizando el IdP como método de autenticación.
2. El IdP obtiene la información de las RP que utiliza el usuario.

Desencadenante: Los pasos pueden ocurrir en cualquier momento debido a que el IdP tiene siempre la información disponible en su almacenamiento.
Caso de ejemplo: Un usuario accede a distintos servicios autenticándose siempre con su cuenta de Facebook (IdP) para evitar abrir una cuenta en cada uno de ellos. Dado el funcionamiento del flujo de autenticación del modelo federado, Facebook conoce todos los servicios que está utilizando el usuario (y cuándo lo hace, desde qué dispositivo, etc.) y puede utilizar esta información para otros fines que no son estrictamente el servicio de autenticación como, por ejemplo, la publicidad personalizada.

Este perfilado no implica necesariamente la revelación de la identidad física o real del usuario.

[A3] **Perfilado del usuario según su localización**
Categoría: *Linkability*.
Elemento: Entidad (U).
Resumen: Se puede obtener la geolocalización de un usuario que utiliza un IdP (no implica necesariamente la revelación de la identidad física o real).
Objetivo: Trazabilidad de la localización de un usuario.
Actores: IdP ó RP.
Precondiciones: El usuario tiene que tener una cuenta en el IdP y no utiliza redes que anonimicen la conexión.
Pasos:

1. El usuario accede a una o varias RP utilizando el IdP como método de autenticación.
2. El IdP o RP obtiene la información de la geolocalización del usuario utilizando cualquier elemento que le permita hacerlo (como la dirección IP de conexión o elementos asociados al uso de una red de telefonía móvil).

Desencadenante: Los pasos pueden ocurrir en cualquier momento, debido a que el IdP tiene la capacidad de obtener los datos de geolocalización siempre que el usuario lo utiliza como método de autenticación.
Caso de ejemplo: Un usuario accede desde su terminal móvil a un servicio de *streaming* utilizando para ello su cuenta de Apple (IdP). Tanto el IdP como la RP podrían obtener la geolocalización del usuario utilizando la dirección IP con la que está realizando las peticiones u otros metadatos asociados al flujo de autenticación.

[A4] **Identificación del usuario en el modelo federado**
Categoría: *Identifiability*.
Elemento: Entidad (U) y Almacenamiento (ASA, ASI, ASD, ASU).
Resumen: Se pueden relacionar, de manera puntual o sostenida en el tiempo, las cuentas que utiliza un usuario (entidad) con su identidad física o real, ya que están asociadas con datos personales que permiten esta identificación.
Objetivo: Obtención de los datos personales del usuario para identificarlo en el mundo físico o real.
Actores: IdP ó RP.

Precondiciones: El usuario tiene que tener una o varias cuentas registradas en el IdP o la RP y debe proporcionar algún dato personal que permita conocer su identidad física o real.
Pasos:

1. El usuario se registra en el IdP o la RP con una o varias cuentas e introduce datos personales que le identifican de manera unívoca.
2. El IdP o la RP pasan a relacionar la cuenta o cuentas de ese usuario con la persona física o real a la que corresponden los datos personales proporcionados.

Desencadenante: Los pasos pueden ocurrir en cualquier momento debido a que el IdP o la RP tienen siempre la información disponible en su almacenamiento (de las cuentas del usuario y de su actividad) en su almacenamiento. Además, en muchos casos no se sigue el principio de minimización y los datos se retienen demasiado tiempo.

Caso de ejemplo: Un usuario accede a una revista digital (RP) utilizando para ello su cuenta de Facebook (IdP). Para el uso de algunos servicios de esta revista digital (como sorteos y promociones) se precisan datos que identifiquen al usuario (nombre, apellidos y documento de identidad). Estos datos se almacenan en el IdP, por lo que tanto Facebook (IdP) como la revista digital (RP) pueden conocer la identidad real de ese usuario en cualquier momento. La pueden obtener cuando se realiza una interacción concreta o cuando se accede posteriormente a los almacenamientos y se audita su actividad, por ejemplo.

[A5] **Revelación de datos del usuario en el modelo federado**
Categoría: *Disclosure of information* y *Detectability*.
Elemento: Proceso (SA,SI, SD, SU) y almacenamiento (ASA, ASI, ASD, ASU).
Resumen: Revelación de datos personales por fuga de información en los procesos de IdP o RP, o en cualquiera de los almacenamientos que estos procesos utilizan. Incluye la amenaza de *Detectability* de todos los almacenamientos (se sabe que un usuario tiene cuenta en un IdP, por ejemplo, aunque no necesariamente se tenga acceso a sus datos).
Objetivo: Obtención de datos personales de uno o varios usuarios.
Actores: Terceros, a través del IdP o la RP.
Precondiciones: El usuario tiene que tener una o varias cuentas registradas en el IdP o la RP.
Pasos:

1. El usuario se registra en el IdP o la RP con una o varias cuentas e introduce datos personales que le identifican.
2. El IdP o RP expone datos de los usuarios por una mala gestión en el manejo de los datos a nivel de proceso (es vulnerable) o a nivel de almacenamiento (datos no cifrados o sin protección).
3. Un tercero obtiene los datos a través de, al menos, un proceso o un almacenamiento, aunque sea simplemente conocer su existencia.

Desencadenante: Un proceso vulnerable a revelación de información o su almacenamiento no está protegido adecuadamente.
Caso de ejemplo: Un usuario tiene sus datos almacenados en Google (IdP) y se han transferido parte de ellos a un servicio de calificación de libros (RP) tras varias interacciones del usuario con este servicio (utilizando para ello su cuenta de Google). Por un fallo en la implementación en la API en Google, un atacante podría obtener todos los datos del usuario desde la RP. También podría obtener los datos del usuario si no se protege y custodia correctamente los sistemas de almacenamiento de los procesos.

[A6] **Falta de información sobre los datos almacenados**
Categoría: *Unawareness*.
Elemento: Entidad (U).
Resumen: Falta de información sobre los datos que se utilizan en el IdP. El usuario desconoce para qué se utilizan, con quién se comparten, etc.
Objetivo: Uso poco transparente de los datos de los usuarios y opacidad acerca de lo que realmente se tiene almacenado sobre ellos.
Actores: IdP.
Precondiciones: El usuario tiene que tener una o varias cuentas registradas en el IdP.
Pasos:

1. El usuario se registra en el IdP con una o varias cuentas e introduce datos personales que lo identifican. El usuario no es consciente de la cantidad y sensibilidad de datos almacenados en el IdP, ni del uso que se hace de ellos (tipo de tratamiento: naturaleza, finalidad, alcance, contexto).
2. El IdP utiliza estos datos para su propio beneficio o el de terceros con los que colabora (como otros IdP, RPs, *data brokers*).

Desencadenante: Una cantidad no controlada de datos en el IdP con asimetría en la información, inexistencia o mal uso de consentimientos

informados.
Caso de ejemplo: Un usuario crea una cuenta en Apple (IdP) e introduce todos los datos que se solicitan. Durante el proceso de creación de la cuenta el usuario acepta la política de privacidad y los términos y condiciones del proveedor. El usuario no termina de comprender lo que está aceptando y sus implicaciones (independientemente de si lee con detalle estos documentos o no). El usuario está claramente en desventaja en esta relación, existe una asimetría obvia.

[A7] **Falta de control sobre los datos almacenados**
Categoría: *Unawareness* y *Non-repudiation*.
Elemento: Entidad (U) y Almacenamiento (ASA, ASI, ASD, ASU).
Resumen: Falta de medidas que permitan que el usuario controle los datos almacenados en el IdP y tome decisiones acerca de su tratamiento. Se relaciona también con la amenaza de *Non-repudiation* de un almacenamiento. Por ejemplo, si un usuario no quiere que quede registro en el almacenamiento de un IdP de una interacción que ha realizado en el pasado con una RP específica.
Objetivo: Uso no controlado de los datos de usuario, abuso de poder y asimetría en la relación.
Actores: IdP.
Precondiciones: El usuario tiene que tener una o varias cuentas registradas en el IdP.
Pasos:

1. El usuario se registra en el IdP con una o varias cuentas e introduce datos personales que le identifican. No hay ninguna herramienta que permita ejercer al usuario sus derechos en términos de privacidad, perdiendo el control sobre ellos.
2. El IdP utiliza estos datos para su propio beneficio o el de terceros con los que colabora (como otros IdP, RPs, *data brokers*).

Desencadenante: Falta de herramientas que permitan controlar los datos en el IdP.
Caso de ejemplo: Un usuario crea una cuenta en Google (IdP) e introduce los datos que se le solicitan. Durante el proceso de creación de la cuenta el usuario acepta la política de privacidad y los términos y condiciones del proveedor. El usuario no termina de comprender lo que está aceptando y sus implicaciones (independientemente de si lee con detalle estos documentos o no). El usuario quiere, después actualizar alguno de sus datos, evitar que se transfieran a terceros, limitar su uso

o eliminarlos. Y se encuentra con que no tiene a su disposición herramientas para realizar estas tareas, ni procedimientos claros y sencillos que le garanticen que se pueden llevar a cabo en un tiempo prudencial.

[A8] **Incumplimiento de las regulaciones, leyes o políticas**
Categoría: *Non-compliance.*
Elemento: Proceso (SA,SI, SD, SU).
Resumen: Incumplimiento total o parcial de las regulaciones, leyes o políticas que afecten a la privacidad de los usuarios.
Objetivo: Ahorro de costes, opacidad en los modelos de negocio.
Actores: IdP ó RP.
Precondiciones: El usuario tiene que tener una o varias cuentas registradas en el IdP.
Pasos:

1. El usuario se registra en el IdP con una o varias cuentas e introduce datos personales que le identifican.
2. El IdP no cumple al completo la regulación, ley o política que protege los derechos de los usuarios.
3. Los datos de los usuarios no están correctamente protegidos conforme a la regulación, ley o política y/o el usuario no puede ejercer correctamente sus derechos.

Desencadenante: Desconocimiento, desinterés, mala intención.
Caso de ejemplo: Un usuario crea una cuenta en Facebook (IdP) e introduce los datos que se le solicitan. Facebook no cumple con el Reglamento General de Protección de Datos (RGPD), por lo que, cuando el usuario europeo necesita ejercer alguno de los derechos que se recogen en la regulación, no puede hacerlo.

3.5. Fase 3: Mitigación de las amenazas

Una vez terminada la fase 2, LINDDUN propone pasar a proponer estrategias y mecanismos de privacidad que puedan lidiar con ellas, evitándolas o al menos, mitigándolas.

En el capítulo 2 de esta tesis ya se han analizado trabajos previos en este sentido, mitigación de amenazas para la privacidad en modelos federados. Pero ya se han analizado también sus principales limitaciones en el contexto de esta tesis doctoral. Y queda aún más de manifiesto tras realizar el modelo de amenazas con LINDDUN de la sección anterior, ya que algunas de las

amenazas encontradas, como la 6 y la 7, relativas a la falta de información o de control por parte del usuario final, o como la 8, relativa al incumplimiento, apenas se han tratado en los trabajos previos en el área.

Por este motivo, los dos siguientes capítulos de esta tesis se centran en proponer mitigaciones para las amenazas encontradas con el enfoque centrado en el usuario, objetivo de esta investigación. En concreto se proponen modificaciones en el proveedor de identidades (portal unificado en el Capítulo 5) y la inclusión de un nuevo agente en las federaciones (*Privacy Advisor* en el capítulo 5).

En la tabla 3.2 se resume qué amenaza mitiga cada una de estas dos estrategias.

Tabla 3.2: Mitigaciones propuestas para las amenazas del modelo federado

Amenaza	Mitigación
A1	*Privacy Advisor*
A2	*Privacy Advisor*
A3	*Privacy Advisor*
A4	*Privacy Advisor*
A5	*Privacy Advisor*
A6	Portal unificado
A7	Portal unificado
A8	Portal unificado y *Privacy Advisor*

Capítulo 4

Modificaciones en el proveedor de identidades para la mitigación de amenazas para la privacidad

El modelo federado para la gestión de identidades no incluye, de manera específica, ninguna capacidad o estrategia para la protección de los datos personales en los proveedores de identidades. De hecho, la mayor parte de especificaciones que se basan en él no mencionan esta protección de datos de manera explícita. Como mucho proporcionan una serie de recomendaciones genéricas y a alto nivel que no suelen traducirse de ninguna manera concreta en los productos que las implementan. Y que, por lo tanto, no proporcionan herramientas para lidiar de manera adecuada con las amenazas identificadas en el capítulo anterior.

Por eso en este capítulo se propone una nueva aproximación para la protección de la privacidad de los usuarios cuando utilizan mecanismos federados para resolver la gestión de la identidad. La solución propuesta permite el cumplimiento del Reglamento General de Protección de Datos (RGPD) [112] con independencia del proveedor de identidad y del servicio accedido o RP.

En concreto, se identifican las capacidades específicas que se tienen que añadir en los esquemas federados para garantizar cada uno de los derechos que recoge dicha regulación en el proveedor de identidades. A continuación se propone una forma específica de cubrir estas capacidades, utilizando para ello un portal web unificado, basado en tecnologías y mecanismos muy aceptados y extendidos en la actualidad, con fácil integración con los esquemas federados actuales. Principalmente, se propone la utilización del cierre de sesión por *back-channel* y del recibo de consentimiento. La solución propuesta

se centra en realizar modificaciones en el proveedor de identidades para, en todo lo posible, empoderar al usuario final permitiéndole ejercer todos sus derechos con garantías. Por último, se implementa un primer prototipo de la solución propuesta para validarla y evaluarla.

4.1. Motivación y casos de uso

El fenómeno que se conoce como la paradoja de la privacidad hace que los usuarios de los proveedores de identidades sospechen o sepan que dichos proveedores están recogiendo sus datos personales en un modelo asimétrico u opaco que puede suponer una amenaza para su privacidad, pero que a pesar de ello no modifiquen sus hábitos [113]. Esto implica que, aunque se haga público el modelo de amenazas propuesto en el capítulo anterior, no se pueda confiar en que el conocimiento de dicho modelo, por sí mismo, mejore la protección de datos personales en las federaciones de identidades.

Por este motivo, esta tesis doctoral comienza por proponer estrategias que garanticen el cumplimiento del RGPD desde los propios proveedores de identidades. El objetivo es identificar un conjunto de capacidades esenciales que deban proveerse en el IdP, desde el diseño, e incorporarse por defecto en sus implementaciones y no únicamente cuando los usuarios las soliciten.

Con respecto a las amenazas identificadas en el capítulo 3, se propone la mitigación de las amenazas:

- Falta de información sobre los datos almacenados [A6]
- Falta de control sobre los datos almacenados [A7]
- Incumplimiento de las regulaciones, leyes o políticas [A8]

Algunos trabajos analizados en el capítulo 2 pueden ayudar a realizar esta propuesta, como punto de partida, pero los pocos que se centran en esta línea de investigación presentan algunas limitaciones fundamentales que se pueden resumir en:

- Dificultad de integración y de adopción. Los trabajos realizados hasta el momento proponen soluciones que implican cambios significativos en las especificaciones actuales o es sus implementaciones. Esto hace que su adopción sea difícil, por lo que no se suelen desplegar en entornos productivos. Otra barrera para la adopción es que, en muchos casos, se basan en el desarrollo de soluciones y componentes a medida, en lugar de en tecnologías o mecanismos estándar que se usen de forma masiva y estén lo suficientemente extendidos.

- Falta de orientación hacia el cumplimiento. En la mayor parte de los trabajos del estado del arte se proponen mejoras que incrementan los niveles de privacidad del modelo federado, pero sin permitir garantías de cumplimiento de las regulaciones actuales como el RGPD para el caso de la Unión Europea. Es decir, las mejoras se guían por criterios tecnológicos o humanos, pero no legales. Y esta dimensión es esencial en todo lo relativo a la privacidad y a la protección de datos personales. Por ejemplo, según la regulación europea el consentimiento informado del usuario cuando se realizan actividades de procesamiento de datos personales es básico. Sin embargo, en la mayoría de los trabajos previos no se aborda el problema de la gestión del consentimiento porque no se tratan los riesgos asociados al incumplimiento.

En este capítulo se pretende avanzar el estado del arte superando ambas limitaciones. Los proveedores de identidad tienen que ser capaces de atender las peticiones explícitas de los usuarios (sin un retraso indebido) cuando desean ejercer los derechos que les proporciona el RGPD. De lo contrario, pueden verse expuestos a quejas ante las autoridades supervisoras y, por tanto, a multas o pérdidas de reputación. Además, los usuarios finales pueden sufrir los impactos asociados a todas las amenazas que tienen que ver con los diferentes tipos de incumplimiento. Para que los proveedores incorporen esta capacidad, se debe proponer una solución fácil de adoptar e integrar con las especificaciones actuales y sus implementaciones, capaz de gestionar el consentimiento de los usuarios.

Los siguientes escenarios pueden ayudar a comprender mejor la motivación de este capítulo:

Escenario 1 Un usuario europeo quiere utilizar Facebook como proveedor de identidades para iniciar sesión en diferentes aplicaciones. El usuario es consciente de la importancia de la privacidad de sus datos, y le gustaría poder dar respuesta a las siguientes preguntas: ¿Qué tipo de datos está procesando Facebook? ¿Cuál es el fundamento jurídico o base de legitimación para el procesamiento de sus datos?¿Cuánto tiempo va a almacenar el proveedor los datos? ¿Va a compartir Facebook la información con otras organizaciones o empresas? ¿Qué visibilidad tendrá el usuario de los procesos de compartición de datos que se produzcan? Además, el usuario quiere ejercer el derecho de acceso (para obtener una copia de sus datos) y de rectificación (para corregir errores en sus datos o para actualizarlos cuando sea necesario).

Escenario 2 Un usuario europeo quiere utilizar el servicio de inicio de sesión con Apple en diferentes servicios y aplicaciones. Sin embargo, quiere

poder ejercer en el futuro el derecho de oposición o restricción a marketing directo (u otros tipos de procesamiento en el proveedor de identidad o RP), quiere poder trasladar sus datos a otro proveedor de identidad (incluyendo el histórico de sus datos sobre la información compartida con otras RP) o quiere poder borrar sus datos.

Escenario 3 Una compañía de SaaS (*Software as a Service*) que confía en GSuite quiere simplificar el proceso de autenticación de sus usuarios utilizando el inicio de sesión con Google en todas sus aplicaciones y servicios corporativos (marketing, desarrollo, recursos humanos, etc.). Sin embargo, esta compañía requiere tener un control preciso sobre los datos compartidos y una trazabilidad detallada de la información compartida con las diferentes RP. Por ejemplo, para la realización de auditorías de seguridad o para la depuración de responsabilidades tras una fuga de datos.

4.2. Modelo a alto nivel para el cumplimiento de RGPD en el esquema federado

El Reglamento General de Protección de Datos (RGPD) [112] es la regulación actual para la protección de datos personales en la Unión Europea. Está diseñada para garantizar una mejor protección de la privacidad y derechos de los ciudadanos europeos y, por tanto, armonizar la ley en todos los países de la Unión Europea. Actualmente, está considerada como la normativa más importante en materia de protección de datos personales. Esta normativa define una serie roles dentro del proceso de protección de los datos personales así como una serie de derechos para todos los ciudadanos europeos.

4.2.1. Roles en el RGPD

El RGPD define cinco roles diferentes para la protección de los datos personales. La figura 4.1 los muestra representados en el escenario de los esquemas federados.

El primero es el interesado, que es la persona física dueña de los datos personales que le identifican o permiten identificarle de forma directa o indirecta. En los esquemas federados, el interesado se corresponde con un usuario europeo.

