Una billetera Blockchain privada y protegida para el intercambio de datos

Subtitle Draft Version 0.8

Daniel Hawthorne y Serafin L. Engel[1]
Alex Norta[2]

[1] *Pnyks Inc. 55E3 rd Ave. San Mateo, CA 94401 EE. UU.*
[2] *Grupo de Sistemas de Gran Escala, Universidad Tecnológica de Tallin, Akadeemia tee15A, 12616 Tallinn, Estonia*

Renuncia de Responsabilidades

Nada de lo aquí contenido constituye una oferta de venta, ni la solicitud de una oferta para comprar, ninguna marca, ni ninguna oferta, solicitud o envío de datos de participación en una jurisdicción en la que dicha oferta, solicitud o venta sea ilegal. Debe leer detenidamente y comprender completamente este documento y las actualizaciones.

Se requerirá que cada posible comprador simbólico se someta a un proceso de incorporación que incluye la verificación de identidad y cierta otra documentación, que debe leer detenidamente y comprender completamente porque estará legalmente obligado. Por favor asegúrese de consultar con asesores apropiados y otros.

Este informe técnico describe nuestra visión actual para la plataforma **Datawallet**. Mientras tenemos la intención de intentar realizar esta visión, reconozca que depende de una gran cantidad de factores y está sujeta a una cantidad considerable de riesgos. Es muy posible que la plataforma **Datawallet** nunca se implemente o adopte, o que solo una parte de nuestra visión se realice. No garantizamos, representamos ni garantizamos ninguna de las declaraciones en este documento, ya que se basan en nuestras creencias, expectativas y suposiciones actuales, sobre las cuales no puede haber seguridad debido a varios eventos anticipados y no anticipados que pueden ocurrir.

Tenga en cuenta que planeamos trabajar arduamente para lograr la visión presentada en este documento, pero que no puede confiar en que se haga realidad. Blockchain, criptomonedas y otros aspectos de nuestra tecnología y estos mercados están en su infancia y estarán sujetos a muchos desafíos, competencia y un entorno cambiante. Intentaremos actualizar nuestra comunidad a medida que las cosas crezcan y cambien, pero no asumimos la obligación de hacerlo.

Resumen

A medida que las personas se involucran más profundamente y, a menudo, con los servicios de Internet, como las redes sociales, la búsqueda y el comercio electrónico, crean cada vez más datos. Las conocidas compañías de medios sociales y una industria de corretaje de datos generan ingresos monetarios de esta información personal sin consultar explícitamente o compartir los ingresos con los usuarios que produjeron el contenido. Tal es el ecosistema de datos personales dominante.

Existe una clara necesidad de un intercambio de datos transparente y justo basado en el expreso consentimiento de los usuarios de Internet que producen los datos. Este documento técnico específica dicho sistema basado en un "**DataWallet**" controlado por el usuario y un sistema de **contrato inteligente** basado en blockchain para permitir el intercambio de datos transparente y mutuamente beneficioso entre las partes que consienten realizarlo.

Detallamos los requisitos de dicho sistema, su estructura y las interacciones dinámicas de sus componentes y partes necesarias a través de un ejemplo centrado en un simple intercambio de datos para **DXT** (el token de intercambio de datos).

Mostramos cómo este intercambio, es solo uno de los muchos casos de uso para el intercambio de datos propuesto, al describir cómo puede proporcionar una experiencia de usuario aumentada para servicios de video y música en línea, así como proporcionar información valiosa a los creadores de datos sobre ellos mismos.

Palabras clave: *intercambio de datos, blockchain, contrato inteligente, descentralizado, punto a punto.*

Índice

1. Introducción

A medida que las personas dedican cada vez más su tiempo a interactuar con servicios en línea, el volumen y la importancia de los datos creados a partir de estas interacciones aumenta simultáneamente.

Los usuarios de Internet están quedando afuera de este sector económico en rápido crecimiento, que tiene el poder de afectar su experiencia tanto en línea como fuera de ella. Existe una oportunidad para que los proveedores de datos recuperen el valor que crean en línea de los propietarios del ecosistema de datos actual.

Así surge una economía basada en datos en la que los usuarios producen una cantidad de contenido cada vez mayor en línea a través de las redes sociales, búsquedas en línea y la compra de productos en plataformas de comercio electrónico.
Sin embargo, la actual economía basada en datos, centrada en la "*industria de corretaje de datos*", es profundamente disfuncional para los creadores de datos y los consumidores de datos. La situación de los proveedores de datos queda bien reflejada en una audiencia judicial del Senado de los EE.UU[3] ., sobre el tema que concluyó que los clientes de la industria de corretaje de datos no tienen control sobre cómo se usan y monetizan sus datos personales y el contenido creado.

Los proveedores de datos lo saben: en un estudio reciente[4], el 81% de los encuestados afirman que no se sienten seguros utilizando las redes sociales para compartir información privada. La economía de intercambio de datos también se rompe para los consumidores de datos. Esta disfunción se puede entender en términos del silo, la calidad y la ética problemas.

El problema del silo es que los datos se encuentran dispersos en un campo de jardines amurallados, y la mayor parte del valor potencial solo está disponible si los datos se pueden recopilar de manera inteligente.
Esto contribuye al problema de calidad. Los datos disponibles a menudo son el resultado de modelos de concordancia probabilística (con resultados poco inspiradores [1], [9], [12], [25]), desactualizados o generalmente incongruentes con un verdadero compromiso en línea. Finalmente, el problema de la ética es claro: los datos son adquiridos y monetizados por las empresas sin informar a los usuarios propietarios.

El advenimiento de la tecnología blockchain[5] plantea un remedio al actual sistema de corretaje de datos disfuncional. Permite sistemas transparentes que pueden dar a los creadores de datos el control de sus datos al mismo tiempo que proporciona a los negocios informados por datos la más alta calidad, con datos de origen ético.

[3] https://tinyurl.com/senate-hearing-2013
[4] http://www.adweek.com/digital/study-pew-public-perceptions-privacy/
[5] Brevemente, el blockchain es una base de datos distribuida para verificar independientemente la propiedad del artefacto de cadena [24] en los valores hash que resultan de los resúmenes criptográficos [13], [19].

Las deficiencias del actual sistema de corretaje de datos y la necesidad de un nuevo enfoque para el intercambio voluntario de datos solo serán más claros dadas las tendencias tecnológicas y sociales actuales. En primer lugar, la producción de datos aumentará drásticamente. Se ha creado un asombroso 90% de todos los datos en los últimos dos años y se prevé que este mercado crezca un 27% por año. Esta explosión de datos se reforzará aún más con la creciente adopción de los sistemas de Internet de las cosas (IoT) [28]. En segundo lugar, las situaciones de seguridad y privacidad de los datos de los usuarios son, en general, insuficientes[6].

El alojamiento en la nube es el predeterminado para IoT y el almacenamiento en redes sociales a pesar de proporcionar un área de superficie de seguridad más amplia [15]. Para garantizar la seguridad y la privacidad de los datos, falta un marco de ingeniería de requisitos adecuado [2] para el desarrollo del sistema[7].

Dadas las deficiencias del sistema de intermediación de datos vigente, es sorprendente la poca disponibilidad de herramientas de gestión de datos personales para los creadores de datos. Una opción prometedora es considerar el desarrollo de navegadores web específicos basados en blockchain[8] que se esfuerzan por permitir el control del usuario sobre el contenido autogenerado.

Sin embargo, esto requiere un cambio importante en el comportamiento del usuario con los desafíos tecnológicos concomitantes. Este documento propone una alternativa que interrumpe el sistema de corretaje de datos, y no la experiencia de Internet de los creadores de datos. Proponemos un **DataWallet** basado en tecnología de cadena de bloqueo que restablece la confianza y el control para los usuarios, al tiempo que proporciona a los negocios informados por datos los perfiles de datos más completos posibles. En este documento, por lo tanto, describimos detalladamente un sistema que permite este ecosistema de intercambio de datos mutuamente beneficioso y de aceptación mutua. Continuamos respondiendo las siguientes preguntas progresivamente refinadas.

¿Cuáles son los requisitos de dicho ecosistema de intercambio de datos?
¿Cuál es la arquitectura estática del ecosistema que cumple estos requisitos? Y, finalmente, ¿cuáles son los protocolos de interacción dinámica de los interesados y la arquitectura que permite la economía de intercambio de datos deseada y mutuamente beneficiosa?

El resto de este trabajo se estructura de la siguiente manera. La *Sección 2* presenta un caso en real en marcha y literatura de antecedentes adicionales. La *Sección 3* define los requisitos para un sistema de **datawallet** basado en la tecnología blockchain. La *Sección 4* muestra la arquitectura del sistema del **datawallet** que se deriva de los requisitos. A continuación, la *Sección 5* explica el compromiso dinámico del sistema. Finalmente, la

[6] Estos problemas de seguridad se extienden a información extremadamente confidencial como ilustra la reciente violación de datos de Equifax [32]
[7] Específicamente para el dominio IoT, la seguridad y la privacidad de la previsión de la producción de datos parabólicamente en crecimiento no están garantizadas ya que los marcos de seguridad existentes no se adaptan a redes grandes de dispositivos heterogéneos. Por ejemplo, la subclase IoT de los denominados dispositivos portátiles [3] sin pilas de hardware y software seguros no puede autenticar el software en ejecución y, por lo tanto, no se pueden validar por sí mismos.
[8] https://www.cryptocoinsnews.com/blockstack-joins-browser-wars-decentralized-tokenizedblockchain-web-browser/

Sección 6 concluye este documento técnico y también analiza las limitaciones, los problemas abiertos y el trabajo futuro.

En consecuencia, la descripción del sistema está en tiempo presente, pero el sistema totalmente descentralizado descrito es el último concepto de la hoja de ruta de implementación (descrito en el *Apéndice A*). Además, para mayor claridad, enfocamos nuestra descripción en el sistema completamente descentralizado (*v3.0 en el Apéndice A.3*). Sin embargo, las versiones alojadas de estos componentes seguirán estando disponibles para los miembros de la comunidad interesados (*ver v2.0 en el Apéndice A.3*)

2. Caso Real y Literatura de soporte

La *Figura 1* muestra el estado actual de la generación y monetización de contenido de usuario. Los usuarios de Internet se muestran a la izquierda, generando y consumiendo datos a través de compras, búsqueda y socialización en la web. Mientras que cada servicio desarrolla su "*ecosistema cerrado*"[9], algunos optan por participar en intercambios de datos a gran escala entre ellos.

El lado derecho de la *Figura 1* muestra los numerosos intereses centralizados que consumen datos personales de los usuarios. Las categorías corporativas a la derecha pueden comprender organizaciones con fines de lucro, como bancos, compañías de publicidad, productores de servicios tangibles.

Además, las organizaciones sin fines de lucro afiliadas al gobierno también consumen datos de las redes sociales, por ejemplo, la policía y las agencias de inteligencia.

Por lo tanto, mientras que la *Figura 1* muestra que los usuarios individuales generan y comunican sus datos de contenido creado sin compensación monetaria, el lado corporativo debe pagar a los miembros de la nube donde se publican los datos, por los datos y contenido. Los usuarios individuales actualmente carecen de una herramienta efectiva que les permita monetizar su propios datos a su discreción