El segundo es el responsable del tratamiento, que es la persona física o jurídica que define el propósito y los mecanismos del procesamiento de los datos. Este rol es el responsable del cumplimiento del RGPD aplicando tanto medidas técnicas como organizativas.

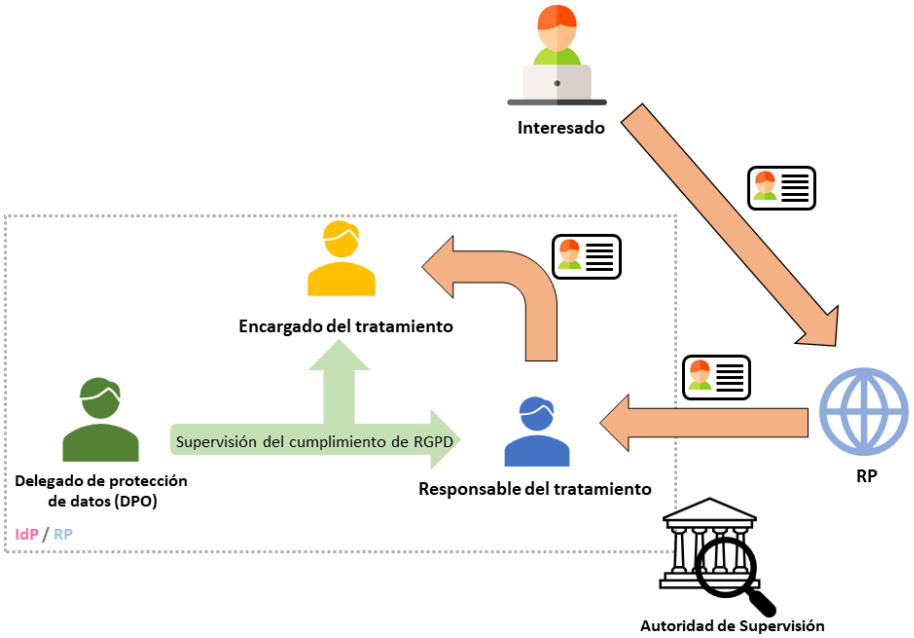

Figura 4.1: Roles RGPD en el esquema federado

El tercero es el encargado del tratamiento, que es la persona física o jurídica que procesa los datos personales en nombre del responsable del tratamiento. Este rol debe cumplir las condiciones especificadas por el responsable del tratamiento firmadas en un contrato o acto jurídico.

En el esquema federado, el responsable del tratamiento es el IdP o la RP y el encargado del tratamiento puede ser también el IdP o la RP. O incluso puede ser otra entidad que no esté incluida en el flujo de identificación y autenticación.

La cuarta es la autoridad de control, que es una organización pública designada por un estado miembro responsable de supervisar el cumplimiento del RGPD aconsejando a las compañías, auditando su cumplimiento, gestionando las quejas de los interesados, etc. Como los proveedores de identidad normalmente operan en varios países de la Unión Europea, habitualmente designan como punto de comunicación una autoridad de control principal. De esta forma, se simplifica la gestión en el cumplimiento de la regulación, entre otros procesos.

La quinta es el delegado de protección de datos DPO, por sus siglas en inglés, *Data Protection Officer*. Esta figura no es obligatoria, sino que depende de las particularidades de la empresa. Además, es responsable de velar por el

cumplimiento del RGPD en la compañía y asesorarla sobre la mejor estrategia a seguir y realizar. Los proveedores de identidad necesitan este tipo de rol debido a que tratan a gran escala datos personales de ciudadanos europeos.

4.2.2. Derechos del usuario en el RGPD

En los esquemas federados se aplican tratamientos de sus datos personales a todos los usuarios pertenecientes a la Unión Europea. Los ciudadanos europeos tienen una serie de derechos RGPD que tanto IdP como RP tienen que garantizar. Para ejercer sus derechos, los usuarios tienen que realizar las peticiones explícitas al responsable del tratamiento. Los derechos de los interesados son el derecho de información, derecho de acceso, el derecho de oposición, el derecho de rectificación, el derecho de supresión, el derecho a la limitación del tratamiento, el derecho a la portabilidad y el derecho a no ser objeto de decisiones individuales automatizadas.

El derecho de información (artículos 13 y 14 y considerandos 60, 61 y 62) permite a los usuarios conocer quién es el IdP o la RP y cómo contactar con ellos, su DPO, el propósito e intereses de los datos, destinatarios, periodos de retención y transferencias internacionales de datos personales.

El derecho de acceso (artículo 15 y considerandos 63 y 64) permite a los usuarios solicitar al IdP o RP una copia de sus datos personales almacenados, además de otra serie de datos adicionales, sin dilación indebida.

El derecho de oposición (artículo 21 y considerandos 69 y 70) permite a los usuarios oponerse al procesamiento de los datos cuando no hay motivos legítimos para ello.

El derecho de rectificación (artículo 16 y considerando 65) permite a los usuarios solicitar la corrección de sus datos personales al IdP o RP cuando éstos sean incompleto o inexactos.

El derecho de supresión, también conocido como derecho al olvido (artículo 17 y considerandos 65 y 66), se puede ejercer cuando los datos personales de los usuarios ya no son necesarios para el fin para el que fueron obtenidos, cuando el interesado retira un consentimiento previamente otorgado, cuando se ejerce el derecho de oposición, cuando el tratamiento no es lícito o en el cumplimiento de una obligación legal.

El derecho a la limitación del tratamiento (artículo 18 y considerando 67) permite a los usuarios suspender el tratamiento de datos personales y la conservación de los mismos. En el caso de la suspensión del tratamiento, se realizará en base a una impugnación por inexactitud, hasta su verificación, o cuando se ejerce el derecho de oposición, hasta que se determina si aplica. En el caso de la conservación de los datos, se realizará cuando el tratamiento no

sea lícito pero el usuario no desea su eliminación, sino la limitación del uso y cuando el responsable no necesite los datos personales pero sí el interesado.

El derecho a la portabilidad (artículo 20 y considerando 68) permite a los usuarios obtener los datos personales en un formato estructurado y procesable cuando el tratamiento se realiza por medios automatizados con el objetivo de transmitir dichos datos a otro responsable, como por ejemplo, a otro proveedor de identidad.

Por último, el derecho a no ser objeto de decisiones individuales automatizadas (artículo 22 y considerandos 71 y 72) permite a los interesados no ser objeto de decisión de forma exclusivamente automatizada, siempre que se produzcan efectos jurídicos sobre ellos o le afecten de forma similar. Este derecho no aplica para el modelo federado ya que no suele incorporar este tipo de decisiones automatizadas (pero si podría ser algo a considerar en un futuro si, por ejemplo, los proveedores de identidades toman parte en decisiones relacionadas con el crédito bancario o las pólizas de seguros, por ejemplo).

4.3. Capacidades del modelo

Teniendo en cuenta los derechos definidos por el RGPD, esta investigación propone que el IdP incorpore una serie de capacidades para asegurar su cumplimiento, principalmente estas tres:

1. Repositorio de información centralizado donde los usuarios puedan visualizar sus datos personales junto con información adicional relacionada con el RGPD que sea de relevancia para los usuarios y donde, además, puedan ejercer sus derechos. Este repositorio tiene que estar siempre actualizado, ser de fácil uso y visualización, y ser conciso para ayudar a los usuarios a entender sus derechos, ayudándoles en el proceso de ejercerlos. De esta forma, el IdP y la RP también pueden cumplir más fácilmente con la regulación, en cuanto a sus obligaciones de facilitar el ejercicio de los derechos de los usuarios (considerando 59), mostrar la información del usuario que se está procesando (considerando 61) y proveer un acceso seguro remoto a los datos personales de los usuarios (considerando 63). El IdP y la RP tienen, de forma centralizada, todo lo necesario para informar a sus usuarios o, de forma sencilla, transferir los datos personales a otros IdP o RP en caso de ser necesario.

2. Revocación de *tokens* para forzar el cierre de sesión de una RP desde un IdP sin la necesidad de interacción del usuario final. Esta capacidad es

necesaria para el cumplimiento de algunos derechos como, por ejemplo, en el derecho a la supresión, si el *token* tiene datos personales afectados por este derecho.

3. Recibos de consentimiento, entendidos como documentos estándar que resumen qué datos personales se han enviado a quién (IdP, RP u otra tercera entidad) y con qué propósitos. Estos recibos de consentimiento se tienen que almacenar en el repositorio de información centralizada, con un formato que sea procesable por máquinas y que esté disponible para su descarga.

4.4. Arquitectura para el cumplimiento del RGPD

Teniendo en cuenta las capacidades definidas en la sección anterior, para el cumplimiento de los derechos RGPD en el esquema federado se propone el uso de un portal web, unificado y estándar que se gestione desde el IdP. Desde el punto de vista del IdP es la forma más sencilla de mantener toda la información necesaria en un mismo sitio y, desde el punto de vista de los usuarios, es la mejor forma de consultar todo lo necesario para el cumplimiento de sus derechos. Por tanto, esta web unificada contiene tanto información relativa a los datos personales de los usuarios como la posibilidad de que éstos ejerzan sus derechos a través de ella.

Además, en el esquema federado el IdP tiene que realizar una serie de acciones adicionales para garantizar el correcto cumplimiento de los derechos RGPD, así como de las capacidades de privacidad definidas.

Para el derecho de información (figura 4.2) se propone el uso de una web unificada, que muestre a los usuarios toda la información requerida para ejercer este derecho. Desde esta web unificada, el usuario puede seleccionar este derecho para ejercerlo donde se le indicará las implicaciones de ejercer este derecho y que, por cómo está diseñado este portal web, ya contiene toda la información necesaria en su pantalla principal. La pantalla principal informa al usuario sobre los datos personales recogidos, periodo de retención y origen de los datos, procesamiento de los datos y motivos de este procesamiento, destinatario y transferencias internacionales de datos e información sobre IdP y RP, incluyendo el DPO y cómo contactar con dicha figura. Además, esta web unificada almacena los recibos de consentimiento así como un registro de los derechos ejercidos y estado de dichas peticiones. De esta forma, se ayuda al usuario a seguir las peticiones de forma sencilla y a saber si se han aceptado o no. También se almacena un registro de las acciones realizadas

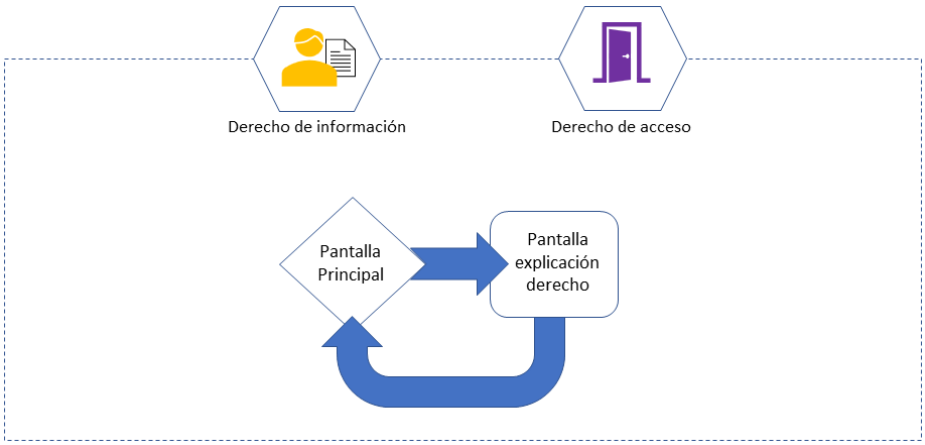

Figura 4.2: Derecho de información y acceso para esquemas federados

por el IdP para el cumplimiento del RGPD.

Para el recibo de consentimientos se propone el uso de *Kantara Consent Receipt* [114]. Siguiendo la especificación, los recibos de consentimiento son legibles para los humanos y se puede representar en JSON, lo que es una ventaja en este tipo de escenario. Por un lado, los usuarios pueden consultar con recibos siempre que quieran hacerlo. Por otro lado, el formato JSON ya se utiliza actualmente en los flujos de autenticación del modelo federado, lo que permite que sea más fácil de utilizar para el IdP y RP en lo que a escritura, lectura, actualización, procesamiento, almacenamiento y transmisión se refiere.

Para el derecho de acceso (figura 4.2) se propone una solución análoga que para el derecho de información, donde se muestre toda la información necesaria relativa a este derecho en la pantalla principal de la web unificada. Además, tener toda la información de ambos derechos en una misma pantalla integrada permite a los usuarios entender mejor cómo manejan los datos personales el IdP y RP. Al igual que en el derecho de información, el usuario puede, desde la pantalla principal, seleccionar ejercer el derecho de acceso. Será informado sobre este derecho y se le indicará que dispone de toda la información en la pantalla principal de la web unificada.

Para el derecho de oposición (figura 4.3) se propone que, una vez el usuario ha seleccionado ejercer dicho derecho en la web unificada, se muestra una explicación de dicho derecho. En dicha explicación se tiene que incluir que la solicitud puede implicar el cierre de sesión en aquellas RP afectadas por esta solicitud. Después, el usuario puede seleccionar los procesamientos de datos sobre los que desea oponerse. Es posible que algunos de estos procesamientos no se puedan seleccionar si existen motivos legítimos para el procesamiento de

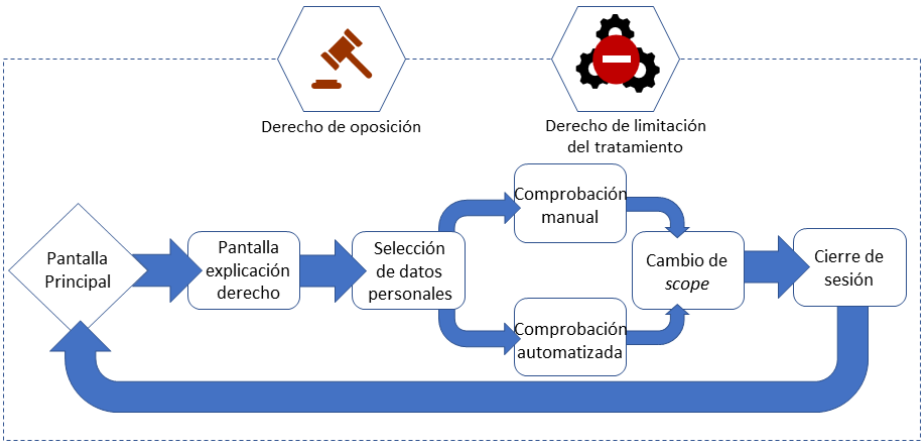

Figura 4.3: Derecho de oposición y limitación del tratamiento para esquemas federados

dichos datos que se impongan a los derechos de los usuarios, tal y como indica la normativa RGPD. Además, se tendrán que explicar claramente dichos motivos. Una vez el usuario ha seleccionado los procesamientos de datos a los que quiere oponerse, el IdP tiene que comprobar, mediante procesos manuales o automáticos, dicha petición. En caso de ser aceptada dicha decisión, el IdP deberá modificar los *scopes* del *ID Token*. Para ello, se propone la utilización del cierre de sesión por *back-channel* definida para OpenID Connect [14] en aquellas RP afectadas por la petición del derecho de oposición.

Se propone un flujo de cierre de sesión de una RP desde el IdP sin la necesidad de interacción del usuario, tal y como se muestra en la figura 4.4 (se recuerda que el detalle de este flujo de comunicaciones está descrito en el apartado 2.2.1). Para ello, es necesario realizar peticiones como las descritas en la figura 4.5. Tras este cierre de sesión, el IdP modificará el *scope* para evitar los tratamientos de datos a los que el usuario se opone. El usuario deberá volver a iniciar sesión cuando vuelva a utilizar las RP, pero el *token* JWT que se genere ya contemplará estos cambios que se han realizado en el *scope*. En el caso en el que la petición de ejercer el derecho se deniegue, el IdP informará al usuario de su denegación, así como de los motivos de la misma.

Para el derecho de limitación del tratamiento (figura 4.3) el proceso es análogo al de oposición. El usuario, tras seleccionar el derecho en la web unificada, accede a una explicación del mismo. La explicación tiene que contener que se cerrará la sesión en las RP afectadas por esta solicitud. Después, el usuario puede seleccionar los tratamientos que quiere limitar, teniendo en cuenta que algunos de éstos pueden no ser seleccionables si existen motivos

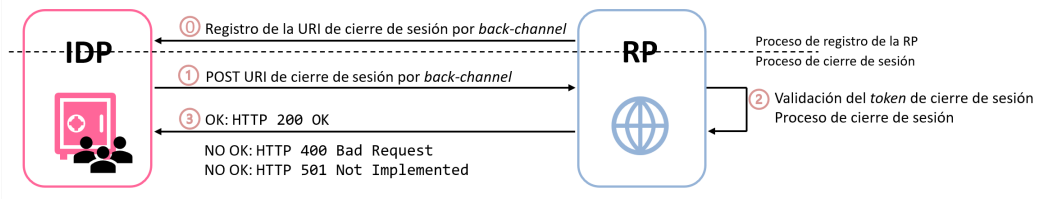

Figura 4.4: Flujo de cierre de sesión por *back-channel*

Petición de cierre de sesión por *back-channel*

```
POST /logout HTTP/1.1
  Host: rp1.com
  Content-Type: application/x-www-form-urlencoded
  Cache-Control: no-cache
  no-store Pragma: no-cache
```

***Token* JWT para el cierre de sesión**

```
{
  "iss": "https://idp.com",
  "sub": "248289761001",
  "aud": “pqR9hmn4rEw",
  "iat": 1611229023,
  "jti": "RftV",
"events": { "http://schemas.openid.net/event/backchannel-logout": {} }
}
```

Figura 4.5: Ejemplo de petición HTTP de cierre de sesión (parte superior) y de *token* JWT de cierre de sesión (parte inferior)

que justifiquen el tratamiento imponiéndose sobre los derechos del usuario. En este tipo de casos, se deberán justificar y explicar los motivos en la web. Tras esta solicitud, el IdP deberá comprobar de forma manual o automática la petición y, en caso de ser aceptada, se procederá a realizar los cambios necesarios en el *scope* para evitar estos tratamientos de datos que se quieren limitar. Para ello, el IdP en primer lugar deberá realizar un cierre de sesión por *back-channel*, siguiendo el flujo descrito anteriormente. El IdP modificará el *scope* para ajustarse a la limitación del tratamiento que se tiene que aplicar y cuando el usuario vuelva a iniciar sesión, ya estará modificado para cumplir con el RGPD. En el caso en el que se deniegue la solicitud, se le indicará la denegación y motivos de ésta.

En el esquema federado, los datos están almacenados en el IdP por lo que en el caso del derecho de rectificación (figura 4.6), los datos se tienen que corregir en el IdP. Se propone que el usuario pueda modificar los datos directamente en el IdP. Para ello, una vez que el usuario ha solicitado ejercer el derecho en la web unificada, se le muestra una explicación del derecho. En esta explicación se deberá indicar que el proceso de modificación de los datos personales puede implicar el cierre de la sesión en algunas RP que utilicen esos datos. De esta forma, se evita que las RP utilicen datos obsoletos y que probablemente no sean de utilidad para el usuario. Después, se muestran los datos personales del usuario para que pueda corregirlos y, tras la modificación, el IdP debe validar que los datos personales son correctos. Esta validación se puede hacer de forma automática o manual. Un ejemplo de una validación automática podría ser enviar un correo electrónico o SMS cuando se modifique la cuenta de correo electrónico o el número de teléfono móvil, respectivamente, que el usuario tiene que introducir en la web para verificar que la modificación es correcta. Un proceso de validación manual puede ser, por ejemplo, que un agente del centro de asistencia llame al usuario para verificar la veracidad de dichos datos. En el caso en el que los datos personales modificados estén almacenados en el *token* que utilizan las RP, el IdP puede esperar a la expiración de dicho *token*, si el tiempo no es excesivamente largo, o forzar el cierre de la sesión del usuario con el proceso de cierre de sesión por *back-channel*. El IdP también podrá ofrecer a los usuarios la opción de elegir con cuál de los dos procesos quiere realizar la modificación dependiendo de la urgencia que tengan éstos por actualizar dichos datos. Tras finalizar el proceso, los usuarios podrán visualizar en la web unificada que los datos han sido correctamente actualizados.

Para el derecho de supresión (figura 4.7), también conocido como derecho al olvido, se propone que el usuario, tras seleccionar ejercer el derecho en la web unificada, pueda leer una explicación del derecho. Esta explicación tiene que incluir que se cerrará sesión en aquellas RP que utilicen los datos

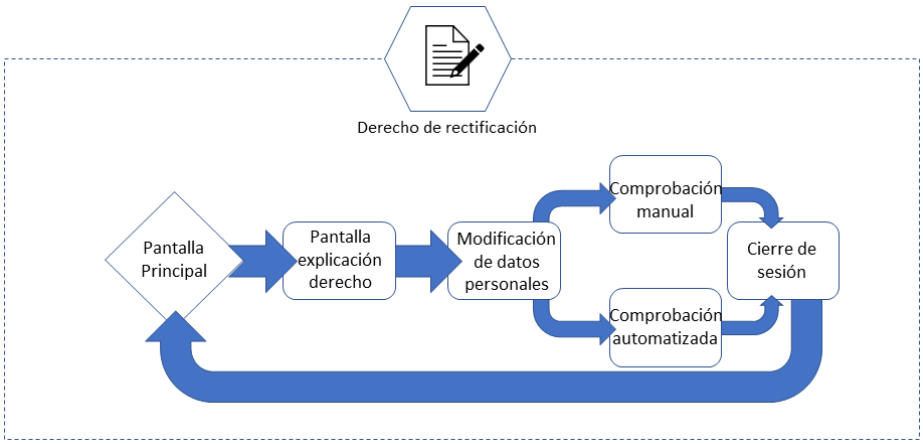

Figura 4.6: Derecho de rectificación para esquemas federados

personales que se van a borrar y se incluirá el tiempo necesario para que el proceso de borrado de datos se complete, incluyendo las copias de seguridad que pueda haber de estos datos personales. Tras la explicación, el usuario puede seleccionar los datos personales que desea eliminar, teniendo en cuenta que alguno de estos datos pueden no ser seleccionables si existen motivos legítimos para mantenerlos que estén por encima del derecho de los usuarios. Tras la selección, se indica que la solicitud se ha completado y se retorna a la web unificada. Por su parte, el IdP procesa la petición y valida, de forma manual o automática, si dichos datos se pueden eliminar. En el caso en el que sí, el IdP realizará el cierre de la sesión por *back-channel* en las RP que utilicen los datos personales eliminados y se procede al borrado de los datos personales en el IdP siguiendo sus propios procedimientos de actuación. Además, se debe tener en cuenta que el proceso de borrado de los datos personales en las copias de seguridad, si fuera viable. Los usuarios pueden comprobar que se han borrado los datos personales porque desaparecen de la web unificada.

En el caso del derecho a la portabilidad (figura 4.8), los usuarios pueden obtener un fichero estructurado, estándar y procesable por máquinas con sus datos personales. Se propone que, tras seleccionar el derecho en la web unificada, se pueda leer una explicación de este derecho. Después, el usuario puede seleccionar la opción de obtener el fichero con sus datos personales o transferirlo a otra IdP que pueda estar integrada con el IdP al que se le solicita este derecho. Tras este paso, el usuario vuelve a la pantalla principal de la web unificada.

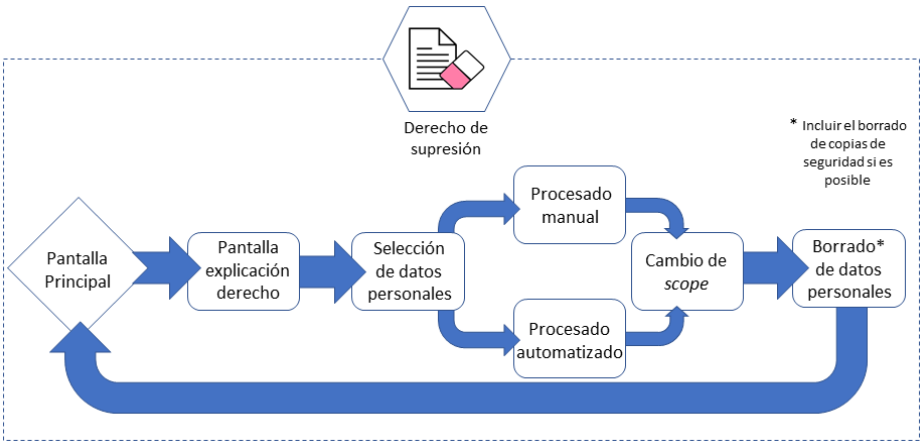

Figura 4.7: Derecho de supresión para esquemas federados

Figura 4.8: Derecho a la portabilidad para esquemas federados

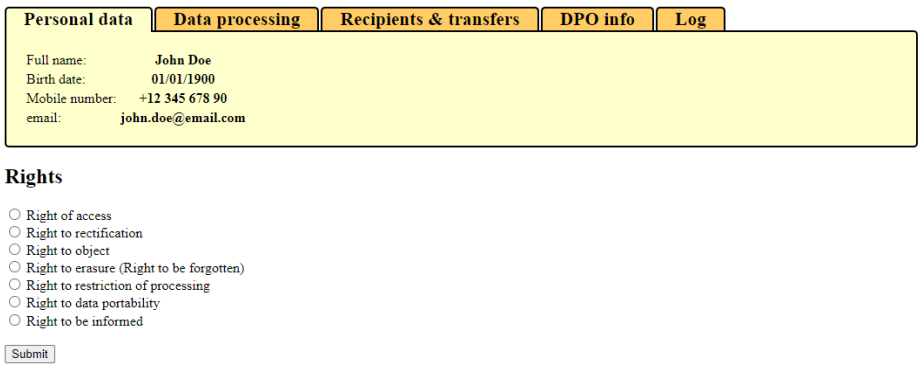

Figura 4.9: Portal web unificado

4.5. Implementación del prototipo

El prototipo implementado para la validación de la solución propuesta se ha realizado en HTML (*HyperText Markup Language*), JavaScript y PHP (*PHP: Hypertext Preprocessor*) ejecutando en un servidor XAMPP (*cross-platform, Apache, MySQL, PHP and Perl*) en Windows 10. También se utiliza JSON como formato estructurado de datos, que es compatible con la especificación del modelo federado. Se ha implementado basándose en un IdP que utiliza OpenID Connect.

El portal unificado (figura 4.9) muestra la información distribuida en pestañas.

- La primera pestaña incluye los datos personales del usuario, recibos de consentimiento, origen y periodo de retención.
- La segunda pestaña muestra los procesamientos de los datos y los motivos de los mismos.
- La tercera pestaña incluye los destinatarios y las transferencias internacionales.
- La cuarta pestaña muestra información sobre el IdP y su DPO.
- La quinta pestaña muestra un registro de los derechos ejercidos, incluyendo los datos personales o tratamientos de datos afectados, el estado de la petición, cuándo se ha solicitado y cuándo se ha resuelto.

En la parte inferior de la web unificada hay disponible un formulario donde los usuarios pueden realizar un nueva solicitud para poder ejercer sus derechos.

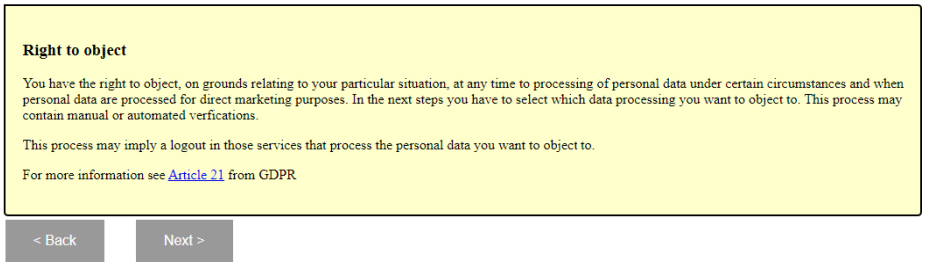

Figura 4.10: Ejemplo de explicación de un derecho

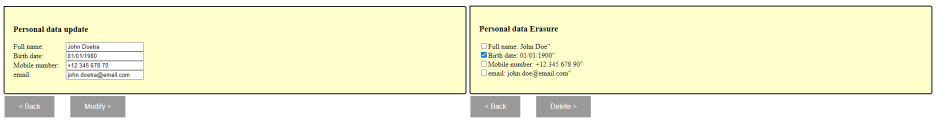

Figura 4.11: Modificación (izquierda) y selección (derecha) de datos personales

Si los usuarios quieren ejercer un derecho, pueden seleccionar el derecho que quieren ejercer y enviar el formulario. En la siguiente pantalla (figura 4.10), el usuario puede leer una explicación del derecho, así como posibles implicaciones que pueda tener (por ejemplo, el cierre de sesión en algunas RP). Una vez que han sido informados, el flujo de la petición varía dependiendo del derecho que se esté solicitando. Si se trata del derecho de acceso o el derecho de información, se vuelve a la pantalla principal (figura 4.9). Si se trata del derecho a la portabilidad, se genera un JSON con todos los datos del usuario y se descarga. Después, se vuelve a la pantalla principal.

Para el derecho de rectificación, tras la explicación del derecho (figura 4.10) aparece una nueva ventana donde el usuario puede corregir sus datos personales (parte izquierda de la figura 4.11). Tras la corrección, se muestra una pantalla donde se produce el proceso de verificación (manual o automática) de los datos personales (figura 4.12) y después, se vuelve a la página principal.

Para el derecho de oposición o limitación del tratamiento, tras la explicación del derecho (figura 4.10), el usuario puede seleccionar los tratamientos a los que quiere oponerse (figura análoga a la parte derecha de la figura 4.11 pero con tratamiento de datos) y, después, se realiza la verificación, manual o automática, del proceso (figura 4.12). Para finalizar se vuelve a la pantalla principal (figura 4.9). Para el derecho de supresión el procedimiento es análogo al descrito anteriormente.

El procesamiento de las peticiones de los usuarios puede implicar la realización de tareas adicionales en el IdP. Por ejemplo, en el derecho de oposición

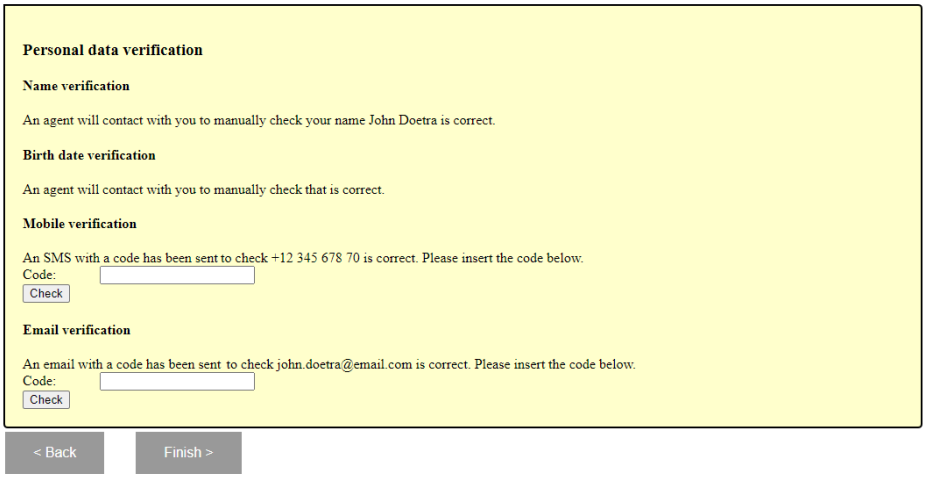

Figura 4.12: Proceso de verificación de los datos personales

y en el de restricción del procesamiento el IdP comprueba qué RP están realizando esos procesamientos de datos. El IdP tiene que realizar un cierre de sesión por *back-channel* enviando una petición HTTP POST (figura 4.5) con un *token* JWT de cierre de sesión a la URI de cierre de sesión de la RP. La RP tiene que responder si el proceso se ha completado satisfactoriamente o no. En el derecho de rectificación, el IdP actualiza los datos de los usuarios finales y, en el caso en el que esta información se encuentre en un *token* JWT válido, se realiza también un cierre de sesión por *back-channel*. En el derecho de supresión, el IdP elimina los datos personales solicitados y realiza un cierre de sesión por *back-channel* en aquellas RP que tengan un *token* JWT válido con dicha información.

4.6. Validación y discusión

Desde el punto de vista funcional, la solución propuesta provee al IdP de los mecanismos suficientes para el cumplimiento de la normativa RGPD. Esto se debe a las siguientes razones:

- Garantiza la legalidad y la equidad. La web unificada propuesta permite a los usuarios ejercer sus derechos en relación con la protección de sus datos personales de forma estandarizada en todos los IdP. Los tres escenarios propuestos en la sección 4.1 se pueden solucionar garantizando el cumplimiento del RGPD gracias a la información disponible en el portal, los permisos para editar, borrar o descargar (o traspasar a otro IdP) los datos personales y la capacidad de oponerse o limitar

el tratamiento de datos personales, incluyendo la revocación de *tokens* válidos y que se utilizan en una o varias RP.

- Facilita la seguridad, la transparencia y la auditoría. Si la autenticación de los usuarios en el IdP se asume como segura (cada vez más se añade autenticación multi-factor para las tareas críticas o incluso autenticación continua). La solución propuesta es segura, ya que únicamente los usuarios autenticados y autorizados por el IdP pueden acceder a la web unificada y ejercer los derechos. Toda la información adicional incluida en el portal como, por ejemplo, la explicación de los derechos de los usuarios o la información de contacto del DPO, mejoran la transparencia dentro de la federación de identidades. Por último, el registro de los derechos ejercidos y el registro de los consentimientos mejoran considerablemente la función de auditoría y proporcionan una trazabilidad que hasta ahora no había sido posible.

- Facilita las adaptaciones, extensiones y ampliaciones. Si la propuesta realizada en esta tesis doctoral se tuviera que adaptar a un marco normativo diferente que el RGPD europeo, bastaría con modificar los derechos que el portal unificado garantiza, adaptando su nomenclatura y los flujos que los soportan (algo muy sencillo al estar basado en tecnologías web fácilmente integrables con los esquemas federados - HTTP, JSON -). De la misma forma, si el propio RGPD evoluciona y extiende o amplía estos derechos, el esfuerzo para evolucionar el portal unificado no debería ser elevado. Por último, hay que mencionar que la propia funcionalidad del portal es fácilmente extensible. Por ejemplo, en el prototipo implementado para su validación, se han añadido a las explicaciones de los derechos algunos consejos o recomendaciones básicos y estáticos (siempre los mismos para todos los usuarios en todas las situaciones), que podrían dar lugar a versiones más sofisticadas del portal en el futuro.

Teniendo en cuenta los razonamientos descritos, se puede afirmar que se mitigan las amenazas identificadas en la sección 4.1 y que se justifican en la tabla 4.1.

Tabla 4.1: Mitigaciones que permite el portal unificado a las amenazas para la privacidad identificadas en el capítulo 3

Amenaza	Descripción	Mitigación
A6	Falta de información sobre los datos almacenados	Con la utilización del portal unificado, se garantiza el cumplimiento del RGPD y que el usuario puede ejercer sus derechos de información y acceso.
A7	Falta de control sobre los datos almacenados	Con la utilización del portal unificado, se garantiza el cumplimiento del RGPD y que el usuario puede ejercer sus derechos de rectificación, oposición, supresión, limitación del tratamiento y portabilidad.
A8	Incumplimiento de las regulaciones, leyes o políticas	Con la utilización del portal unificado tal y como se ha propuesto, se garantiza directamente el cumplimiento del RGPD. Si fuera necesario cumplir con otra regulación, ley o política bastaría con rediseñar la arquitectura del portal unificado para incluir las capacidades requeridas que no estén cubiertas en el diseño actual.

El prototipo demuestra que la solución propuesta constituye una herramienta válida, que permite a los interesados ejercer sus derechos en las soluciones basadas en esquemas federados, sin la necesidad de realizar modificaciones en las especificaciones. Para ello se emplean tecnologías y mecanismos accesibles para todos los proveedores y que son ampliamente extendidos. Sin embargo, como se discutirá en el último capítulo de esta tesis doctoral, estas soluciones requieren una mayor investigación en dos vías: incentivos para el IdP e incentivos para los usuarios, como en cualquier solución de privacidad centrada en el usuario.

Capítulo 5

Nuevo agente en la federación de identidades para la mitigación de amenazas para la privacidad

En el capítulo anterior se ha comenzado a explorar el concepto de privacidad centrada en el usuario para las actuales federaciones de identidades, realizando modificaciones sencillas en los proveedores de identidades de manera que se garantice que los usuarios europeos pueden ejercer los derechos recogidos en el RGPD.

Al final del capítulo se apuntaba una posibilidad que se explora en profundidad en el presente capítulo: proporcionar consejos o recomendaciones a los usuarios acerca de cómo tomar las mejores decisiones para su privacidad.

Ya se ha comentado con anterioridad que, debido a la paradoja de la privacidad, informar a los usuarios acerca de los riesgos que corren, exclusivamente, no suele ser suficiente. Es por eso, que en este capítulo se pretende proponer un sistema de recomendación que les ayude en su toma de decisiones. No tiene por qué realizarse desde el proveedor de identidades, del que muchos usuarios desconfiarán (con razón), sino que este sistema de recomendación podrá ser un nuevo agente en las federaciones de identidades sin que su aparición implique ningún cambio significativo en las especificaciones actuales.

De nuevo la estructura del capítulo se basa en realizar una propuesta de alto nivel basada en unas capacidades y arquitectura genéricas, para detallar a continuación cómo cubrir estas capacidades y validar las propuestas realizadas mediante la implementación de un primer prototipo.

5.1. Motivación y casos de uso

En esta tesis doctoral se persigue el compromiso del usuario con la protección de su propia privacidad: una correcta y apropiada comunicación de los riesgos de privacidad permite evitar amenazas o mitigar sus impactos, pero debe ir acompañada de otras capacidades como la de ejercer los derechos RGPD de manera sencilla (capítulo anterior) o la de disponer de recomendaciones personalizadas que ayuden en la toma de decisiones (capítulo actual).

En esta tesis doctoral se propone incluir en las federaciones de identidad un nuevo agente, el *Privacy Advisor*, capaz de recoger información e informar a los usuarios sobre las prácticas que se realizan para la protección de sus datos personales, proveer un servicio personalizado de recomendación para la toma de decisiones e incluso actuar en nombre de los usuarios en casos específicos. Este tipo de aproximación extiende las arquitecturas de privacidad tradicionales más allá del IdP o RP. Hay que destacar que la personalización de la recomendaciones es esencial ya que en trabajos previos se ha demostrado que aumenta significativamente la intención de los usuarios de seguir las recomendaciones proporcionadas porque incrementa en gran medida la confianza en su utilidad [115], [116], [117].