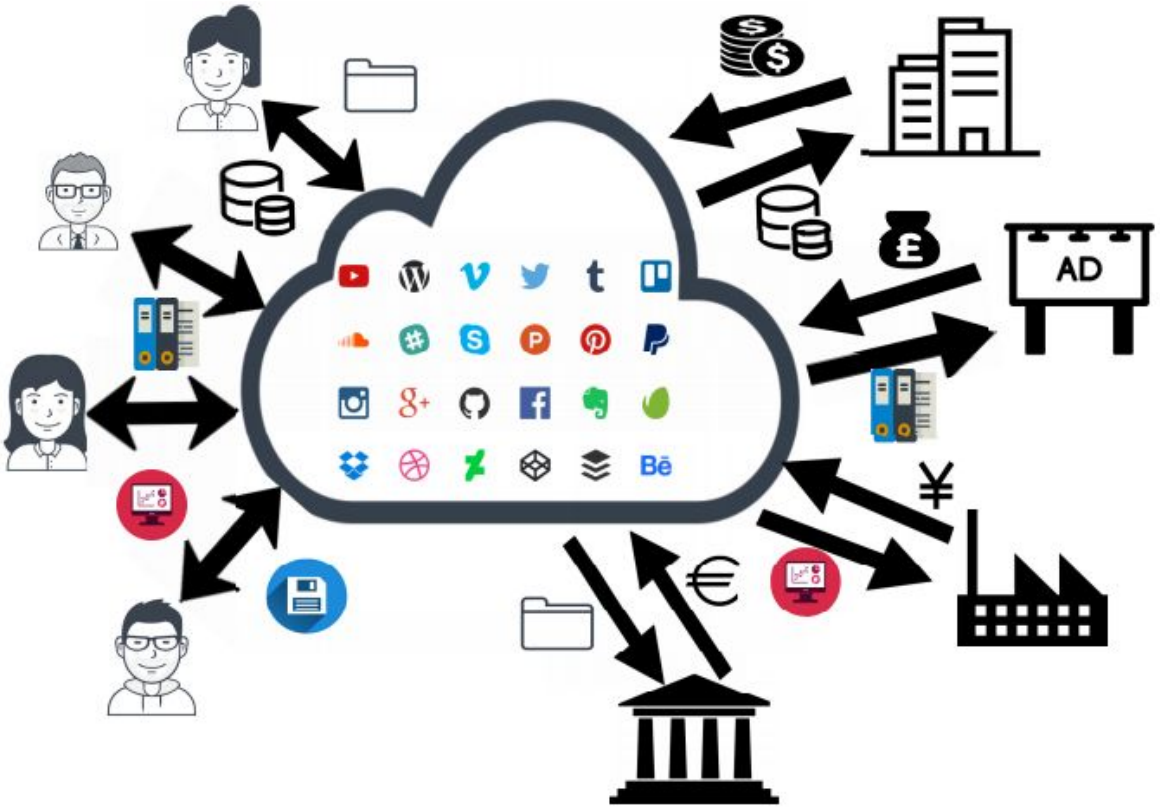

Figura 1: Situación de monetización actual del contenido de datos generado por el usuario en las redes sociales.

[9] *Es un sistema de software donde el operador o proveedor de servicios tiene control sobre las aplicaciones, el contenido y los medios, y restringe el acceso conveniente a aplicaciones o contenido no aprobado. Esto está en contraste con una plataforma abierta, donde los consumidores generalmente tienen acceso ilimitado a aplicaciones y contenido.*

Es por esto que hemos desarrollado una plataforma de administración de datos personales: **DataWallet** y una forma de obtener beneficios directamente de los datos personales, dx-Insights, que es la aplicación de datos del mercado y del conocimiento del consumidor de **DataWallet**. Utilizaremos dx-Insights para proporcionar ejemplos fundamentados en las secciones técnicas posteriores. Este caso es solo un ejemplo de un producto basado en datos personales[10].

En la *Sección 5* destacamos la extensión de posibles productos basados en datos personales al describir dos amplias categorías de productos: productos de datos para el proveedor de datos (*por ejemplo, para la introspección*). , (*toma de decisiones y auto-mejora*) y productos de datos que aumentan los servicios existentes personalizándolos con los datos del usuario.

2.1 Literatura de soporte

Ahora destacamos la literatura que motiva nuestro enfoque y los términos necesarios para comprender la metodología empleada en las secciones siguientes. Nuestra presentación del ecosistema **DataWallet** sigue la metodología establecida de diseño impulsado por modelos (**MDD**) [5] para garantizar el diseño, la arquitectura y la implementación del sistema de calidad[11].

De acuerdo con esta metodología, presentamos una serie de formalizaciones que brindan una profundidad cada vez mayor.

Tratamiento de las características estáticas y dinámicas del ecosistema propuesto. Empleamos :

1. Un modelo de objetivos para describir los requisitos del sistema,
2. Un diagrama de componentes **UML** para delinear la arquitectura estática, y finalmente
3. Un diagrama de secuencia **UML** para ilustrar el comportamiento dinámico del sistema. Ahora describimos estas formalizaciones a su vez.

Primero, los modelos de objetivos son parte del método de modelado orientado a agentes (**AOM**) [29] que empleamos para especificar formalmente las características del ecosistema **DataWallet** basado en blockchain.

Los modelos de objetivos de **AOM** utilizan la notación de la *Figura 2* para capturar los requisitos funcionales de un sistema en forma de objetivos funcionales, requisitos no funcionales o "*objetivos de calidad*" y agentes (*que pueden ser humanos o artificiales, es decir, agentes de software*) con roles especificados.

[10] De hecho, dx-Insights es simplemente el primer producto disponible en un intercambio de productos de datos abiertos
[11] El MDD implica una carga frontal de actividades en un proceso de diseño del sistema que da como resultado un alto grado de instancias de código y pruebas generadas.

Adoptamos un símbolo adicional, la flecha apuntando hacia la derecha, para denotar herencia múltiple de objetivos funcionales de nivel superior a un objetivo funcional de menor nivel, como se ve en la *Figura 6*.

Figura 2: Iconos del modelo de objetivo - AOM [29].

La parte central, de un modelo objetivos AOM, es un objetivo funcional denominado '*propuesta de valor*'. Esto captura el objetivo general de los sistemas y requiere una descomposición jerárquica para establecer una complejidad de desarrollo manejable. Los objetivos emocionales y de calidad se adjuntan a los objetivos funcionales respectivos y, dependiendo de la posición en la jerarquía, todos los objetivos funcionales de refinación también deben cumplir un objetivo de calidad asignado.

Lo mismo vale para los roles y su relación con la jerarquía de objetivos funcionales.