Con este nuevo agente se pretende mitigar las siguientes amenazas modeladas en el capítulo 3:

- Perfilado del usuario en el modelo federado [A1]
- Perfilado del usuario en el uso de servicios [A2]
- Perfilado del usuario según su localización [A3]
- Identificación del usuario en el modelo federado [A4]
- Revelación de datos del usuario en el modelo federado [A5]
- Incumplimiento de las regulaciones, leyes o políticas [A8]

El *Privacy Advisor* se puede ofrecer en el propio proveedor de identidad, como un servicio con valor añadido (por ejemplo, a través del portal unificado propuesto en el capítulo anterior). En aquellos casos en los que la neutralidad del IdP no esté garantizada, también puede ser un nuevo agente en el esquema federado, completamente independiente a los tres roles tradicionales de este tipo de esquemas. Lo pueden ofrecer tanto organizaciones sin ánimo de lucro, como grupos que pretenden proteger los derechos de los usuarios en Internet, agencias gubernamentales u otras empresas privadas con modelos de negocio diferentes a los del IdP o las RP.

De nuevo se ejemplifica la utilidad de la propuesta que se realiza en este capítulo, el *Privacy Advisor*, mediante diferentes escenarios o casos de uso:

Escenario 1 Un usuario quiere utilizar un servicio que necesita autenticación y utiliza Facebook como proveedor de identidades para iniciar sesión, de manera que se ahorra crear una cuenta nueva para ese servicio. Sin embargo, el usuario desea proteger su privacidad lo máximo posible y desconoce si existen vulnerabilidades conocidas y no resueltas en Facebook o en el servicio, si Facebook y el servicio cumplen las regulaciones acerca de protección de datos o si tienen algún certificado de terceros relacionado con su manera de gestionar la privacidad.

Escenario 2 Un usuario quiere utilizar un servicio autenticado del ámbito de la medicina y la salud. Para ello, puede realizar la autenticación con Google o con Facebook. El usuario necesita introducir datos sensibles (como lo son sus datos médicos) pero desconoce a cuál de las dos entidades confiar dicha información, ya que no tiene datos objetivos que le permitan comparar ambas opciones.

Escenario 3 Un usuario necesita realizar una transacción económica (un traspaso de efectivo) y para ello tiene varios servicios disponibles que ofrecen la misma funcionalidad. El usuario utiliza Google como servicio que provee la identidad y en su cuenta tiene registrados datos sensibles (como su tarjeta de crédito). Sin embargo, el usuario no sabe cuál de los servicios posibles es más respetuoso con la privacidad de sus datos, no sabe si alguno emplea patrones de diseño engañosos para los usuarios, desconoce su reputación más allá de algunas opiniones que ha encontrado buscando en Internet. Le gustaría saber cuál escoger y si, antes de iniciar sesión en el elegido con Google, debería modificar alguna configuración de su proveedor de identidades o no.

5.2. Modelo a alto nivel del sistema de recomendación

El modelo propuesto en esta tesis doctoral se basa en añadir un sistema de recomendación dedicado a mejorar la protección de la privacidad en los sistemas de autenticación federados. El principal propósito del *Privacy Advisor* (PAdv) es mostrar a los usuarios cómo se ve afectada su privacidad cuando utilizan un IdP o servicio (RP) en el ámbito de una federación de identidades. El PAdv puede realizar recomendaciones para ayudar a los usuarios a

decidir si utilizar un IdP o servicio o no y, además, si deberían utilizarlo con restricciones o únicamente tras modificar alguna configuración en el IdP. Estas recomendaciones se generan teniendo en cuenta a cada usuario específico (características, atributos y preferencias de privacidad) y el procesamiento de los datos que hace cada IdP y RP (datos relativos a los usuarios recogidos por éstos y tratamiento que se hace de ellos).

El PAdv se basa en mecanismos de tipo *content-based filtering*, teniendo en cuenta un conjunto de características de usuario, IdP y RP discretas y etiquetadas para proporcionar una recomendación [118], [119].

5.3. Capacidades del modelo

El recomendador propuesto debe incorporar las siguientes capacidades:

1. Recolección de características, atributos y preferencias de usuario relativas a la protección de los datos y a su privacidad.

2. Recolección automatizada acerca de cómo un recurso, aplicación o proveedor de servicios maneja los datos de los usuarios. El PAdv puede utilizar, para ello, diferentes fuentes de datos abiertas.

3. Recomendación de privacidad multicapa, que muestre toda la información relevante obtenida de forma sencilla e ilustrativa para los usuarios, haciendo fáciles de interpretar las recomendaciones proporcionadas de manera que se consiga un impacto tangible en las decisiones de los usuarios.

4. Granularidad o niveles de detalle diferenciados para usuarios con diferentes niveles de conocimiento y concienciación o necesidades.

5.4. Arquitectura del *Privacy Advisor*

Para el *Privacy Advisor* se propone una arquitectura totalmente modular (descrita en la figura 5.1) que provea de las capacidades mencionadas anteriormente. Cada módulo tiene su propia funcionalidad y se puede ejecutar de forma distribuida, incluso diferentes módulos podrían pertenecer a diferentes organizaciones o compañías. La integración de nuevos módulos es también muy sencilla, ya que están diseñados para evitar al máximo dependencias entre ellos.

La implementación del PAdv no está ligada a un diseño concreto, se puede realizar, por ejemplo, como un *plugin*, servicio o aplicación.

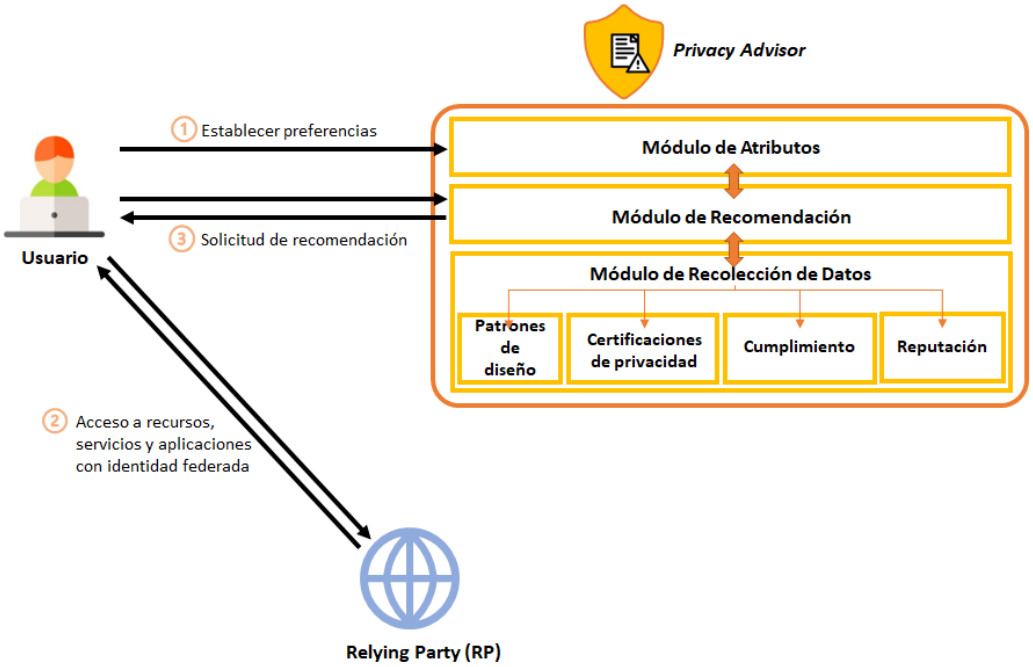

Figura 5.1: Arquitectura del *Privacy Advisor*

Como se muestra en la figura 5.1, hay un Módulo de Atributos responsable de obtener y almacenar las características, atributos y preferencias de los usuarios en relación con su privacidad (paso 1). Esta información ayuda al *Privacy Advisor* a adaptar la recomendación a las necesidades de los usuarios, de los dispositivos que usan, etc. Es decir, a que las recomendaciones estén realmente personalizadas. Cuando un usuario interacciona con una RP (paso 2), el Módulo de Recomendación recibe una petición y éste pide al Módulo de Recolección de Datos toda aquella información necesaria para realizar la recomendación. El Módulo de Recolección de Datos trabaja recogiendo información sobre la RP a la que el usuario desea acceder desde diferentes fuentes y empleando para ello diferentes mecanismos. Una vez que consigue toda la información necesaria la devuelve al Módulo de Recomendación para que produzca su salida y le devuelva una respuesta al usuario (paso 3).

Adicionalmente, el usuario podrá acceder al Módulo de Recomendación del PAdv para solicitar recomendaciones sobre un IdP.

Las recomendaciones se presentan en la primera capa con un código de colores tipo semáforo: verde (OK, buena), amarillo (parcialmente OK, advertencia) y rojo (NO OK, mala). Se emplea la información obtenida del Módulo de Recolección de Datos y del Módulo de Atributos para obtener la recomendación. Adicionalmente, hay una categoría "Desconocida" de color

gris cuando es imposible dar una recomendación personalizada o fundada debido a la falta de información del IdP o RP, de las preferencias del usuario o de ambas.

5.4.1. Integración con la federación de identidad

Añadir el *Privacy Advisor* como nuevo agente no afecta a los flujos de identificación, autenticación y autorización definidos en las especificaciones federadas. Este punto es esencial, ya que facilita su incorporación en entornos productivos en los que no se quiere modificar dichas especificaciones o en productos que las implementan.

Cuando un usuario quiere utilizar una RP y necesita autenticarse utilizando un esquema federado, en el modelo propuesto, se lanzan dos flujos diferentes. El primero se corresponde con el flujo habitual de autenticación. El segundo solicita al PAdv una recomendación. Para ello, un *plugin* en el navegador, una llamada a un servicio externo o una aplicación instalada realizan una petición de recomendación sobre la RP al PAdv (paso A de la figura 5.2). El PAdv responde con una recomendación sobre esta RP diseñada especialmente para el usuario teniendo en cuenta los atributos del usuario (paso B de la figura 5.2). Este flujo se puede realizar antes o durante el proceso de autenticación (resto de pasos en la figura 5.2). En el ejemplo de la figura 5.2 se utiliza el flujo *Authorization Code* de la especificación OpenID Connect pero podría utilizarse para cualquier flujo de OpenID Connect o en especificaciones federadas similares (SAML, OAuth, etc.).

El PAdv puede operar en dos modos diferentes dependiendo de las preferencias del usuario. La primera lanza el proceso de autenticación una vez el usuario ha revisado y aceptado la recomendación de privacidad. Este modo de proceder es más intrusivo, pero garantiza que el usuario conozca cómo se ve afectada su privacidad cuando utiliza un servicio. El segundo ejecuta en paralelo al proceso de autenticación (por ejemplo, en una nueva pestaña o ventana en el navegador). Este proceso es completamente independiente, no obliga al usuario a comprobar la recomendación antes de utilizar el recurso, aplicación o servicio. Trabajar con un modo u otro se puede configurar en el Módulo de Atributos del PAdv.

5.4.2. Diseño de los módulos del *Privacy Advisor*

Módulo de Atributos

Tal y como se ha comentado anteriormente, este módulo es parte esencial en el modelo propuesto, ya que con las características, atributos y preferencias

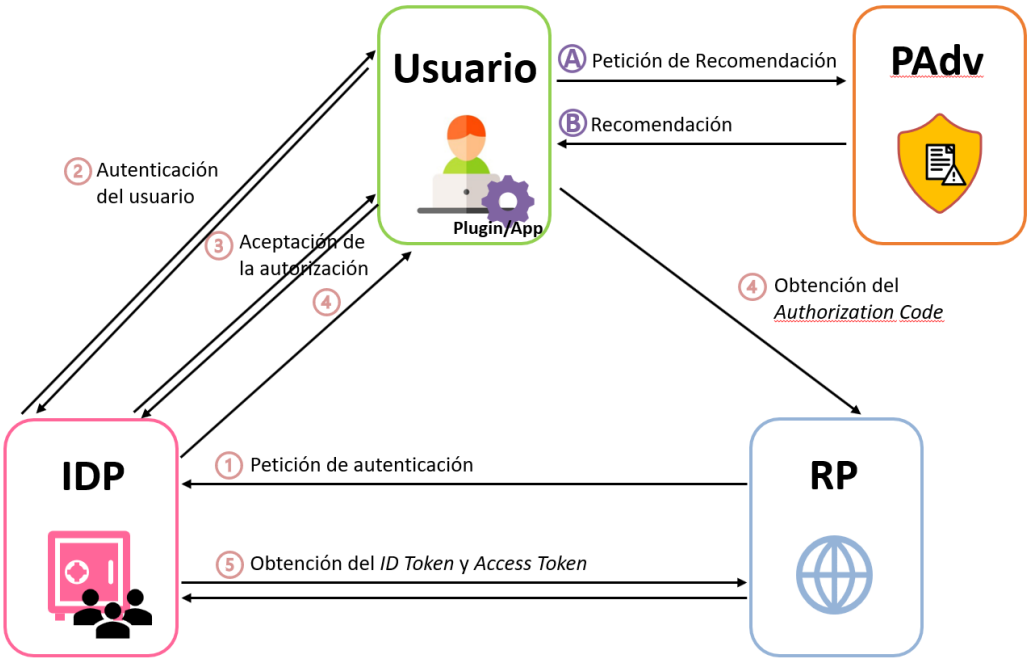

Figura 5.2: Ejemplo de flujo de autenticación con PAdv y OpenID Connect

del usuario se puede personalizar la respuesta para este usuario, creando una actitud positiva con respecto a la recomendación y aumentando su intención de uso porque se confía más en las recomendaciones obtenidas del PAdv.

La configuración de las características, preferencias y atributos es un proceso guiado que lleva entre 20-30 minutos la primera vez que se realiza y que se puede actualizar tantas veces el usuario necesite. Este proceso permite al PAdv conocer mejor al usuario preguntando explícitamente:

- Qué proveedores de identidad utiliza el usuario (lista cerrada).
- Los datos que comparte con cada uno de ellos, en concreto los datos personales, y tipo de cuenta.
- Dispositivos desde los que los utiliza.
- Categorías de los servicios (RP) que utiliza y los más visitados o utilizados. Se propone utilizar un algoritmo de clasificación de servicios usando las categorías de SimilarWeb [120]: adultos, arte y entretenimiento, informática y tecnología, noticias y medios de comunicación, etc.

Tabla 5.1: Nivel de exposición

Puntuación	Explicación
1 - Improbable	Utilizar IdP sin vulnerabilidades conocidas (no resueltas) y sin compartir información sensible con los IdP: nombre, apellidos, fotografía, códigos o números identificativos, etc.
2 - Bajo	Usar un IdP sin vulnerabilidades conocidas (no resueltas) y compartiendo datos sensibles con los IdP: número de teléfono, número de tarjeta de crédito, número de cuenta bancaria, etc.
3 - Alto	Utilizar al menos un IdP con vulnerabilidades conocidas (no resueltas) y sin compartir información sensible con los IdP: nombre, apellidos, fotografía, códigos o números identificativos, etc.
4 - Muy alto	Utilizar al menos un IdP con vulnerabilidades conocidas (no resueltas) y compartiendo datos sensibles con los IdP: número de teléfono, número de tarjeta de crédito, número de cuenta bancaria, etc.

Todos estos atributos permiten al PAdv modelar el nivel de exposición de un usuario específico. Hay cuatro niveles definidos indicados en la tabla 5.1: improbable (1), bajo (2), alto (3) o muy alto (4).

Un usuario que, por ejemplo, utiliza un solo IdP sin vulnerabilidades conocidas, al que únicamente le ha proporcionado detalles de contacto, y que utiliza un portátil con el que accede a categorías poco sensibles de la RP, se le etiquetará con un nivel "improbable" de exposición. Por el contrario, un usuario que utiliza varios IdP, que tengan vulnerabilidades conocidas, donde tiene información de sus tarjetas de crédito y que de manera habitual utiliza dispositivos móviles para acceder a muchas RP que utilizan estos datos sensibles (por ejemplo, finanzas, comercio electrónico o salud), se le etiquetará con un nivel "muy alto" de exposición.

Además, el usuario muestra, a través de ejemplos, a qué servicios (RP) le gustaría dar distintos niveles de acceso a sus datos personales. Estos niveles permiten al PAdv modelar la concienciación de privacidad del usuario dependiendo de las elecciones. Se proponen cuatro niveles de nuevo: muy comprometido (1), comprometido (2), interesado (3) y sin conocimiento (4).

A partir de este punto, se etiqueta a los usuarios con una puntuación de riesgo inherente (exposición x concienciación), con valores que oscilan entre 1 (improbable x muy comprometido) a 16 (muy alto x sin conocimiento). Este riesgo no varía a no ser que el usuario modifique sus características y atributos, de ahí utilizar el término riesgo inherente. Se ha seleccionado esta aproximación, que se basa en un sistema de puntuación y una matriz de riesgos, debido a que muchos organismos de estandarización y *frameworks* de gestión de riesgos en la seguridad y privacidad han demostrado ya sus beneficios en el pasado. Incluyendo evaluaciones de impacto de privacidad y procesos de cuantificación de riesgos para la privacidad [121], [122], [123].

Los datos obtenidos de este proceso se agrupan en un panel de control y se pueden modificar navegando a través de diferentes pestañas o pantallas. Se utilizan también mapas de calor como mecanismo de visualización, ya que son concisos y fáciles de entender. Los mapas muestran la exposición en una dimensión y con la concienciación de privacidad del usuario en la otra. Además, sirven para comparar a unos usuarios con otros, mostrando el comportamiento más común y señalando mejoras potenciales en la puntación del riesgo inherente. Todos estos aspectos son muy útiles para ayudar a los usuarios a mejorar sus conocimientos en relación con la protección de datos, así como su sensibilización.

Por último, en el apartado de preferencias, los usuarios pueden configurar cómo quieren usar el PAdv (el modo de funcionamiento antes explicado dentro de los flujos de autenticación) y escoger opciones de seguridad como la autenticación multi-factor, etc.

Módulo de Recolección de Datos

Este módulo es el encargado de obtener la información acerca de cómo trabajan los IdP o las diferentes RP con los que interactúan los usuarios. Para ello se clasifican tal y como se ha explicado anteriormente, en las categorías Adultos, Arte y entretenimiento, Negocios y servicios al consumidor, Comunidad y sociedad, Ordenadores y tecnología, Comercio electrónico y compras, Finanzas, Comida y bebida, Juegos de azar, Juegos, Salud, Industria pesada e ingeniería, Hobbies y ocio, Hogar y jardín, Trabajo y carrera profesional, Derecho y gobierno, Estilo de vida, Noticias y medios de comunicación, Mascotas y animales, Material de consulta, Ciencia y educación, Deportes, Viajes y turismo y Vehículos. Se podrían utilizar otras formas de clasificación utilizando otras APIs como *Alexa top site* porque difieren muy ligeramente unas de otras.

En este trabajo se proponen cuatro submódulos para la recolección de datos acerca de un IdP o RP: Revisión de patrones de diseño, Certificaciones

de privacidad, Cumplimiento y Reputación.

Todos estos submódulos se han diseñado de forma independiente para poder ser modificados, reemplazados, ejecutar de forma distribuida o que se añadan nuevos submódulos sin afectar el modelo propuesto. Cada submódulo recibe una entrada de datos diferente cuando se les consulta por un IdP o RP, producen un resultado y devuelven la información al Módulo de Recolección de Datos. Se pueden utilizar diferentes combinaciones de submódulos dependiendo de la categoría del IdP o RP por el que se pregunta y la información que tiene disponible.

Revisión de patrones de diseño Los patrones oscuros (*dark patterns*), también conocidos como *deceptive designs*, [124] [125] se basan en el uso de textos ambiguos, opciones premarcadas, distracciones con *pop-ups*, colores, y otros atributos de la interfaz, ocultando opciones o información relevante que las organizaciones no quieren que el usuario localice rápidamente (por ejemplo, la cancelación de una suscripción). Por tanto, estos patrones están diseñados para engañar al usuario y que realice acciones que no son necesariamente buenas o positivas para él. Los aspectos psicológicos relacionados con el uso de este tipo de patrones se discuten en profundidad en [126].

Este submódulo del PAdv tiene que identificar qué patrones oscuros utiliza un IdP o RP en relación con la protección de datos. Con esta información disponible en el *Privacy Advisor*, los usuarios pueden estar más atentos a sus interacciones con este sistema en concreto. Además, pueden evitar el propósito del patrón oscuro o incluso decidir utilizar otro IdP o RP que sea más respetuosa con la privacidad.

El submódulo propuesto debería ser capaz de identificar las siguientes categorías de patrones oscuros siguiendo la clasificación proporcionada en [124]:

- *Tricky Questions*: Se realizan preguntas confusas a los usuarios (por ejemplo, con dobles o triples negaciones) para que den la respuesta que ellos podrían no querer dar.

- *Sneak into Basket*: Normalmente aparece con opciones premarcadas (*checkboxes* o *radio buttons*). Por ejemplo, mediante la confirmación de una configuración de privacidad, otras configuraciones adicionales se confirman también.

- *Roach Motel*: La suscripción en un servicio es fácil para el usuario pero la cancelación del servicio es complicada para el usuario.

- *Privacy Zuckering*: Se confunde al usuario para que comparta más información de la que quiere. Por ejemplo, los términos de uso ocultan una cláusula donde se indica que toda la información se comparte con el *data broker*.

- *Misdirection*: Mediante un diseño específico en el servicio se intenta confundir a los usuarios para que pasen por alto información relevante.

- *Bait and Switch*: El usuario decide algo pensando que la decisión es diferente.

- *Confirmshaming*: El servicio trata de evitar que ciertos usuarios realicen alguna acción, por ejemplo, una cancelación de una suscripción, apelando al sentimiento de culpa.

- *Disguised Ads*: Hay anuncios ocultos entre el contenido de servicio para que los usuarios pulsen más fácilmente o de forma no intencionada.

- *Forced Continuity*: Intenta mantener la suscripción de un servicio después de un periodo de prueba sin la confirmación del usuario.

- *Friend Spam*: Cuando un usuario da el consentimiento a un servicio para que acceda a su lista de amigos y el servicio la utiliza para enviarles correo *spam*.

Certificaciones de privacidad Los sellos de privacidad y las certificaciones permiten a los proveedores certificar el cumplimiento de unos requisitos de privacidad y protección de datos. Estos sellos y certificaciones normalmente implican auditorías externas al principio y, también, periódicamente durante el proceso de renovación.

El *Privacy Advisor* recolecta esta información para los usuarios, ya que un tercero certifica unos niveles de protección de un IdP o RP en particular. La principal funcionalidad de este submódulo es recolectar estos sellos y certificaciones de privacidad, listarlos e informar al usuario sobre las implicaciones que tiene tener uno de estos activo.

Se propone que se comprueben, al menos, las siguientes certificaciones:

- *European Privacy Seal* (EuroPriSe) [127] certifica el cumplimiento de la regulación europea de protección de datos, centrándose en los datos de navegación públicos de los usuarios. Esta certificación tiene definidos dominios de privacidad, donde los auditores pueden comprobar *cookies*, *hosting* web, transferencias internacionales de datos, etc.

- *TrustArc company* [128] tiene nueve certificaciones centradas en diferentes escenarios de privacidad. Estas certificaciones requieren del uso del *framework TrustArc Privacy & Data Governance* (P&DG). Este *framework* cumple con las guías de privacidad OECD (*Organisation for Economic Co-operation and Development*), la regulación de protección de datos (RGPD), la ISO27001 e HIPAA (*Health Insurance Portability and Accountability*).

    - *APEC Cross Border Privacy Rules* (CBPR): Certifica el libre paso de datos personales entre los participantes de este programa y América o Asia. También comprueba que el nivel de protección de datos personales es apropiado en términos de seguridad y privacidad.
    - *Enterprise Privacy & Data Governance Practices*: Con esta certificación, el organismo cumple con los requirimientos incluidos en el *framework TrustArc Privacy & Data Governance*.
    - *Data Collection*: Certifica que el proceso de recolección de datos personales es respetuoso con la privacidad cuando actúan como tercera parte en una web o aplicación móvil.

- *ePrivacy* [129] se centra en el RGPD europeo y la oficina de publicidad interactiva (*Interactive Advertising Bureau Europe for Online Behavioural Advertising Framework*)

    - *ePrivacySeal*: Se centra en la revisión del cumplimiento del RGPD (ePrivacyEU) incluyendo las particularidades de la ley de protección de datos de Suiza (ePrivacySeal CH) y Alemania (ePrivacySeal DE).
    - *ePrivayApp*: Es equivalente al *PrivacySeal* pero en el escenario de aplicaciones móviles.

Estas certificaciones se han seleccionado porque están relacionadas con la privacidad en la Unión Europea y en sus relaciones con terceros. EuroPriSe es un proyecto fundado por la Unión Europea (UE), ePrivacy está reconocido como una autoridad de control experta para la protección del dato [130] y TrustArc existe desde hace más de veinte años con más de mil clientes [131] y premios en ciberseguridad [132].

Cumplimiento Las especificaciones OAuth 2.0 [13], OpenID Connect [14] y *JSON Web Token* (JWT) [133], utilizadas en los esquemas federados, incluyen apartados específicos dedicados a resolver aspectos de seguridad y

privacidad. Estos apartados proponen las buenas prácticas relativas a la mitigación de amenazas y vulnerabilidades y que también se pueden encontrar en otros trabajos como [134] o [19].

Las regulaciones RGPD europea, eIDAS (*electronic IDentification, Authentication and trust Services*) y PSD2 (*Payments Services Derivative 2*) también incluyen normas que los proveedores tienen que cumplir. eIDAS se centra en infraestructuras de clave pública (PKI, *Public Key Infrastructures*), gestión del ciclo de vida de los certificados y firmas electrónicas. Esta regulación aplica a todos los miembros de la UE, permitiendo marco legal común para los certificados y firmas electrónicas. De esta forma, los certificados y firmas electrónicas son válidos en todos los estados miembros de la UE. PSD2 es una regulación europea para los servicios de pago electrónico. Esta regulación se centra en la seguridad en los pagos electrónicos y facilita los pagos en la Eurozona. También permite a terceras entidades gestionar las finanzas de los consumidores. El RGPD regula la protección de la privacidad y seguridad de los datos personales de las personas de la Unión Europea. Esta regulación define una serie de roles para asegurar su cumplimiento.

La principal funcionalidad de este submódulo es comprobar el cumplimiento de las buenas prácticas en seguridad y privacidad que se incluyen en las especificaciones subyacentes, así como el cumplimiento con las regulaciones mencionadas. Cuando el Módulo de Recolección de Datos pregunta sobre un IdP o RP, este submódulo contesta con información específica sobre cumplimientos e incumplimientos.

Reputación Este submódulo del *Privacy Advisor* es opcional y cuantifica el nivel de confianza que se tiene en un IdP o RP utilizando agentes independientes (otros proveedores, usuarios, recomendadores de privacidad, etc.). Por tanto, este componente recurre a evaluadores externos que tienen información sobre la reputación del IdP o RP: facilidad para ejercer los derechos, quejas y reclamaciones, brechas de datos previas, opiniones y comentarios. Cada evaluador externo basará su cuantificación de reputación en aspectos diferentes.

Este submódulo solicita una recomendación de reputación de un servicio a una RP, IdP ó PAdv de confianza. Estas entidades responden, en un formato procesable por máquinas, con una puntuación de recomendación entre cero y cien (V_i) y cualquier otra información adicional que justifique esa evaluación.

El *Privacy Advisor* pueden no confiar de igual forma la evaluación de cada una de las entidades a las que se ha preguntado. Se propone un sistema de pesos donde el peso P_i toma valores entre cero y uno por cada evaluador consultado. Cuanto más mayor sea el peso, más confiable será para el PAdv.

Esta confianza se puede configurar en las preferencias del usuario para darle la oportunidad de representar su confianza en cada entidad integrada en el submódulo.

La puntuación final de la reputación se obtiene aplicando la siguiente expresión:

$$Reputación = \sum_{i=1}^{n} V_i \cdot P_i \tag{5.1}$$

El diseño del sistema de reputación y la decisión sobre los pesos que se deben asignar a cada agente evaluador no forma parte del objetivo de esta tesis. Aunque hay que remarcar, que incluir este tipo de sistemas de reputación es crucial, ya que se utilizan en contextos complejos donde hay multitud de RP como en computación en nube [135], [61] o IoT [136], [137] y están ganando una mayor fuerza [138], incluso recurriendo a uso en entornos descentralizados basados en *blockchain* [139].

Módulo de Recomendación

Como ya se ha mencionado, los usuarios tienen distinto nivel de conocimiento y sensibilización, percepciones del riesgo y necesidades. El PAdv debe ser capaz de generar recomendaciones útiles para todos ellos. Para ello, tiene la capacidad de recoger la información del resto de módulos para generar una recomendación y permite la solicitud de recomendaciones sobre IdP.

La diversidad y heterogeneidad de todos los usuarios potenciales en el dominio considerado hacen que la meta del modelo propuesto y la decisión informada de los usuarios sea un reto. Por lo tanto, la siguiente lista de características es primordial para que la recomendación tenga los efectos deseados:

- Personalización: La recomendación proporcionada debe ser individualizada para ser de ayuda, informando a los usuarios sobre las prácticas de protección de datos del IdP o RP y el grado de cumplimiento de las preferencias personales.

- Complejidad reducida: La recomendación proporcionada tiene que tener valor y ser directa para evitar la fatiga para obtener los resultados esperados.

- Tener en cuenta la diversidad de usuarios: Enseñar todos los detalles de la recomendación a la vez rara vez es práctico. Sin embargo, los usuarios más expertos pueden requerir una explicación de la recomendación proporcionada. La recomendación, por tanto, tiene que darse en capas,

desde la recomendación más directa hasta los diferentes niveles de detalle provenientes del Módulo de Recolección de Datos y el Módulo de Atributos.

- Ofrecer elecciones significativas: La recomendación proporcionada debe ser procesable y útil, permitiendo a los usuarios tomar decisiones informadas sobre privacidad.

Las primeras tres características se cumplen por cómo está el sistema de recomendación diseñado, considerando las características, atributos y preferencias almacenadas en el Módulo de Atributos y empleando métodos visuales multicapa que combinan texto, imágenes e iconos. Las recomendaciones multicapa están compuestas por un conjunto de detalles complementarios que se ajustan a diferentes públicos. Están diseñados con especial cuidado en términos de presentación y tiene la capacidad de aumentar la información proporcionada gradualmente de forma que se mejore la atención de los usuarios y la comprensión de la recomendación. Los colores también son algo esencial. El PAdv utiliza el código de colores de tipo semáforo (verde, amarillo y rojo) y utiliza el gris para aquellas categorías o atributos no conocidos.

La cuarta característica de la lista, ofrecer elecciones significativas, se asegura mediante la integración del PAdv en los flujos de identificación, autenticación y autorización (tal y como se ha explicado en la sección 5.4.1). El usuario puede realizar decisiones de peso en base a la recomendación recibida, pudiendo limitar o abortar la interacción con el IdP o RP, o incluso cambiar la configuración de la cuenta en el IdP.

Primera capa de recomendación Como se ha indicado anteriormente, en la primera capa de la recomendación se utiliza el código de color tipo semáforo. El verde significa OK (Buena). Por tanto, el usuario puede continuar con la interacción con el IdP o RP con garantías, considerando sus atributos y la información disponible del IdP o RP. Amarillo significa parcialmente OK, además de una recomendación de "Advertencia". Los usuarios deberán limitar las interacciones con el IdP o RP (únicamente si no hay alternativas o si es estrictamente necesario) o modificar la configuración de la cuenta en el IdP para mejorar el control de los datos personales compartidos. Finalmente, el color rojo indica NO OK (Mala) y con la recomendación de abortar toda interacción con el IdP o RP. El color gris es para resultados desconocidos o no concluyentes, lo que significa que no es posible dar una recomendación personalizada o bien argumentada debido a la falta de información sobre el usuario, el IdP o la RP.

Tal y como se ha mencionado, se utiliza un mecanismo de tipo *content-based filtering* para obtener la recomendación, en concreto, uno basado en

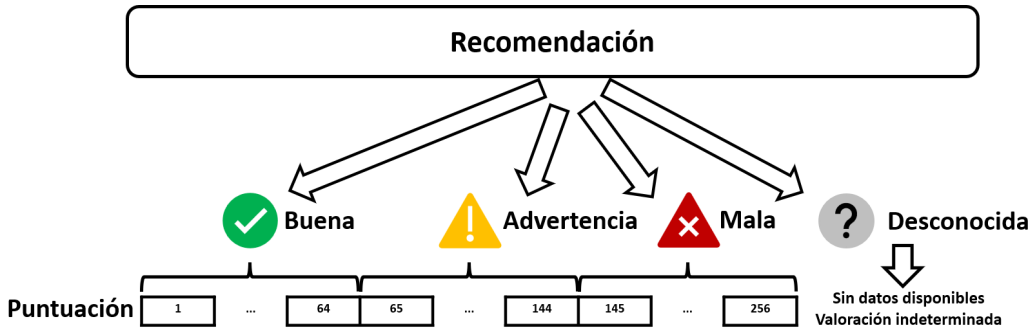

Figura 5.3: Ejemplo de un árbol de decisión para una RP en la categoría Comercio electrónico y Compras

árboles de decisión. Se pueden tener diferentes árboles de decisión para generar la recomendación dependiendo de la categoría del IdP o RP, ya que la recomendación puede variar dependiendo de si el IdP o RP pertenece a la categoría Noticias, Finanzas o Salud, por ejemplo.

La figura 5.3 muestra un ejemplo del árbol de decisión utilizado para la categoría de Comercio electrónico y Compras. La recomendación puede ser "Desconocida" si no hay suficientes datos disponibles o "Buena", "Advertencia" o "Mala", dependiendo de la puntuación total de riesgo que se utiliza en el árbol de decisión (con valores desde 1 a 256). Cuando se cambia de una categoría a otra, con una mayor o menor criticidad de los datos manejados, el único elemento que cambia en el árbol de decisión son las puntuaciones umbral para cada recomendación.

Los datos obtenidos del Módulo de Recolección de Datos se utilizan para obtener una puntuación de riesgo de transacción para un determinado IdP o RP. En este caso, el riesgo está asociado, no con el usuario como era el caso del riesgo inherente, sino con el uso de un IdP o RP en concreto en un momento específico. Es por estos motivos por lo que se le ha denominado riesgo de transacción.

De nuevo se utilizan mapas de calor, en este caso considerando las características del IdP o RP para el que se solicita la recomendación. Los atributos relativos a las certificaciones y cumplimiento se representan en una dimensión (tabla 5.2, 1: Excelente, 2: Buena, 3: Suficiente y 4: Pobre). Se combinan en una única dimensión, ya que representan el mismo objetivo, la conformidad con la ley, regulación, especificación técnica, certificación asociada a buenas prácticas, etc. Los atributos relativos a los patrones de diseño se representan en la otra dimensión (tabla 5.3, 1: Excelente, 2: Buena, 3: Suficiente y 4: Pobre). Los resultados obtenidos del submódulo de Reputación (si están

Tabla 5.2: Certificaciones y cumplimiento

Puntuación	Explicación
1 - Excelente	Al menos una certificación de privacidad y al menos un 90 % de cumplimiento de las buenas prácticas y regulaciones
2 - Buena	Al menos una certificación de privacidad y al menos un 60 % de cumplimiento de las buenas prácticas y regulaciones
3 - Suficiente	Sin certificaciones de privacidad y al menos un 40 % de cumplimiento de las buenas prácticas y regulaciones
4 - Pobre	Sin certificaciones de privacidad y menos de un 40 % de cumplimiento de las buenas prácticas y regulaciones

Tabla 5.3: Patrones de diseño

Puntuación	Explicación
1 - Excelente	Sin patrones oscuros identificados
2 - Buena	Se utiliza un patrón oscuro
3 - Suficiente	Se utilizan entre dos y cuatro patrones oscuros
4 - Pobre	Se utilizan más de cuatro patrones oscuros

disponibles) se utilizan para ajustar la puntuación obtenida. Por ejemplo, una RP con una buena puntuación en certificaciones y cumplimiento, que no tiene patrones de diseño oscuros identificados, 2 (Buena) X 1 (Excelente) = 2, podría acabar con un riesgo de transacción más alto si la reputación de la RP es terrible porque ha sufrido muchas fugas de información en los últimos meses o porque se ha demostrado que hace perfilado de sus usuarios de forma agresiva.

En particular, la reputación se utiliza para ajustar la puntuación en el eje de certificación y cumplimiento. Es decir, se matiza la evaluación estática de las buenas prácticas en cumplimiento con información sobre lo que otros usuarios o agentes independientes han observado en casos de uso reales. La tabla 5.4 muestra la propuesta de calibración. En ningún caso la puntuación puede ser menor a 1 o superior a 4. Si hay un IdP o RP que ya tiene la puntuación más baja o más alta posible, la puntuación de riesgo inicial no se modificará hacia un valor menor o mayor, respectivamente, ya que no son valores posibles en la matriz.

Tabla 5.4: Calibración de la puntuación de certificaciones y cumplimiento basado en la reputación del IdP o RP implicado

Calibración	Explicación
0	La puntuación de riesgo inicial no se modifica porque no hay información suficiente sobre su reputación o porque la reputación está entre 40 y 60 (valores intermedios o en la media)
+1	La puntuación de riesgo original se desplaza un nivel arriba debido ya que tiene una reputación por debajo de 40 (el IdP o RP no tiene buena reputación)
-1	La puntuación de riesgo original se desplaza un nivel abajo debido ya que tiene una reputación por encima de 60 (el IdP o RP tiene buena reputación)

El riesgo total, que se utiliza en los árboles de decisión del recomendador se evalúa como $R = Riesgo\,inherente \times Riesgo\,de\,transacción$, con valores entre 1 (1x1) hasta 256 (16x16).

Segunda capa de recomendación La segunda capa muestra la puntuación del riesgo inherente y de transacción para el usuario y el IdP o RP, respectivamente, así como el árbol de decisión empleado en función de la categoría del IdP o RP y los mapas de calor (figura 5.4) para que el usuario pueda entender como se ha generado la recomendación. Si la recomendación es "Advertencia", en esta segunda capa también se muestran las acciones recomendadas para el usuario: como restringir la interacción con el IdP o RP o como cambiar la configuración de la cuenta en el IdP. Estos cambios puede variar las características, atributos o preferencias de los usuarios (su exposición, por ejemplo, compartiendo menos datos personales con el IdP) y, por tanto, las recomendaciones en interacciones futuras.

Tercera capa de recomendación La tercera capa muestra todos los datos recogidos por el Módulo de Recolección de Datos y el origen de dichos datos. En la figura 5.5 se muestra un ejemplo de la información mostrada, indicando las tres capas de recomendación propuestas.