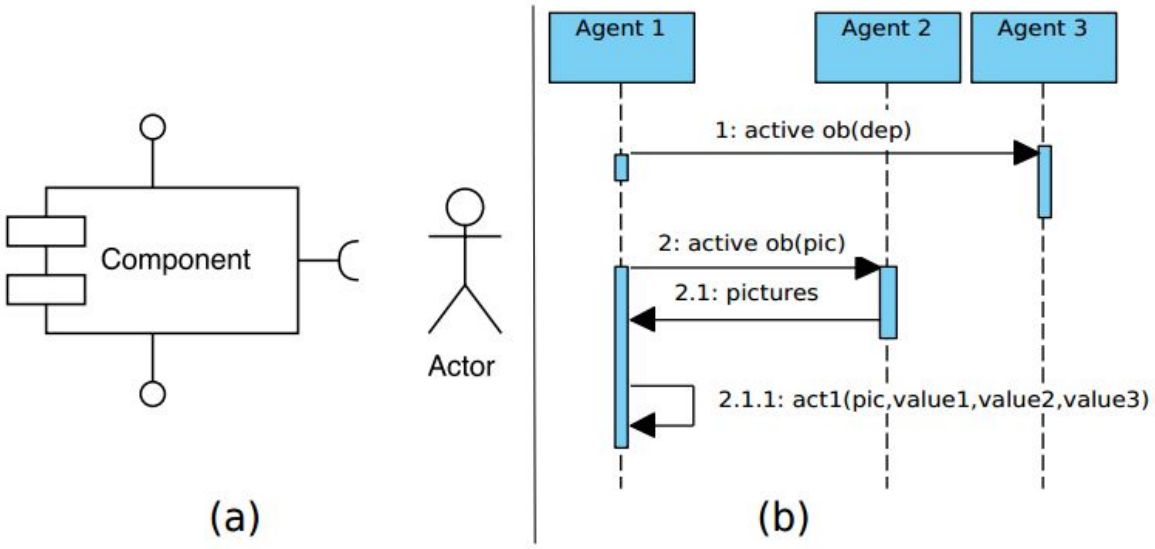

Figura 3: Notación del diagrama de componentes UML en (a) y diagramas de secuencia en (b).

En segundo lugar, desarrollamos un diagrama de componente UML independiente de la tecnología [6] de los objetivos funcionales de AOM especificados, para explicitar la estructura estática del sistema **DataWallet**. La *Figura 3(a)* muestra los elementos de notación UML donde los componentes están etiquetados como rectángulos.
Los componentes proporcionan interfaces para intercomunicación y son refinables con subcomponentes. También empleamos actores para indicar cómo los componentes interactúan con los diferentes interesados de acuerdo con las posiciones de los actores del modelo de objetivos de **AOM**.

Finalmente, en tercer lugar, usamos diagramas de secuencia UML [27] para especificar el comportamiento dinámico del sistema **DataWallet**.
La *Figura 3(b)* muestra un ejemplo con entidades comunicantes con rectángulos etiquetados. Debajo de cada entidad hay una línea punteada que indica las líneas de tiempo en las que se muestran las barras cuando las entidades respectivas están activas. Los arcos dirigidos etiquetados entre esas barras muestran el intercambio de mensajes de comunicación a lo largo de las líneas de tiempo.

El objetivo del enfoque **MDD** descrito, es capturar la naturaleza de los sistemas distribuidos [30] del ecosistema **DataWallet**. Si bien la *Figura 1* refleja el ecosistema de datos actual, en el que la mayoría de los usuarios producen datos dentro de las redes sociales, el advenimiento de IoT(Internet de las cosas), o sistemas ciber físicos (**CPS**) [26] dará como resultado una nueva dimensión en la que los usuarios crearán más datos cada vez.

Las **CPS** son conjuntos de objetos inteligentes integrados en Internet, como dispositivos IoT (Internet de las cosas), que pueden ser dirigidos por Clouds en tareas complejas de procesamiento. Como tal, **CPS** que integra capacidades computacionales y físicas que permiten la interacción con los humanos a través de diversos medios [4].

Estos nuevos vectores de interacción tienen la particularidad de expandir las capacidades y la realización del ser humano a través del cálculo, la comunicación y el control, para permitir una próxima generación de aviónica y vehículos, ciudades inteligentes, sistemas de producción industriales tipo 4.0, Servicios de salud, etc.

Finalmente, en lo que respecta a la tecnología blockchain, la tecnología de contrato inteligente es esencial para que el sistema **DataWallet** logre trazabilidad y seguridad.

Esto es esencial para los contratos inteligentes es una validación descentralizada de transacciones para las cuales la llamada prueba de trabajo (PoW) [31] es la más utilizada.
Los contratos inteligentes emplean un ledger distribuido y público llamado blockchain que registra eventos de transacción sin una autoridad central de confianza. El estándar actual de contrato inteligente es el Ethereum [33], a pesar de algunos inconvenientes ampliamente reconocidos.

- Primero, la validación de la transacción **PoW** no se escala y, por lo tanto, Ethereum no es factible para la mayoría de las aplicaciones industriales.
-
- En segundo lugar, el contrato inteligente **Solidity** afiliado a Ethereum no puede ser formalmente verificado [8] y fue recientemente pirateado[12] debido a fallas de seguridad que resultaron en una pérdida de aproximadamente u$s 50 millones[13].

[12] https://www.wired.com/2016/06/50-million-hack-just-showed-dao-human/
[13] https://bitcoinsmagazine.com/articles/ethereum-classic-hard-forks-diffuses-difficulty-bomb-1484350622/

Ser más escalable es una solución de contrato inteligente que utiliza validación de transacción de prueba de participación (**PoS**) [7] y fragmentación de cadena de bloque [16].

Por ejemplo, el sistema de contrato inteligente **Qtum** [11] usa **PoS** con éxito en su aplicación.

3 Metas del Ecosistema de Intercambio de Datos

Definimos las metas y los requisitos de un ecosistema de intercambio de datos basado en el consentimiento expreso utilizando el modelo de objetivo del modelado orientado a agentes (AOM). El **método AOM** es un enfoque de ingeniería de requisitos socio-técnicos utilizado para modelar sistemas dinámicos y complejos compuestos por agentes humanos y de software.

Para ayudar a la comprensión de los requisitos de todo el sistema, descomponemos el modelo en piezas coherentes centradas en las dos partes interesadas principales, es decir, los solicitantes de datos y los proveedores. Comenzamos describiendo los requisitos conjuntos de alto nivel del intercambio de datos en la *Sección 3.1*, y luego volvemos a los requisitos de los proveedores de datos en la *Sección 3.2*. Finalmente, concluimos describiendo en la *Sección 3.3* los requisitos de los solicitantes de datos[14]

3.1 Requisitos de alto nivel

Como se ve en la *Figura 4*, la propuesta principal de los interesados en el ecosistema, los solicitantes de datos y los proveedores de datos, es participar en un intercambio consensuado, transparente y seguro en torno a los datos. Para alimentar este intercambio, los solicitantes de datos necesitan un método claro y fácil para crear una solicitud de datos e incentivar su cumplimiento. Del mismo modo, los proveedores de datos que ofrecen personalizaciones a servicios existentes o crean nuevos servicios basados en datos necesitan una forma de ofrecer servicios junto con la solicitud de los datos necesarios y la recepción de compensación.