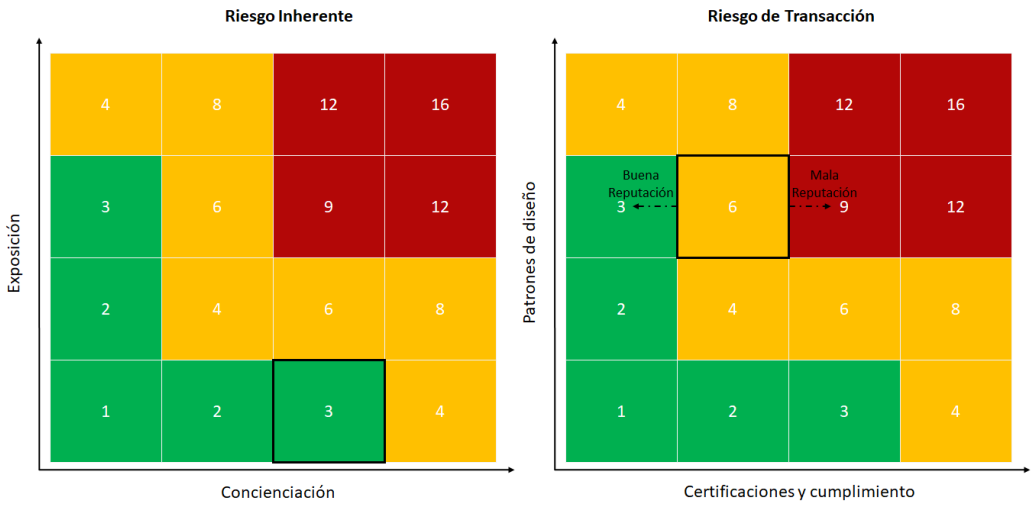

Figura 5.4: Ejemplo de mapas de calor para el riesgo inherente y de transacción

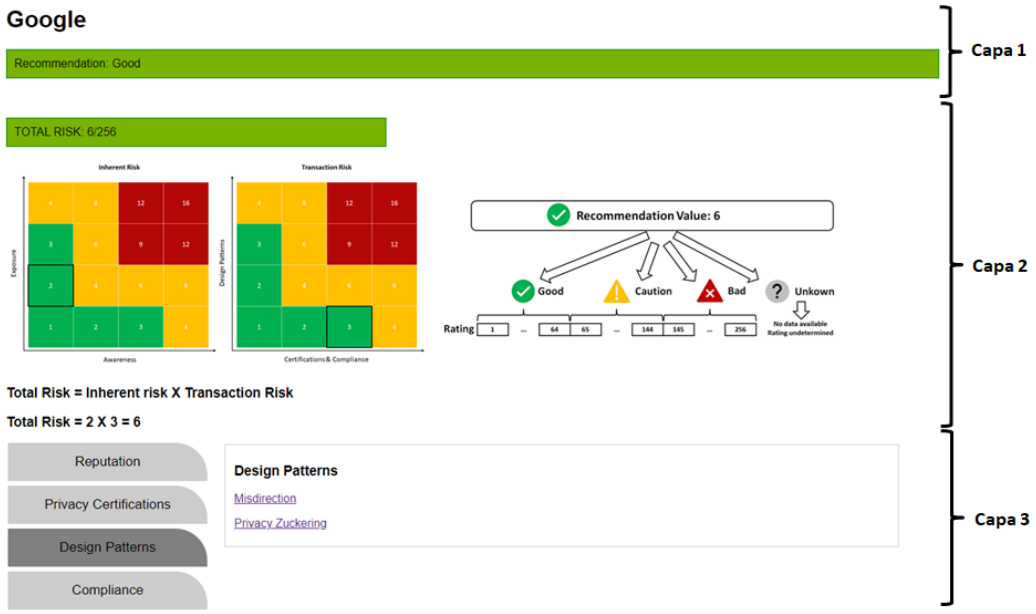

Figura 5.5: Ejemplo de una recomendación del PAdv

5.5. Implementación del prototipo

Se ha desarrollado un nuevo *plugin* para Google Chrome que abre una nueva pestaña cuando el usuario pulsa en un *link* de inicio de sesión en un IdP. Esta pestaña muestra la recomendación de privacidad para el servicio que el usuario quiere utilizar. La implementación del PAdv se ha realizado utilizando Python. El Módulo de Recomendación recibe una petición HTTPS (*Hypertext Transfer Protocol Secure*) de un usuario final y, una vez que ha consultado a los otros módulos, compone una página web HTML de respuesta con toda la información sobre el servicio por el que se le está preguntando siguiendo la propuesta de multicapa explicada anteriormente. También se ha provisto de una pequeña interfaz gráfica en el Módulo de Recomendación para solicitar este tipo de recomendaciones sobre los IdP.

El proceso de autenticación se ejecuta de forma independiente y se ha certificado que no se ve afectado por esta nueva entidad, el *Privacy Advisor*.

Además, se almacenan en una caché las peticiones de los usuarios para reducir el tiempo de respuesta medio. La respuesta de cada submódulo del Módulo de Recolección de Datos se almacena asociándola al IdP o RP. Cuando un usuario solicita una recomendación, el *Privacy Advisor* consulta si esta información ya ha sido obtenida recientemente. Si lo ha sido, el PAdv utiliza la información almacenada en la caché. La caché tiene un tiempo de expiración de 10 minutos para evitar el almacenamiento masivo y el consumo de recursos innecesarios así como la obsolescencia de los datos almacenados.

5.5.1. Implementación de los submódulos del Módulo de Recolección de Datos

Se debe aclarar que ciertos aspectos de bajo nivel relativos a la implementación de los submódulos están fuera del alcance de esta tesis, como por ejemplo el empleo de las técnicas de *web crawling* o *web scraping* para detectar el uso de patrones oscuros en un IdP o RP. El modelo propuesto se ha validado utilizando técnicas, servicios y herramientas disponibles actualmente. Se podrían mejorar y proponer innovaciones si se tratara de una puesta en producción, pero son suficientes para la validación de la propuesta realizada y de un primer prototipo.

El submódulo de Revisión de patrones de diseño se ha implementado utilizando el *hall of shame* propuesto en [124] como punto de partida. En el primer prototipo del PAdv, se categorizaron manualmente un total de cincuenta patrones oscuros de este *hall of shame*, extrayendo cada tipo de patrón oscuro, la página web donde aparecía y un enlace al *tweet* con la descripción del patrón oscuro. Cuando el Módulo de Recomendación pregunta por los

patrones oscuros identificados para un IdP o RP, este submódulo, utilizando Python, busca en su base de conocimiento todos los patrones oscuros y los devuelve en formato JSON, si hubiera. Esta base de conocimiento se podría actualizar con nuevos sitios web o categorías de patrones oscuros en el futuro.

En relación al submódulo de Certificaciones de privacidad, toda la información sobre los IdP o RP certificados se obtiene directamente de las páginas web de los certificadores. En este sentido, la información sobre las certificaciones actuales es completa y está actualizada. Las certificaciones que se han considerado en este prototipo son las previamente explicadas: EuroPrise [127], TrustArc [128] y ePrivacy [129]. La obtención de las certificaciones del IdP o RP se han realizado utilizando Python para procesar las páginas web que contienen estas certificaciones o para el uso de servicios mediante los cuales también se puede obtener la certificación. Se ha utilizado la librería BeautifulSoap [140] para parsear el HTML de las páginas web de EuroPriSe y ePrivacy. Para la integración de TrustArc se ha utilizado un PHP que devuelve un JSON con las certificaciones asociadas a una compañía.

El submódulo de Cumplimiento se ha validado comprobando algunas de las buenas prácticas de OAuth 2.0 [13] y de OpenID Connect [14] así como el uso de certificados eIDAS [141].

La primera verificación consiste en comprobar que el código de autorización se envíe sobre un canal seguro. Se ha comprobado si la URI de redirección utiliza HTTPS. La segunda verificación implementada es la protección contra CSRF (*Cross-Site Request Forgery*, del inglés). La RP tiene que utilizar una protección anti-CSRF en la URI de redirección. La RFC de OAuth 2.0 [13] recomienda la utilización del parámetro *state* con un valor no adivinable en la petición de autorización. Se ha comprobado si el parámetro está incluido y cumple dicho requisito.

Además, para los requisitos de TLS de la especificación de OpenID Connect [14], también se ha comprobado que la negociación del cifrado es segura. Esta información se obtiene a través de la herramienta pysslscan [142]. La implementación se considera insegura si utiliza SSL (*Secure Sockets Layer*) 2.0 o SSL 3.0. El uso de TLS 1.0 y TLS 1.1 se considera como cifrado mejorable con algunas vulnerabilidades y el uso de TLS 1.2 o superior se considera suficientemente seguro.

Adicionalmente, tal y como se ha explicado anteriormente, se comprueba que una entidad de certificación de confianza (como las registradas en eIDAS) es la emisora de los certificados que se utilizan. Se ha utilizado una API [143] que comprueba la entidad de certificación, de manera que se obtiene el *Common Name* (CN) del certificado emisor y se muestra información detallada del emisor.

El submódulo de Reputación se ha desarrollado preguntando a una RP y

a un IdP de confianza, suponiendo que son capaces de ofrecer evaluaciones de este tipo. Estos dos elementos son capaces de recibir peticiones de reputación y de responder con su evaluación. Se han registrado como ejemplo algunas reputaciones de algunos IdP o RP dando valores entre 0 y 100, simplemente para poder realizar una validación funcional del submódulo.

5.6. Validación y discusión

Se han realizado diferentes experimentos con una implementación centralizada en Python 3 del prototipo descrito en la sección anterior, ejecutando en un ordenador Windows 10 con procesador Intel Core i5-4670K y 8GB de RAM.

5.6.1. Análisis del rendimiento

Cada uno de los experimentos ha consistido en consultar al *Privacy Advisor* de manera que tenga que producir una recomendación completamente nueva. Los resultados del rendimiento se han obtenido ejecutando los experimentos cincuenta veces y obteniendo una media aritmética obteniendo los siguientes resultados:

- Latencia: El tiempo medio de ejecución de una sola recomendación, sin el uso de la caché, es de 16,1 segundos (con una desviación estándar de 510 ms). De media, se han consumido 700 ms en la comunicación entre los distintos submódulos y módulos, y el resto de tiempo se ha consumido en los distintos módulos para la obtención de los resultados desde cero (algunos de ellos son procesos que consumen mucho tiempo, como el parseo HTML y el análisis). Teniendo en cuenta que los usuarios suelen utilizar los mismos servicios, la latencia se puede mejorar en gran medida utilizando una caché o incluso analizando algunos de los IdP y RP más utilizados por adelantado (por ejemplo, el conjunto de IdP y RP que utiliza el usuario normalmente y otros usuarios parecidos). El tiempo medio de obtención de una recomendación disminuye a 5.4 ms (con una desviación estándar de 0.15 ms) con esta mejora.

- Consumo de recursos: Durante la ejecución del *Privacy Advisor*, el uso de recursos en media es de 30MB de RAM y un 11 % de CPU. El uso más alto de RAM ha sido 38MB y el de CPU, un 13 %.

Se han detectado algunos problemas de escalabilidad con la implementación realizada en Python, ya que el prototipo no soportaría una gran cantidad

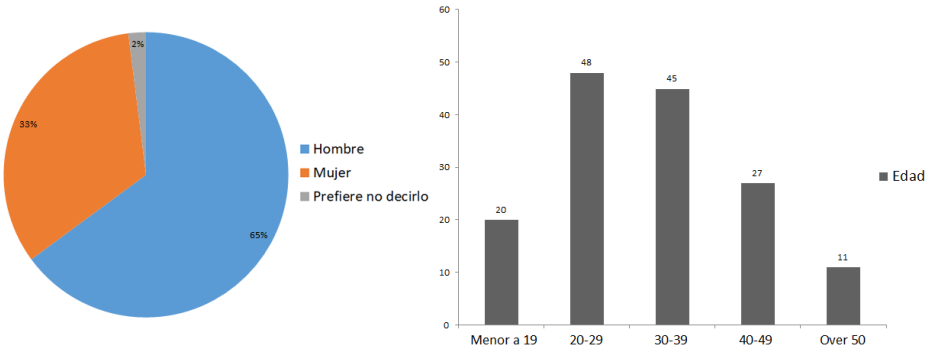

Figura 5.6: Género de los participantes (izquierda) y su edad (derecha)

de usuarios. Sin embargo, la propuesta de una arquitectura modular permitiría distribuir la ejecución del *Privacy Advisor* cuando fuera necesario en entornos productivos.

5.6.2. Análisis de efectividad y usabilidad

Para comprobar la aceptación potencial del *Privacy Advisor* se ha realizado una encuesta con 151 participantes. En este caso, el objetivo es evaluar sus beneficios, efectividad y usabilidad.

En relación a los datos demográficos de los participantes, el 65 % de los participantes eran varones y el 33 % eran mujeres (parte izquierda de la figura 5.6). Todos ellos de España. También en la figura 5.7 se muestra el sector laboral de los participantes agrupados usando *The Global Industry Classification Standard* [144]. La categoría otros incluye otros estados de empleo como estudiantes o desempleados. Un gran número de ellos trabaja en el sector de las tecnologías de la información (figura 5.7), por lo que se espera que su grado de sensibilización con la privacidad sea ligeramente superior a la media en otros sectores. La parte derecha de la figura 5.6 muestra la edad de los participantes, todos ellos mayores de edad. Todos los participantes en esta evaluación de la solución propuesta han recibido información sobre la finalidad del experimento y han participado en él de manera completamente voluntaria, sin recibir ningún tipo de presión y proporcionando para ello su consentimiento explícito. A todos se les ha informado de los mínimos riesgos que corrían, ya que solo se recogían datos como su edad, género y sector de actividad, que no permiten identificarles.

Se ha preparado un cuestionario (tabla 5.5) que consulta sobre la calidad de la recomendación para 25 IdP y RP diferentes, la calidad de las decisiones que se han tomado gracias a la recomendación proporcionada y el esfuerzo

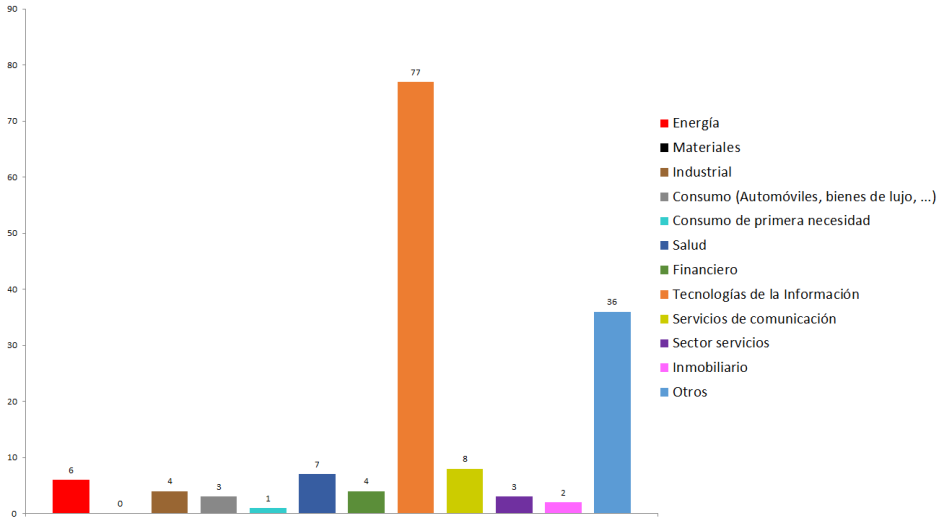

Figura 5.7: Sector laboral de los participantes

Tabla 5.5: Cuestionario y resultados medios

#	Pregunta	Métrica	Resultado medio
1	La recomendación encaja con tus necesidades y es personalizada	Calidad de la recomendación	4.12
2	Las recomendaciones son comprensibles y te permiten tomar decisiones sobre la protección de tu privacidad que no habrías tomado sin el PAdv	Calidad de la recomendación	4.28
2	Estás satisfecho con las decisiones finales que has tomado	Calidad en la decisión	4.05
3	El proceso de decisión ha sido sencillo	Esfuerzo en la decisión	3.84
4	El tiempo dedicado a la toma de la decisión ha sido razonable	Esfuerzo en la decisión	3.23

que ha supuesto tomar estas decisiones para el usuario. Los usuarios han respondido cada pregunta utilizando estrellas y el sistema *Likert scale*: Muy en desacuerdo/En desacuerdo/Ni en acuerdo ni en desacuerdo/De acuerdo/Muy de acuerdo [145].

Las preguntas que se muestran en la tabla 5.5 se refieren al PAdv como un todo y a las recomendaciones que proporciona. Se ha añadido al cuestionario una pregunta específica que valora los diferentes tipos de información proporcionada en la tercera capa del PAdv, permitiendo a los participantes evaluar la utilidad de la información proporcionada por los submódulos del Módulo de Recolección de Datos:

- Revisión de patrones de diseño ha obtenido un 4.31 sobre 5.
- Certificaciones de privacidad ha obtenido un 3.77 sobre 5.

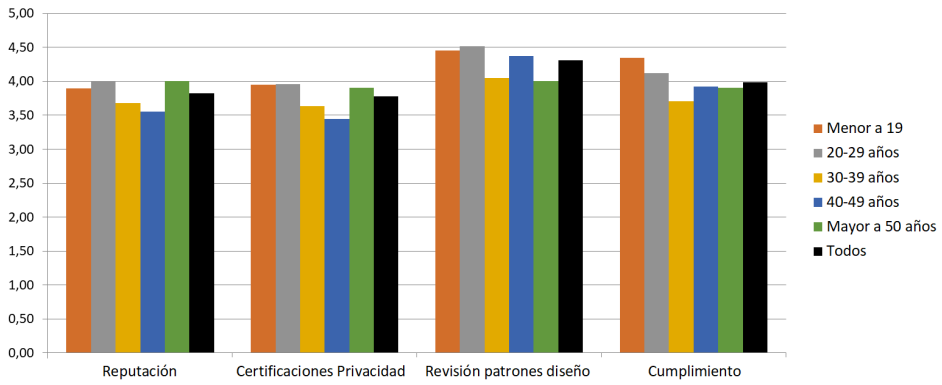

Figura 5.8: Evaluación de los submódulos del Módulo de Recolección de Datos

- Cumplimiento ha obtenido un 3.99 sobre 5.
- Reputación ha obtenido un 3.82 sobre 5.

La figura 5.8 muestra el resultado en detalle, se observa cómo todos los submódulos se perciben como útiles y valiosos independientemente de la edad de los participantes.

5.6.3. Discusión

Teniendo en cuenta el PAdv propuesto, se puede afirmar que se mitigan las amenazas identificadas en la sección 5.1 y justificadas en la tabla 5.6, teniendo en consideración que de la amenaza A1 a A5 la mitigación proporcionada sería la misma para todas. Como se puede observar, el enfoque es complementario al del capítulo anterior, ya que no se trata de modificar la manera de trabajar del proveedor de identidades para mitigar las amenazas para la privacidad sino de empoderar al usuario de manera que pueda ser una parte activa en este proceso de mitigación tomando decisiones de calidad gracias a las recomendaciones recibidas.

Tabla 5.6: Mitigaciones que permite el PAdv a las amenazas para la privacidad identificadas en el capítulo 3

Amenaza	Descripción	Mitigación
A1	Perfilado del usuario en el modelo federado	El uso del PAdv permite, no evitar que los proveedores materialicen estas cinco amenazas relacionadas con el perfilado, la identificación o la revelación de datos a terceros, sino que los propios usuarios gestionen el riesgo que corren según su tolerancia a este riesgo. De esta forma, podrán mitigar estas cinco amenazas realizando cambios en sus manera de interactuar con determinados proveedores (datos proporcionados, consentimientos, configuraciones, etc.) o incluso evitarlas por completo dejando de trabajar con ellos.
A2	Perfilado del usuario en el uso de servicios	
A3	Perfilado del usuario según su localización	
A4	Identificación del usuario en el modelo federado	
A5	Revelación de datos del usuario en el modelo federado	
A8	Incumplimiento de las regulaciones, leyes o política	El uso del PAdv permite que un usuario realice cambios en sus manera de interactuar con determinados proveedores si hay riesgo de incumplimiento (por ejemplo, no compartiendo datos personales o escogiendo una configuración determinada) o incluso dejar de trabajar con aquellos que no garantizan este cumplimiento.

Capítulo 6

Conclusiones y trabajo futuro

En este capítulo se presentan las conclusiones más relevantes que se pueden extraer del trabajo de investigación realizado. Primero se analizan las conclusiones generales que provienen del estudio teórico realizado y del análisis de todos los resultados experimentales obtenidos en relación con los dos objetivos generales para la tesis que se propusieron en el capítulo 1. A continuación se resumen las conclusiones específicas, relacionadas con los objetivos específicos identificados en ese mismo capítulo, y se detallan las líneas más interesantes de trabajo futuro que se han detectado.

6.1. Conclusiones generales

La conclusión más importante que se puede extraer de esta tesis doctoral es la confirmación de la hipótesis de partida presentada en el capítulo 1. Es decir, se ha demostrado que es posible modelar las amenazas para la privacidad que supone el uso de esquemas federados para la gestión de identidades y accesos. Y que es posible, a partir de un modelo de este tipo, completo, exhaustivo y sistemático, proponer estrategias de privacidad centradas en el usuario que permitan evitar o mitigar las amenazas encontradas.

Para llegar a esta conclusión se ha realizado el análisis exhaustivo del problema de la gestión de identidades y accesos federada y de las estrategias de privacidad propuestas en la metodología para llevar a cabo esta investigación, que también ha demostrado ser adecuada. Pero las conclusiones que se han extraído no se basan exclusivamente en este análisis o en el proceso de modelado de amenazas para la privacidad que se ha realizado, sino que se apoya en gran cantidad de resultados experimentales obtenidos en entornos controlados. Para realizar todos los experimentos necesarios se han implementado prototipos de las estrategias de privacidad propuestas (portal

unificado y *Privacy Advisor*) que han permitido validarlas, evaluarlas y corroborar todas las conclusiones extraídas, así como estudiar los aspectos más prácticos del problema considerado.

En este plano más práctico, la conclusión más importante ha sido que es posible incorporar las estrategias propuestas a los flujos definidos por las especificaciones federadas actuales sin necesidad de modificarlos, lo que facilita enormemente que las propuestas realizadas en esta tesis doctoral se trasladen a entornos de producción en el corto plazo.

6.2. Conclusiones específicas

Las conclusiones asociadas a los seis objetivos específicos identificados para este trabajo de investigación se pueden resumir en:

1. En el capítulo 3 se ha decidido que LINDDUN es la metodología de modelado de amenazas adecuada para identificar de manera exhaustiva y sistemática las amenazas para la privacidad en el contexto de esta investigación.

2. También en este capítulo se ha propuesto un diagrama de flujo de datos común a todos los esquemas federados. Se ha decidido además trabajar con el modelo *honest but curious* para todos los proveedores de la federación (tanto IdP como RP), ya que es el que mejor se ajusta a la posición actual de estos agentes tanto en contextos de *Social Login* como de *Single Sign-On*.

3. La aplicación de esta metodología ha permitido modelar hasta ocho amenazas diferentes para los usuarios del modelo federado de gestión de identidades y accesos, teniendo en cuenta los impactos para ellos de estas amenazas, como personas o individuos (no sólo los impactos técnicos tradicionales) y el ciclo de vida de los datos que se producen en estos flujos. Estas amenazas se han encontrado para todas las categorías del modelo, dadas las condiciones y asunciones que se han tenido en cuenta durante el proceso de modelado. Por este mismo motivo, las amenazas encontradas lo son para entidades, procesos y almacenamiento, ninguna de ellas afecta al flujo de datos.

4. En los capítulos 4 y 5 se han propuesto estrategias de privacidad centradas en el usuario que permitan evitar o mitigar estas ocho amenazas encontradas en el modelo. El portal unificado (capítulo 4) y el *Privacy Advisor* (capítulo 5) han demostrado proporcionarle al usuario,

dentro de la federación de identidades, control sobre sus propios datos y decisiones. Y esto contribuye en gran medida a evitar o mitigar los distintos tipos de perfilado, identificación, revelación de datos, falta de información, de control o incumplimiento que deben preocupar a los usuarios cuando utilizan esquemas federados.

5. En el caso del capítulo 4, el portal unificado ha demostrado, mediante una sencilla modificación en la manera de implementar el proveedor de identidades, facilitar al usuario el ejercicio de sus derechos relacionados con la protección de datos. En el caso del capítulo 5, un nuevo agente en las federaciones denominado *Privacy Advisor* ha sido validado como un mecanismo eficiente y eficaz, basado en un sistema de recomendación, para la ayuda en la toma de decisiones relacionadas con el uso de proveedores específicos o con las configuraciones de privacidad. Las dos estrategias propuestas presentan ciertos aspectos en común que son esenciales para cumplir con los objetivos planteados para esta tesis doctoral como puedan ser:

   - Facilidad de integración con las especificaciones actuales: El modelo federado está muy extendido en la actualidad, tanto como soporte para el *Social Login* como para el *Single Sign-On*. Las estrategias de privacidad propuestas en esta tesis no implican cambios en las especificaciones actuales, sino que se pueden incorporar directamente en la fase de implementación. No hay que modificar las especificaciones ni aspectos de las mismas ya implementados y en funcionamiento en la actualidad. Esta forma de evitar o mitigar las amenazas para la privacidad, mediante extensiones a lo que ya está en producción, ayudará mucho a facilitar la adopción de las propuestas realizadas.
   - Personalización: Las estrategias propuestas, para ser realmente centradas en el usuario, permiten un control o soporte a la decisión individualizado, que tiene en cuenta que existen distintos tipos de usuarios con distintos niveles de conocimientos y de necesidades.
   - Complejidad reducida: Las estrategias propuestas son sencillas de utilizar y resultan útiles, tanto en el ejercicio de los derechos como en la toma de decisiones.
   - Granuralidad: Las dos estrategias propuestas permiten el trabajo por capas, gradual, con diferentes niveles de detalle según las necesidades y preferencias de cada usuario y según el caso de uso.

6. En los capítulos 4 y 5 se han implementado prototipos de las estrategias propuestas y se han integrado en flujos realizados con la especificación OpenID Connect, una de las más extendidas en la actualidad dado que es la base de Google Connect, Facebook Connect, Apple ID o Mobile Connect, por mencionar sólo algunos ejemplos. Estos prototipos han permitido validar y evaluar las estrategias propuestas en escenarios reales.

Algunas de estas conclusiones han permitido desarrollar actividades de docencia y divulgación en el ámbito de la privacidad y la protección de datos. Otras han sido publicadas en congresos científicos nacionales [146], congresos científicos internacionales [147] o revistas de impacto [148].

6.3. Líneas de trabajo futuro

A lo largo de esta tesis doctoral se han identificado algunas líneas de trabajo, que han quedado fuera del alcance de la presente investigación pero que sería muy interesante abordar en el futuro, como pueden ser:

- Se puede investigar en cómo utilizar el portal unificado propuesto en el capítulo 4 para mejorar la transparencia de los proveedores, de manera que no sólo se permita ejercer los derechos de los usuarios sino que sea parte integral de su propuesta de valor, por ejemplo, incluyendo las peticiones que se han admitido o rechazado en el pasado y los motivos del rechazo o permitiendo auditoría continua.

- Es necesario profundizar en cómo obtener información acerca de los proveedores que forman parte de la federación de identidades (IdP o RP), sobre cómo protegen la privacidad de sus usuarios, incluso cuando sus métodos son opacos. Y también en cómo completar o mejorar las técnicas contempladas en esta tesis doctoral para recoger información sobre los atributos ya propuestos en esta investigación (patrones de diseño, certificaciones de privacidad, cumplimiento de la normativa).

- En lo que se refiere a la incorporación de nuevos atributos, como podrían ser por ejemplo, la designación de un responsable de la protección de datos o el país en el que se almacenan los datos, las técnicas de extracción automática de información no serían complicadas. Pero en otros casos, como el análisis de las políticas de privacidad, se podría explorar el uso de técnicas de procesamiento del lenguaje natural.

- También es necesario avanzar en la construcción de sistemas de reputación de proveedores, tanto en los protocolos que los soportan como en la definición de las métricas de reputación o los pesos asignados a estas métricas en función del evaluador.

- Igualmente, sería interesante investigar en sistemas de recomendación híbridos que tengan en cuenta al usuario para el que se hace la recomendación y a otros usuarios similares. En otras palabras, que fusionen el filtrado colaborativo y el basado en el contenido para combinar los puntos fuertes de ambos enfoques.

- Por último, se debería profundizar en formas de fomentar la colaboración de los distintos agentes de las federaciones de identidad para que las estrategias centradas en los usuarios sean lo más útiles posible, incentivando su uso por parte de los diferentes actores, resolviendo las asimetrías de los modelos federados, incentivando modelos de negocio más éticos, etc., Estas soluciones en todos los casos deben proponer estrategias, mecanismos y soluciones que no obliguen a modificaciones significativas de las especificaciones que ya existen y que están tan extendidas en Internet en la actualidad.

Bibliografía

[1] Lau Kung Wei and S. Jarzabek, "A generic discretionary access control system for reuse frameworks," *Proceedings. The Twenty-Second Annual International Computer Software and Applications Conference (Compsac '98) (Cat. No.98CB 36241)*, pp. 356–361, 1998.

[2] K. Harsha, Bharath M. Palavalli, Shrisha Rao, and Ashwin, "Lothlorien: Mandatory Access Control using Linux Security Modules," *IEEE International Conference on Internet Multimedia Services Architecture and Applications (IMSAA)*, pp. 1–6, 2009.

[3] SELinux, "SELinux project." `https://selinuxproject.org`.

[4] David Ferraiolo and Richard Kuhn, "Role-Based Access Controls," *15th National Computer Security Conference*, pp. 554–563, 1992.

[5] Microsoft, "Windows - Active Directory Permissions role." `https://docs.microsoft.com/en-us/exchange/active-directory-permissions-role-exchange-2013-help`.

[6] National Institute of Standards and Technology (NIST), "NIST Special Publication 800-162 - Guide to Attribute Based Access Control (ABAC) Definition and Considerations." `https://doi.org/10.6028/NIST.SP.800-162`, 2014.

[7] OASIS, "eXtensible Access Control Markup Language (XACML) Version 3.0." `https://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.pdf`, 2013.

[8] Xin Jin, Ravi Sandhu, and Ram Krishnan, "RABAC: Role-Centric Attribute- Based Access Control," *International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security (MMM-ACNS)*, vol. 7531, pp. 84–96, 2012.

[9] Yuan Cao and Lin Yang, "A survey of Identity Management technology," *2010 IEEE International Conference on Information Theory and Information Security*, pp. 287–293, 2010.

[10] Gail-Joon Ahn and Moonam Ko, "User-centric Privacy Management for Federated Identity Management," *International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pp. 187–195, 2007.

[11] Abhilasha BhargavSpantzel, Jan Camenisch, Thomas Gross, and Dieter Sommer, "User Centricity: A Taxonomy and Open Issues," *Journal of Computer Security*, vol. 15, no. 5, pp. 493–527, 2007.

[12] OASIS, "Security Assertion Markup Language (SAML) v2.0 Technical Overview." `http://docs.oasis-open.org/security/saml/Post2.0/sstc-saml-tech-overview-2.0.pdf`, 2008.

[13] Internet Engineering Task Force (IETF), "RFC 6749: The OAuth 2.0 Authorization Framework." `https://tools.ietf.org/html/rfc6749`, 2012.

[14] OpenID Foundation, "OpenID Connect Core 1.0 incorporating errata set 1." `https://openid.net/specs/openid-connect-core-1_0.txt`, 2014.

[15] OpenID Foundation, "OpenID Connect back-channel logout 1.0." `https://openid.net/specs/openid-connect-backchannel-1_0.txt`, 2022.

[16] Frank J. Ohlhorst, *Big Data Analytics: Turning Big Data into Big Money*. John Wiley & Sons, 1st edition ed., 2012.

[17] Lee Kah Moey, Norliza Katuk, and Omar Mohd Hasbullah, "Social login privacy alert: Does it improve privacy awareness of Facebook users," *IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, pp. 95–100, 2016.

[18] Marios Isaakidis, Harry Halpin, and George Danezis, "UnlimitID: Privacy-Preserving Federated Identity Management using Algebraic MACs," *Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society (WPES)*, pp. 139–142, 2016.

[19] Jorge Navas and Marta Beltrán, "Understanding and mitigating OpenID Connect threats," *Computers & Security*, vol. 84, pp. 1–16, 2019.

[20] Marc Langheinrich, "Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems," *Ubiquitous Computing (Ubicomp)*, vol. 2201, pp. 273–291, 2001.

[21] Jeroen van Rest, Daniel Boonstra, Maarten Everts, Martin van Rijn, and Ron van Paassen, "Designing Privacy-by-Design," *Privacy Technologies and Policy*, pp. 55–72, 2014.

[22] National Institute of Standards and Technology (NIST), "NIST Special Publication 800-122 - Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)." `https://doi.org/10.6028/NIST.SP.800-122`, 2010.

[23] Sabrina De Capitani di Vimercati, Sara Foresti, Giovanni Livraga, and Pierangela Samarati, "Data Privacy: Definitions and Techniques," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20, no. 6, pp. 793–817, 2012.

[24] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity," *IEEE 23rd International Conference on Data Engineering*, pp. 106–115, 2007.

[25] Cynthia Dwork, "Differential Privacy," *Automata, Languages and Programming*, vol. 4052, pp. 1–12, 2006.

[26] Rui Chen, Noman Mohammed, Benjamin C. M. Fung, Bipin C. Desai, and Li Xiong, "Publishing SetValued Data via Differential Privacy," *Proceedings of the VLDB Endowment*, vol. 4, no. 11, pp. 1087–1098, 2011.

[27] Aaron Roth and Tim Roughgarden, "Interactive Privacy via the Median Mechanism," *Proceedings of the Forty-Second ACM Symposium on Theory of Computing (STOC)*, pp. 765–774, 2010.

[28] Jordi Soria-Comas, Josep Domingo-Ferrer, David Sánchez, and David Megías, "Individual Differential Privacy: A Utility-Preserving Formulation of Differential Privacy Guarantees," *IEEE Transactions On Information Forensics And Security*, vol. 12, no. 6, pp. 1418–1429, 2017.

[29] Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil Vadhan, "Computational Differential Privacy," *Advances in Cryptology - CRYPTO 2009*, pp. 126–142, 2009.

[30] National Institute of Standards and Technology (NIST), "NIST Special Publication 800-57 Part 1 Revision 5 - Recommendation for Key Management Part 1: General." `https://doi.org/10.6028/NIST.SP.800-57pt1r5`, 2020.

[31] John Bethencourt, Amit Sahai, and Brent Waters, "Ciphertext-Policy Attribute-Based Encryption," *2007 IEEE Symposium on Security and Privacy (SP)*, pp. 321–334, 2007.

[32] D. Nuñez, Isaac Agudo, and Javier Lopez, "Proxy Re-Encryption: Analysis of Constructions and its Application to Secure Access Delegation," *Journal of Network and Computer Applications*, vol. 87, pp. 193–209, 2017.

[33] Matthew Green and Giuseppe Ateniese, "Identity-Based Proxy Re-encryption," *Applied Cryptography and Network Security (ACNS)*, vol. 4521, pp. 288–306, 2007.

[34] Dan Boneh and Matt Franklin, "Identity-Based Encryption from the Weil Pairing," *Advances in Cryptology (CRYPTO)*, vol. 2139, pp. 213–229, 2001.

[35] Roberta Mayumi Matsunaga, Ivan Ricarte, Tania Basso, and Regina Moraes, "Towards an Ontology-Based definition of Data Anonymization Policy for Cloud Computing and Big Data," *47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops*, pp. 75–82, 2017.

[36] David Chaum, "Group Signatures," *Advances in Cryptology (EUROCRYPT)*, vol. 547, pp. 257–265, 1991.

[37] Jan Camenisch and Markus Stadler, "Efficient Group Signature Schemes for Large Groups," *Advances in Cryptology (CRYPTO 1997)*, vol. 1294, pp. 410–424, 1997.

[38] X. Ding, Gene Tsudik, and Shouhuai Xu, "Leak-free Group Signatures with Immediate Revocation," *24th International Conference on Distributed Computing Systems (ICDCS)*, pp. 608–615, 2004.

[39] Stephen R. Tate and He Ge, "A Group Signature Scheme with Signature Claiming and Variable Linkability," *IEEE International Performance Computing and Communications Conference*, pp. 8 pp.–504 pp., 2006.

[40] V. Kumar, H. Li, J.-M. Park, K. Bian, and Y. Yang, "Group Signatures with Probabilistic Revocation: A Computationally-Scalable Approach for Providing Privacy-Preserving Authentication," *22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 1334–1345, 2015.

[41] Ronald L. Rivest, Adi Shamir, and Yael Tauman, "How to Leak a Secret," *Advances in Cryptology (ASIACRYPT)*, vol. 2248, pp. 552–565, 2001.

[42] Ronald L. Rivest, Adi Shamir, and Yael Tauman, "How to Leak a Secret: Theory and Applications of Ring Signatures," *Theoretical Computer Science: essays in Memory of Shimon Even*, pp. 164–186, 2006.

[43] Aggelos Kiayias and Hong-Sheng Zhou, "Hidden Identity-Based Signatures," *IET Information Security*, vol. 3, pp. 119 – 127, Oct. 2009.

[44] Liu Xin and Xu Qiu-liang, "Practical Hidden Identity-based Signature Scheme from Bilinear Pairings," *3rd International Conference on Computer Science and Information Technology*, vol. 6, pp. 97–102, 2010.

[45] Sherman S.M. Chow, Haibin Zhang, and Tao Zhang, "Real Hidden Identity-Based Signatures," *Financial Cryptography and Data Security*, vol. LNCS 10322, pp. 21–38, 2017.

[46] Cynthia Dwork, Moni Naor, and Amit Sahai, "Concurrent Zero-Knowledge," *Journal of the ACM*, vol. 51, no. 6, pp. 851–898, 2004.

[47] Moni Naor, "Deniable Ring Authentication," *Advances in Cryptology (CRYPTO 2002)*, pp. 481–498, 2002.

[48] Confidential Computing Consortium, "A Technical Analysis of Confidential Computing." `https://confidentialcomputing.io/wp-content/uploads/sites/85/2022/01/CCC-A-Technical-Analysis-of-Confidential-Computing-v1.2.pdf`, 2021.

[49] Christian Mainka, Vladislav Mladenov, Jörg Schwenk, and Tobias Wich, "SoK: Single Sign-On Security - An Evaluation of OpenID Connect," *IEEE European Symposium on Security and Privacy (EuroS&P)*, pp. 251–266, 2017.

[50] Daniel Fett, Ralf Küsters, and Guido Schmitz, "The Web SSO Standard OpenID Connect: In-Depth Formal Security Analysis and Security

Guidelines," *IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 189–202, 2017.

[51] Wanpeng Li, Chris Mitchell, and Thoman Chen, "OAuthGuard: Protecting User Security and Privacy with OAuth 2.0 and OpenID Connect," *Proceedings of the 5th ACM Workshop on Security Standardisation Research Workshop (SSR)*, pp. 35–44, 2019.

[52] Jose M. del Alamo, Miguel A. Monjas, Juan C. Yelmo, Beatriz San Miguel, Ruben Trapero, and Antonio M. Fernandez, "Self-service Privacy: User-Centric Privacy for Network-Centric Identity," *4th IFIP WG 11.11 international conference on Advances in Information and Communication Technology (IFIPTM)*, vol. 321, pp. 17–31, 2010.