Para habilitar esta solicitud de datos de forma estructurada, los proveedores de datos requieren un método fácil, seguro y transparente para crear y administrar un perfil de datos intercalados. Luego necesitan un medio equitativo y transparente para explorar las solicitudes de datos disponibles y participar en las que elijan.

Del mismo modo, requieren un medio para personalizar sus servicios existentes y participar en nuevos servicios basados en su perfil de datos. Estos son los objetivos de calidad del sistema, ilustrados en la *Figura 4* y estructurados de acuerdo con [10], [14]. Ahora pasamos a describir cada objetivo de calidad en detalle.

[14] Tenga en cuenta que cuando se consideran los requisitos del solicitante de datos, el foco está en aquellos que desean una solución descentralizada de gestión de perfiles, que es el paso final en el desarrollo tecnológico descrito en el *Apéndice A* y es la solución de gestión de perfiles centralizada desarrollada actualmente por DataWallet

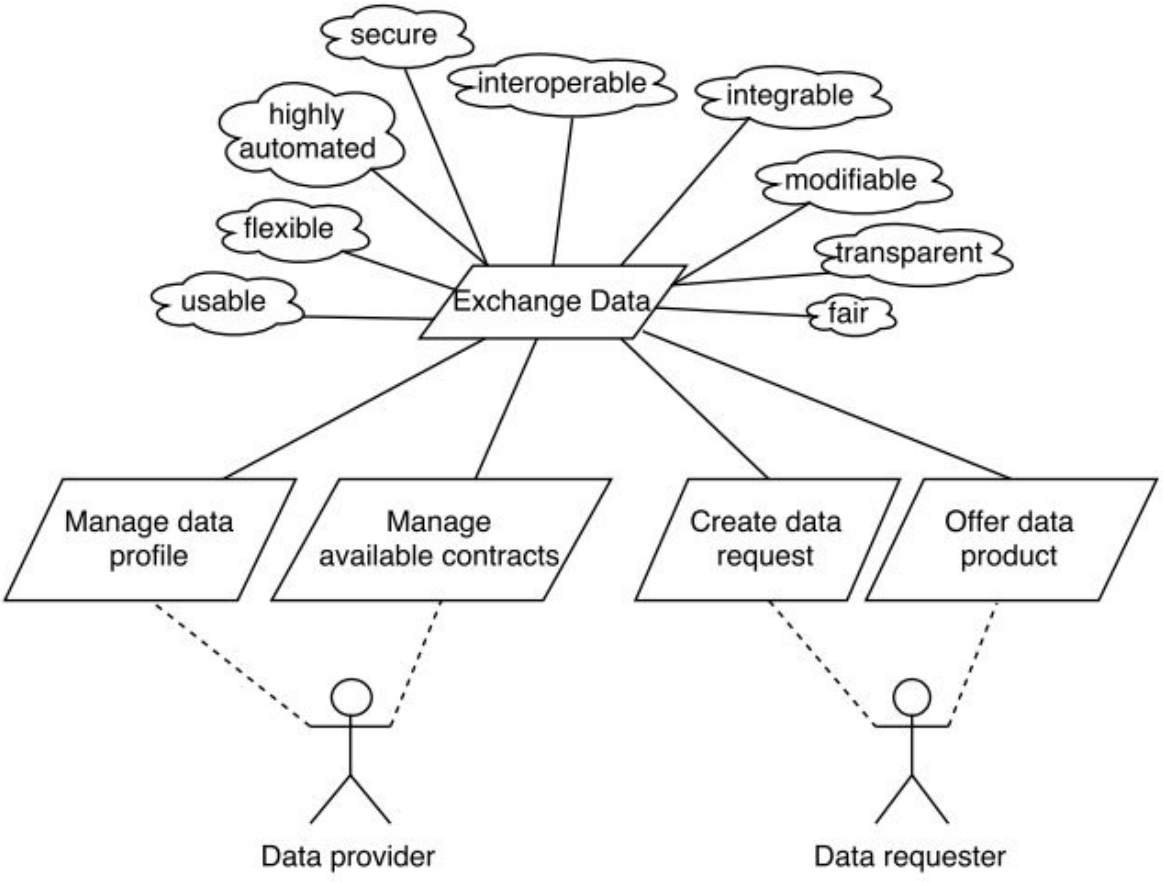

Figura 4.Modelo de objetivos para el ecosistema de la billetera de datos.

- ***Interoperable*** significa que el ecosistema central, debe poder interactuar y componerse con componentes externos, como fuentes de datos y aplicaciones.

Este es un objetivo desafiante dada la heterogeneidad y el dinamismo de las interfaces externas para las fuentes de datos y las aplicaciones existentes. Esto también indica que el ecosistema **DataWallet** debe resistir intentos de uso no autorizados y ataques de denegación de servicio al tiempo que proporciona servicios a los proveedores de datos y solicitantes de confianza.

El intercambio flexible de datos, muestra la capacidad del sistema para integrar datos heterogéneos y diversos y productos de datos con una interfaz uniforme.

- ***Utilizable*** significa que el ecosistema **DataWallet** debe ser claro y accesible tanto para los proveedores de datos como para los solicitantes de datos. La prevención de errores se refiere a la anticipación proactiva y la prevención de errores de colaboración que ocurren comúnmente. El manejo de errores asegura el soporte del sistema para el registro y la gestión de errores de software en el ecosistema.

La capacidad de aprendizaje se refiere a minimizar el tiempo que les toma a los usuarios dominar el ecosistema **DataWallet**.

El intercambio transparente de datos, se relaciona con garantizar que tanto el proveedor de datos como el solicitante tengan, ambos, una comprensión clara de los elementos puntuales, la compensación y los servicios que se intercambian, y que cada uno se encuentre alejado de la reputación previa de los demás.

Finalmente, la equidad del ecosistema **DataWallet** significa que las partes interesadas son tratadas sin favoritismo o discriminación y que todos los intercambios están claramente definidos y explícitamente aceptados.

Además, existen objetivos de calidad que no son discernibles durante una interacción particular con el ecosistema **DataWallet**.

- ***Integrable*** indica que todos los componentes deben ajustarse a una interfaz específica para facilitar la interacción de los componentes funcionalmente contenidos.

- ***Modificable*** significa que el sistema debe adaptarse durante su ciclo de vida al contexto de la aplicación, por ejemplo, para adaptarse a los nuevos estándares de formato de datos.

3.2 Requisitos del proveedor de datos

La jerarquía de la *Figura 5* representa los requisitos básicos del proveedor de datos que gestionan el perfil de datos y los contratos disponibles.

La gestión del perfil de datos, requiere garantías de que los datos son seguros y privados, lo que requiere una gestión y almacenamiento de clave privada, que se utiliza para encriptar los datos.
El proveedor debe controlar qué datos se incluyen en el perfil y qué campos están expuestos al intercambio de datos en ese momento. La recopilación de datos del perfil requiere la vinculación de las fuentes de datos, posiblemente a través de claves de acceso **API** que deben almacenarse. Este perfil de datos intercalados debe actualizarse de acuerdo con las directivas de datos del proveedor.

La segunda jerarquía de refinación de la *Figura 5* ilustra los requisitos de los proveedores de datos para contratar y gestionar los datos disponibles y los contratos de productos.

La funcionalidad principal deseada es encontrar contratos disponibles de acuerdo con las directivas de datos del proveedor[15]. Los proveedores requieren acceso a diferentes tipos de contratos:

- Simples solicitudes de datos como las publicadas por **dx-Insights**,

[15] dependiendo del balance óptimo entre privacidad y eficiencia, se pueden optar por algunos campos demográficos estén disponibles para la búsqueda de contratos para agilizar el proceso.

- Productos de datos novedosos o
- experiencia aumentada basada en datos personales de productos existentes.

Cada uno de estos tres tipos de contratos puede estar dirigido/restringido a una población particular de proveedores que necesita ser administrada automáticamente por el proveedor.

Una vez que se accede, el proveedor debe poder aceptar o rechazar contratos.

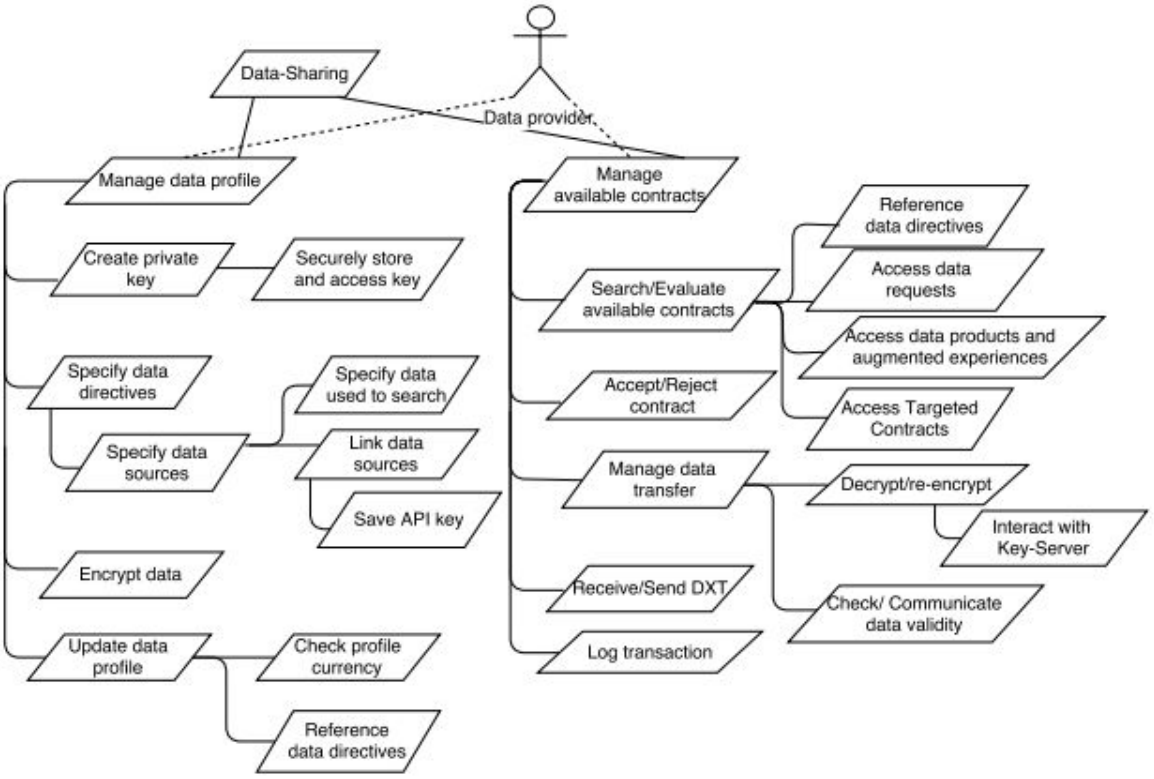

Figura 5: Gestión del perfil del proveedor de datos y objetivos de compromiso del contrato.

Para cumplir un contrato aceptado, los proveedores deben poder enviar los elementos puntuales acordados, lo que requiere aislar esos elementos puntuales, descifrarlos, verificar y comunicar la validez de los datos para garantizar la fidelidad, por ejemplo, a través de una coincidencia de *'datos hash*', y idealmente volver a cifrarlos con la clave pública del solicitante para garantizar la inviolabilidad de la transferencia.

La gestión de contratos también necesita interactuar con un servicio de custodia para enviar y recibir **DXT** después de que se haya completado el intercambio. Finalmente, todas las transacciones deben registrarse en un registro persistente y de fácil acceso.

3.3 Requisitos del solicitante de datos

El refinamiento del modelo de objetivos, para el solicitante de datos en la *Figura 6* muestra sus dos requisitos principales.

En primer lugar, requieren poder crear una solicitud de datos a cambio de una compensación (**DXT**).

En segundo lugar, desean poder ofrecer un producto de datos que requiera elementos puntuales específicos y, por lo tanto, gran parte de la misma funcionalidad de una simple solicitud de datos. Por lo tanto, empleamos el símbolo de herencia múltiple para resaltar los objetivos superpuestos, y usamos objetivos coloreados para indicar donde divergen los requisitos (r*ompiendo la herencia múltiple; consulte la Sección 2.1 para la discusión*).

Como muestra la *Figura 6*, tanto las solicitudes de datos como las ofertas de productos, necesitan la capacidad de especificar los elementos puntuales requeridos y, opcionalmente, especificar una población particular para la cual el contrato es válido a través de parámetros demográficos o mediante un código de acceso/código tipo **QR**.