[53] José M. del Álamo, Antonio M. Fernández, Rubén Trapero, Juan C. Yelmo, and Miguel A. Monjas, "A Privacy-Considerate Framework for Identity Management in Mobile Services," *Mobile Networks and Applications*, vol. 16, no. 4, pp. 446–459, 2011.

[54] Bernd Zwattendorfer, Daniel Slamanig, Klaus Stranacher, and Felix Hörandner, "A Federated Cloud Identity Broker-Model for Enhanced Privacy via Proxy Re-Encryption," in *Communications and Multimedia Security*, (Berlin, Heidelberg), pp. 92–103, Springer Berlin Heidelberg, 2014.

[55] D. Nuñez, I. Agudo, and J. Lopez, "Integrating OpenID with proxy re-encryption to enhance privacy in cloud-based identity services," *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, pp. 241–248, 2012.

[56] Patricia Cabarcos, Florina Almenárez-Mendoza, Felix Gomez Marmol, and Andrés Marín-López, "To Federate or Not To Federate: A Reputation-Based Mechanism to Dynamize Cooperation in Identity Management," *Wireless Personal Communications: An International Journal*, vol. 75, no. 3, pp. 1769–1786, 2014.

[57] Jorge Werner, Carla Merkle Westphall, Tahleen Rahman, R. Weingartner, Guilherme Geronimo, and Carlos Westphall, "An Approach to IdM with Privacy in the Cloud," *IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pp. 168–175, 2015.

[58] Lucas Bodnar, Carla Merkle Westphall, Jorge Werner, and Carlos Westphall, "Towards Privacy in Identity Management Dynamic Federations," *The Fifteenth International Conference on Networks (ICN)*, vol. 6, 2016.

[59] María Villarreal, Sergio Villarreal, Carla Merkle Westphall, and Jorge Werner, "Privacy Token: A Mechanism for User's Privacy Specification in Identity Management Systems for the Cloud," *The Sixteenth International Conference on Networks (ICN)*, pp. 53–58, 2017.

[60] Farzaneh Karegar, Nina Gerber, Melanie Volkamer, and Simone Fischer-Hübner, "Helping John to Make Informed Decisions on Using Social Login," *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, pp. 1165–1174, 2018.

[61] Rosa Sanchez, Florina Almenares, Patricia Arias, Daniel Díaz-Sánchez, and Andrés Marín, "Enhancing Privacy and Dynamic Federation in IdM for Consumer Cloud Computing," *IEEE Transactions on Consumer Electronics*, vol. 58, pp. 95 –103, 2012.

[62] R. Weingärtner and W. Carla Merkle, "A design towards personally identifiable information control and awareness in OpenID Connect identity providers," *IEEE International Conference on Computer and Information Technology (CIT)*, pp. 37–46, 2017.

[63] J. Werner and C. Merkle Westphall, "A Model for Identity Management with Privacy in the Cloud," *IEEE Symposium on Computers and Communication (ISCC)*, pp. 463–468, 2016.

[64] David Núñez and Isaac Agudo, "BlindIdM: A privacy-preserving approach for identity management as a service," *International Journal of Information Security*, vol. 13, pp. 199–215, 2014.

[65] David Núñez, Isaac Agudo, and Javier Lopez, "Privacy-Preserving Identity Management as a Service," *International Journal of Information Security*, vol. 13, pp. 199–215, 2014.

[66] Arkajit Dey and Stephen Weis, "PseudoID: Enhancing Privacy in Federated Login," *3rd Hot Topics in Privacy Enhancing Technologies*, pp. 95–107, 2010.

[67] Wanpeng Li and Chris Mitchell, "User Access Privacy in OAuth 2.0 and OpenID Connect," *IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pp. 664–672, 2020.

[68] Daniel Fett, Ralf Küsters, and Guido Schmitz, "SPRESSO: A Secure, Privacy-Respecting Single Sign-On System for the Web," *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pp. 1358–1369, 2015.

[69] Sven Hammann, Ralf Sasse, and David Basin, "Privacy-Preserving OpenID Connect," *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (ASIA CCS)*, pp. 277–289, 2020.

[70] M. Potthast, C. Forler, E. List, and S. Lucks, "Passphone: Outsourcing Phone-based Web Authentication while Protecting User Privacy," *21st Nordic Conference (NordSec)*, vol. LNCS 10014, pp. 235–255, 2016.

[71] S. Zeng, S. Tan, Y. Chen, He, Mingxing, Xia, Meichen, and Li, Xiao, "Privacy-preserving Location-based Service based on Deniable Authentication," *IEEE/ACM 9th International Conference on Utility and Cloud Computing (UCC)*, pp. 276–281, 2016.

[72] H. Hu, J. Xu, Q. Chen, and Z. Yang, "Authenticating Location-based Services without Compromising Location Privacy," *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, pp. 301–312, 2012.

[73] Muhammad Rizwan Asghar, Michael Backes, and Milivoj Simeonovski, "PRIMA: privacy-preserving identity and access management at internet-scale," *IEEE International Conference on Communications (ICC)*, pp. 1–6, 2018.

[74] H. Halpin, "NEXTLEAP: Decentralizing Identity with Privacy for Secure Messaging," *Proceedings of the 12th International Conference on Availability, Reliability and Security*, no. 92, pp. 1–10, 2017.

[75] Paul Fremantle and Benjamin Aziz, "OAuthing: Privacy-enhancing Federation for the Internet of Things," *Cloudification of the Internet of Things (CIoT)*, pp. 1–6, 2016.

[76] M. S. Ferdous and R. Poet, "Portable Personal Identity Provider in Mobile Phones," *12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, pp. 736–745, 2013.

[77] Michael Hölzl, Michael Roland, and René Mayrhofer, "Real-World Identification: Towards a Privacy-Aware Mobile eID for Physical and

Offline Verification," *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media (MoMM)*, pp. 280–283, 2016.

[78] Yuto Iso and Takamichi Saito, "A Proposal and Implementation of an ID Federation That Conceals a Web Service from an Authentication Server," *IEEE 29th International Conference on Advanced Information Networking and Applications*, pp. 347–351, 2015.

[79] Takamichi Saito, Yuta Tsunoda, and Daichi Miyata, "An Authorization Scheme Concealing Client's Access from Authentication Server," *10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 593–598, 2016.

[80] Martin Schanzenbach and Georg Bramm, "reclaimID: Secure, Self-Sovereign Identities using Name Systems and Attribute-Based Encryption," *17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 946–957, 2018.

[81] John Maheswaran, David Isaac Wolinsky, and Bryan Ford, "CryptoBook: An Architecture for Privacy Preserving Online Identities," *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks*, no. 14, pp. 1–7, 2013.

[82] Alexandre B. Augusto and Manuel Eduardo Correia, "OFELIA – A Secure Mobile Attribute Aggregation Infrastructure for User-Centric Identity Management," *IFIP Advances in Information and Communication Technology*, vol. 376, pp. 61–74, 2012.

[83] Liberty Alliance, "Liberty Alliance Project." `http://www.projectliberty.org/`.

[84] Bart P Knijnenburg and Jin Hongxia, "The persuasive effect of privacy recommendations," *Special Interest Group on Human-Computer Interaction (SIGHCI)*, no. 16, 2013.

[85] Kevin Huguenin, Igor Bilogrevic, Joana Soares Machado, Stefan Mihaila, Reza Shokri, Italo Dacosta, and Jean-Pierre Hubaux, "A predictive model for user motivation and utility implications of privacy-protection mechanisms in location check-ins," *IEEE Transactions on Mobile Computing*, vol. 17, no. 4, pp. 760–774, 2017.

[86] Jane Henriksen-Bulmer, *Incorporating contextual integrity into privacy decision making: a risk based approach*. PhD thesis, Bournemouth University, 2019.

[87] Kambiz Ghazinour, Stan Matwin, and Marina Sokolova, "Monitoring and recommending privacy settings in social networks," *Proceedings of the Joint EDBT/ICDT 2013 Workshops*, pp. 164–168, 2013.

[88] Jose Alemany, Elena del Val, Amit Sahai, J. Alberola, and A. García-Fornes, "Enhancing the privacy risk awareness of teenagers in online social networks through soft-paternalism mechanisms," *International Journal of Human-Computer Studies*, vol. 129, pp. 27–40, 2019.

[89] Kambiz Ghazinour, Stan Matwin, and Marina Sokolova, "YourPrivacyProtector: A Recommender System for Privacy Settings in Social Networks," *International Journal of Security, Privacy and Trust Management*, vol. 2, no. 4, pp. 11–25, 2013.

[90] Yang Zhang, Mathias Humbert, Tahleen Rahman, Cheng-Te Li, Jun Pang, and Michael Backes, "Tagvisor: A Privacy Advisor for Sharing Hashtags," *Proceedings of the 2018 World Wide Web Conference (WWW)*, pp. 287–296, 2018.

[91] Tribhuvanesh Orekondy, Bernt Schiele, and Mario Fritz, "Towards a Visual Privacy Advisor: Understanding and Predicting Privacy Risks in Images," *IEEE International Conference on Computer Vision (ICCV)*, pp. 3706–3715, 2017.

[92] Matius Chairani, Mathieu Chevalley, Abderrahmane Lazraq, and Sruti Bhagavatula, "By the user, for the user: A user-centric approach to quantifying the privacy of websites," *arXiv preprint arXiv:1911.05798*, 2019.

[93] Karin Bernsmed, Inger Anne Tøndel, and Åsmund Ahlmann Nyre, "Design and Implementation of a CBR-based Privacy Agent," *Seventh International Conference on Availability, Reliability and Security*, no. 317-326, 2012.

[94] Cheng Chang, Huaxin Li, Yichi Zhang, Suguo Du, Hui Cao, and Haojin Zhu, "Automated and personalized privacy policy extraction under GDPR consideration," *Proceedings of the International Conference on Wireless Algorithms, Systems, and Applications*, pp. 43–54, 2019.

[95] Rui Liu, Jiannong Cao, Kehuan Zhang, Wenyu Gao, Junbin Liang, and Lei Yang, "When privacy meets usability: unobtrusive privacy permission recommendation system for mobile apps based on crowdsourcing," *IEEE Transactions on Services Computing*, vol. 11, no. 5, pp. 864–878, 2018.

[96] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun (Aerin) Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti, "Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions," *Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS)*, pp. 27–41, 2016.

[97] Benjamin Andow, Samin Yaseer Mahmud, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Serge Egelman, "Actions Speak Louder than Words: Entity-Sensitive Privacy Policy and Data Flow Analysis with PoliCheck," *Proceedings of the 29th USENIX Conference on Security Symposium*, no. 56, pp. 985–1002, 2020.

[98] Norman Sadeh, Bin Liu, Anupam Das, Martin Degeling, and Florian Schaub, "Personalized Privacy Assistant," 2021.

[99] Odnan Ref Sanchez, Ilaria Torre, Yangyang He, and Bart P. Knijnenburg, "A recommendation approach for user privacy preferences in the fitness domain," *User Modeling and User-Adapted Interaction*, vol. 30, pp. 513–565, 2020.

[100] Mahsa Keshavarz and Mohd Anwar, "Towards Improving Privacy Control for Smart Homes: A Privacy Decision Framework," *16th Annual Conference on Privacy, Security and Trust (PST)*, pp. 1–3, 2018.

[101] imec-DistriNet Research Group, "LINDDUN Privacy Threat Modeling." `https://linddun.org/`.

[102] Antonio Robles-González, Javier Parra-Arnau, and Jordi Forné, "A LINDDUN-Based Framework for Privacy Threat Analysis on Identification and Authentication Processes," *Computers & Security*, vol. 94, no. 101755, 2020.

[103] K. Wuyts, L. Sion, and W. Joosen, "LINDDUN GO: A Lightweight Approach to Privacy Threat Modeling," *IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*, pp. 302–309, 2020.

[104] Leonardo Horn Iwaya, Muhammad Ali Babar, Awais Rashid, and Chamila Wijayarathna, "On the Privacy of Mental Health Apps: An Empirical Investigation and its Implications for Apps Development," *arXiv preprint arXiv:2201.09006*, 2022.

[105] Microsoft, "STRIDE." https://docs.microsoft.com/en-us/previous-versions/commerce-server/ee823878(v=cs.20).

[106] Microsoft, "Damage, Reproducibility, Exploitability, Affected users and Discoverability (DREAD)." https://docs.microsoft.com/en-us/windows-hardware/drivers/driversecurity/threat-modeling-for-drivers#the-dread-approach-to-threat-assessment.

[107] European Network and Information Security Agency (ENISA), "Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE)." https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory/rm-ra-methods/m_octave.html.

[108] MITRE, "Threat Assessment and Remediation Analysis (TARA)." https://www.mitre.org/publications/technical-papers/threat-assessment-and-remediation-analysis-tara.

[109] MITRE, "MITRE ATT&CK." https://attack.mitre.org/.

[110] MITRE, "CAPEC." https://capec.mitre.org/.

[111] imec-DistriNet Research Group, "LINDDUN Privacy Threat Trees catalog." https://www.linddun.org/linddun-threat-catalog.

[112] Parlamento Europeo y del Consejo de la Unión Europea, "Reglamento General de Protección de Datos (RGPD)." https://www.boe.es/doue/2016/119/L00001-00088.pdf, 2016.

[113] Nina Gerber, Paul Gerber, and Melanie Volkamer, "Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior," *Computers & Security*, vol. 77, pp. 226–261, 2018.

[114] Kantara Initiative, "Consent Receipt Specification 1.1.0." https://kantarainitiative.org/file-downloads/consent-receipt-specification-v1-1-0/, 2018.

[115] Sherrie Y. X. Komiak and Izak Benbasat, "The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents," *Management Information Systems Research Center, University of Minnesota*, vol. 30, no. 4, pp. 941–960, 2006.

[116] Jingjing Zhang and Shawn P. Curley, "Exploring Explanation Effects on Consumers' Trust in Online Recommender Agents," *International Journal of Human-Computer Interaction*, vol. 34, no. 5, pp. 421–432, 2018.

[117] Bo Xiao and Izak Benbasat, "An empirical examination of the influence of biased personalized product recommendations on consumers' decision making outcomes," *Decision Support Systems*, vol. 110, pp. 46–57, 2018.

[118] Robin Van Meteren and Maarten Van Someren, "Using content-based filtering for recommendation," *Proceedings of the machine learning in the new information age (MLnet/ECML2000) workshop*, vol. 30, pp. 47–56, 2000.

[119] Poonam B. Thorat, R. M. Goudar, and Sunita Barve, "Survey on collaborative filtering, content-based filtering and hybrid recommendation system," *International Journal of Computer Applications*, vol. 110, no. 4, pp. 31–36, 2015.

[120] SimilarWeb, "Top Websites Ranking." `https://www.similarweb.com/top-websites/`.

[121] National Institute of Standards and Technology (NIST), "NIST Special Publication (SP) 800-30, Revision 1 - Guide for Conducting Risk Assessments." `https://www.nist.gov/privacy-framework/nist-sp-800-30`, 2012.

[122] Yang Yang, Xuehui Du, and Zhi Yang, "PRADroid: Privacy Risk Assessment for Android Applications," *IEEE 5th International Conference on Cryptography, Security and Privacy (CSP)*, pp. 90–95, 2021.

[123] Peter Katsumata, Judy Hemenway, and Wes Gavins, "Cybersecurity Risk Management," *MILCOM 2010 Military Communications Conference*, pp. 890–895, 2010.

[124] Harry Brignull, "Dark Patterns (Deceptive Design)." `https://www.darkpatterns.org/`.

[125] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs, "The Dark (Patterns) Side of UX Design," *Proceedings of the 2018 CHI Conference on Human Factors in Computing (CHI)*, no. 534, pp. 1–14, 2018.

[126] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher, "Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns," *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 4, pp. 237–254, 2016.

[127] EuroPriSe GmbH, "European Privacy Seal (EurPriSe)." `https://www.euprivacyseal.com/`.

[128] T. Inc., "TrustArc." `https://www.trustarc.com`.

[129] ePrivacy GmbH, "ePrivacy." `https://www.eprivacy.eu/`.

[130] Unabhängiges Landeszentrum für Datenschutz, "Unabhängiges Lande center for Data protection (ULD) Register Experts." `https://www.datenschutzzentrum.de/guetesiegel/register-sachverstaendige/`.

[131] TrustArc Inc., "TrustArc - About." `https://trustarc.com/why-trustarc/`.

[132] TrustArc Inc., "TrustArc - Awards." `https://trustarc.com/trustarc-awards/`.

[133] Internet Engineering Task Force (IETF), "RFC 7519: JSON Web Token (JWT)." `https://tools.ietf.org/html/rfc7519`, 2015.

[134] Manuel Urueña, Alfonso Muñoz, and David Larrabeiti, "Analysis of privacy vulnerabilities in single sign-on mechanisms for multimedia websites," *Multimed Tools Appl*, vol. 68, no. 1, pp. 159–176, 2014.

[135] A. K. Jaithunbi, S. Sabena, and L. SaiRamesh, "Trust evaluation of public cloud service providers using genetic algorithm with intelligent rules," *Wireless Personal Communications*, vol. 121, no. 4, pp. 3281–3295, 2021. Publisher: Springer.

[136] E. Kokoris-Kogias, O. Voutyras, and T. Varvarigou, "TRM-SIoT: A scalable hybrid trust & reputation model for the social Internet of Things," *IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–9, 2016.

[137] L. Barakat, P. Taylor, N. Griffiths, and S. Miles, "A Reputation-based Framework for Honest Provenance Reporting," *ACM Transactions on Internet Technology*, 2021. Publisher: Association for Computing Machinery (ACM).

[138] R. Govindaraj, P. Govindaraj, S. Chowdhury, D. Kim, D.-T. Tran, and A. N. Le, "A Review on Various Applications of Reputation Based Trust Management," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 10, 2021.

[139] Z. Zhou, M. Wang, C.-N. Yang, Z. Fu, X. Sun, and Q. J. Wu, "Blockchain-based decentralized reputation system in E-commerce environment," *Future Generation Computer Systems*, vol. 124, pp. 155–167, 2021.

[140] Leonard Richardson, "Beautiful Soup - Python Library." `https://www.crummy.com/software/BeautifulSoup/`.

[141] Parlamento Europeo y del Consejo de la Unión Europea, "Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC." `http://data.europa.eu/eli/reg/2014/910/oj`, 2014.

[142] Phibos, "pysslscan - Python Library." `https://pysslscan.readthedocs.io`.

[143] European Commission, "Trusted List Browser." `https://esignature.ec.europa.eu/efda/swagger-ui.html`.

[144] MSCI Inc., "The Global Industry Classification Standard (GICS)." `https://www.msci.com/gics`.

[145] James T. Croasmun and Lee Ostrom, "Using Likert-Type Scales in the Social Sciences," *Journal of Adult Education*, vol. 40, no. 1, pp. 19–22, 2011.

[146] Carlos Alberto Villarán and Marta Beltrán, "Agente de recomendación para mejorar la privacidad de los usuarios cuando usan gestión de identidades federada," *VI Jornadas Nacionales de Investigación en Ciberseguridad (JNIC)*, vol. 34, pp. 273–280, 2021.

[147] Carlos Alberto Villarán and Marta Beltrán, "Protecting End User's Privacy When using Social Login through GDPR Compliance," *Proceedings of the 18th International Conference on Security and Cryptography (SECRYPT)*, pp. 428–435, 2021.

[148] Carlos Alberto Villarán and Marta Beltrán, "User-Centric Privacy for Identity Federations Based on a Recommendation System," *Electronics*, vol. 11, no. 8, 2022.