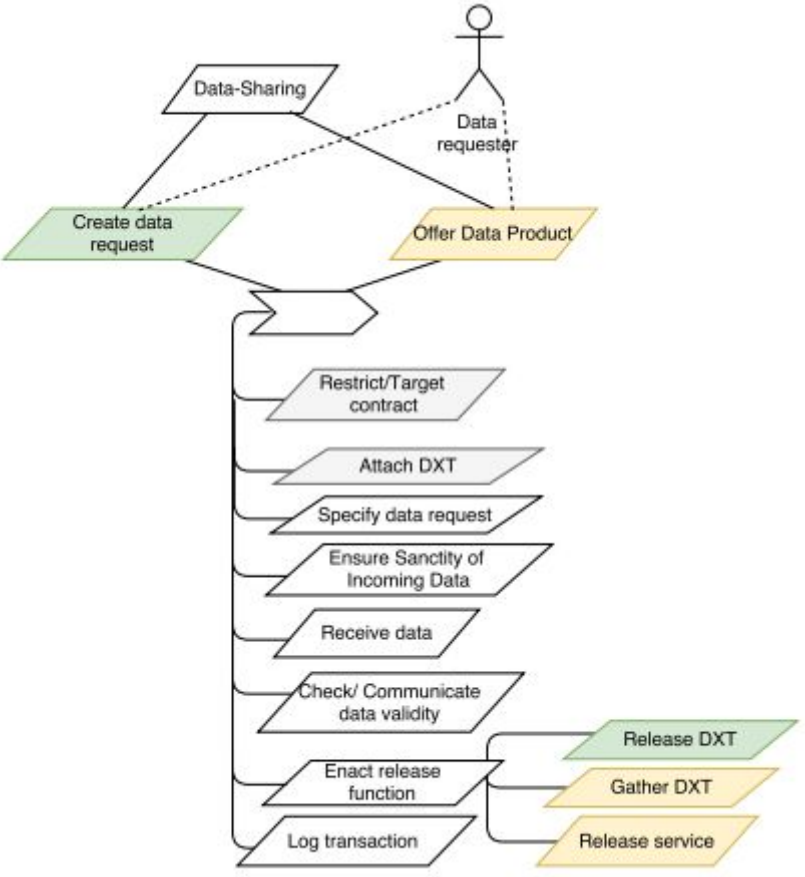

Figura 6. Objetivos del solicitante de datos . Los sombreados en gris son opcionales, y los objetivos de color dependen del principal precedente. (Esto rompe la herencia múltiple).

También necesitan la capacidad de garantizar que los datos que reciben no se puedan interpretar a lo largo del proceso de transferencia, lo cual se logra al enviar una clave pública al proveedor de datos, para encriptar los datos antes de la transferencia, de una manera que solo se puede descifrar con la clave privada del solicitante. Los datos que reciben deben poder verificarse, por ejemplo, mediante una coincidencia hash.

Los datos válidos deben promulgar la función de liberación especificada. Para la solicitud de datos, es el lanzamiento de **DXT** asignado en custodia.

Para un producto de datos, esto requiere el lanzamiento del servicio prometido.

Finalmente, todas las transacciones deben registrarse en un registro persistente y de fácil acceso. Habiendo especificado los requisitos del solicitante de datos y el proveedor en términos de modelos de objetivos de **AOM**, pasamos ahora a la arquitectura que satisface los requisitos.

4 Arquitectura de intercambio de datos

Resumimos la arquitectura central de un sistema que satisface los requisitos, tanto para los proveedores de datos como para los solicitantes. La *Figura 7* muestra el diagrama general de componentes del ecosistema **DataWallet**. Tenga en cuenta que hay tres partes en las figuras que agrupan los componentes para el proveedor de datos:

- **El solicitante de datos** y los
- **componentes de intercambio de datos**, respectivamente.

Cada parte contiene un componente de administrador central y crucial:

- El **Administrador de perfiles de datos**,
- El **Gestor de contratos inteligentes** y
- El **Administrador de solicitudes**, cuyas funcionalidades múltiples también se representan.

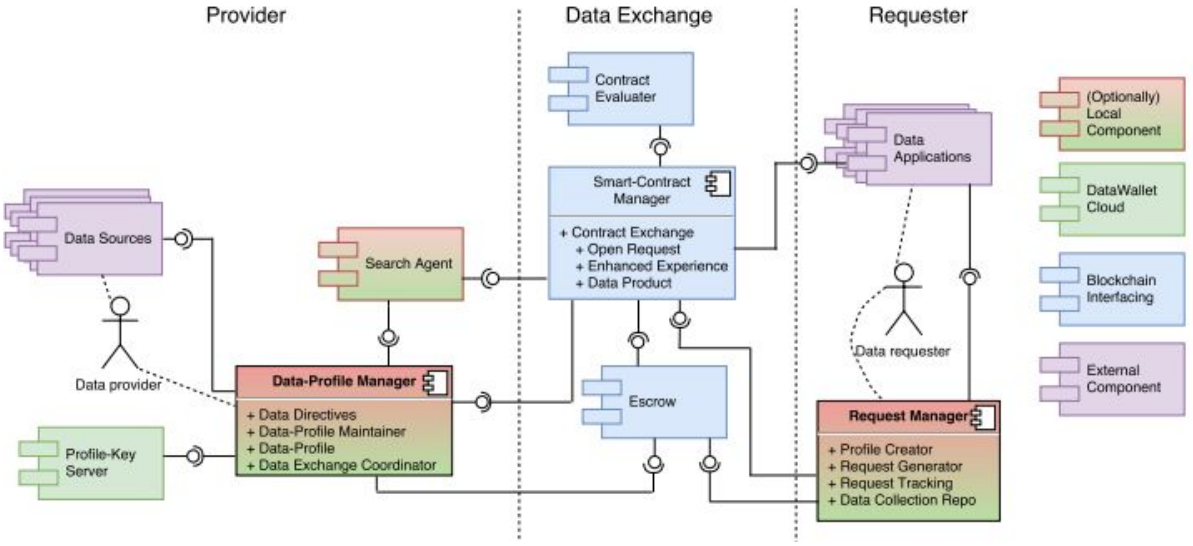

Figura 7: diagrama UML de componentes del ecosistema de billetera de datos

Describimos cada parte de la arquitectura, a su vez la *Sección 4.1* describe los componentes que soportan el solicitante de datos. La *Sección 4.2* detalla los componentes correspondientes del proveedor de datos y finalmente, la *Sección 4.3* detalla los componentes de intercambio de datos que emplea blockchain y que media los datos y el intercambio de servicios entre proveedores y solicitantes.

4.1 Componentes Solicitantes de Datos

El componente central con el que el solicitante de datos se relaciona directamente, es el Administrador de solicitudes y proporciona dos interfaces:

- Una API para solicitudes programáticas de aplicaciones de datos y usuarios expertos y un
- Sistema basado en la web fácil de usar.

El Administrador de solicitudes permite dos comportamientos principales.

En primer lugar, permite la creación de una nueva solicitud de datos y, en **segundo lugar**, permite consultar el estado de las solicitudes existentes y acceder a los datos resultantes.

El *Administrador de solicitudes* primero guía a los nuevos solicitantes a través de la creación de un perfil y al menos un par de claves públicas y privadas para facilitar la transferencia de datos encriptados.

El *Administrador de solicitudes* luego permite a los solicitantes crear solicitudes de datos rellenando una 'plantilla o formulario' de contrato de datos (plantilla o formulario **dx-contrac**t), especificando la información en la *Tabla 1*.

- Una interfaz abstracta, una *'API de datos*', para recopilar elementos puntuales que abordan casos de uso comunes, proporciona una alternativa a la especificación del contrato manual. **Data API** se basa en un conjunto de contratos preparados desarrollados y actualizados por **DataWallet** y la comunidad para casos de uso común.

Para convertir la plantilla o formulario, en un contrato completo de intercambio de datos y publicar el contrato **dx-Insights,** que muestra la *Tabla 1*, el intercambio de contratos requiere que el solicitante asigne el **DXT** especificado en el depósito de garantía y pegue la clave pública de cifrado al contrato.

Estas solicitudes de datos se colocan en la sección "*Solicitud abierta"* del Intercambio de contratos. Por ejemplo, la aplicación **dx-Insights**, el primer producto en **Data Application Exchange**, genera solicitudes de datos para empresas que buscan información sobre clientes y mercados basada en datos de redes sociales recopiladas de poblaciones específicas de proveedores, todas las cuales se pueden ver en el contrato abierto..

		Tipo de Contrato	
		Solicitud de datos	Producto de datos
Campos	OBLIGATORIOS	Elementos puntuales Requeridos (E)	
		que dato es usado para:	
		Compensaciones en DXT	Precio del Producto (DXT
		Cantidad de proveedores de datos	categoría del producto (E)
		Umbral de rodaje o adopción	Descripción del producto (enlaces descriptivos externos opcionales)
	OPCIONALES	parámetros demográficos del proveedor (es) (E)	
		Fecha de vencimiento para la expiración de la solicitud	
		Formato de datos o preferencias de preprocesamiento (E)	
		Puerto de enrutamiento de datos (si no es el administrador de datos de Solicitud estándar).	
	Ejecución de la acción	Asignar el DXT especificado al fideicomiso (y un porcentaje adicional para respaldar el ecosistema activo)	
		Adjuntar una llave de encriptación pública	

Tabla 1: Lista de los campos de plantilla para un contrato **dx-Insights**. La primera columna especifica todos los campos asociados con una solicitud de datos, mientras que la segunda columna indica todos los campos asociados con una oferta de datos y productos. **[E]** indica un campo de enumeración, es decir, un campo elegido de una lista de opciones previamente especificada.

Además de la sección de Solicitud abierta del Intercambio de contratos, está la sección de Productos de datos, cuyos contratos se crean de manera similar. Como se ve en la *Tabla 1*, comparte algunos de los mismos campos que una solicitud de datos, pero requiere la especificación de campos adicionales, como el precio del producto (en **DXT**) y una especificación de cómo se entrega la aplicación de datos en la ejecución del contrato.

El mecanismo de entrega de la aplicación puede ser una lógica interna del contrato que, por ejemplo, proporciona un código de acceso y un enlace al proveedor de datos, o una descripción de la entrega fuera de la cadena de acceso al producto. Por ejemplo,el **DataWallet** proporciona un producto de análisis de personalidad que se basa en los mensajes de un proveedor en las redes sociales.
El contrato en la sección de producto de Datos, del Intercambio de Contrato, indica los elementos puentuales requeridos, el precio del producto en **DXT** y el mecanismo de entrega. Este último es un localizador uniforme de recursos (**URL**) transferido al Administrador de perfiles del proveedor que está conectado al perfil de personalidad con una contraseña inicial implícita, que puede modificarse posteriormente.

A través del **Administrador de solicitudes**, los solicitantes también pueden administrar los contratos que han realizado en el Intercambio de contratos. Para los contratos *'renovables'* que permiten a los proveedores ejecutar el contrato de inmediato, los solicitantes pueden ver el número actual de contratos ejecutados y los datos resultantes *(si se utiliza Repo de recopilación de datos del administrador de solicitudes*). Para los contratos con umbral, los solicitantes pueden ver si se ha ejecutado (o establecer una alerta cuando el umbral o un número arbitrario de proveedores han señalado el contrato).

4.2 Componentes del proveedor de datos

Los proveedores de datos interactúan con el intercambio de datos, a través de su administrador de perfiles de datos que tiene dos responsabilidades principales. En primer lugar, construye y mantiene el perfil de datos del proveedor, ya sea a través de un servicio basado en la nube o en el sistema local del proveedor.

En segundo lugar, permite al proveedor utilizar este perfil y comprometerse con contratos en el Administrador de intercambio de datos.

El Administrador de perfiles de datos construye y mantiene el perfil de datos del proveedor de acuerdo con sus Directivas de datos. Las **Directivas de datos** del proveedor determinan qué datos se pueden recuperar de orígenes externos para personalizar el perfil y el comportamiento del agente de búsqueda.

Por ejemplo, un proveedor comparte “me gusta” y “tiempos de compromiso de una empresa de medios sociales”, pero no sus mensajes privados. Además, un proveedor de datos puede especificar que desea que su respectivo perfil de datos se actualice a intervalos regulares especificados, o bajo demanda. También podrían personalizar su agente de búsqueda para tener acceso a su edad e información geográfica, con el objetivo de buscar contratos viables más rápidamente.

Dadas las directivas de datos, el Administrador del perfil de datos se coordina con el Solicitante y el Servidor de clave de perfil para construir y mantener el perfil de datos del proveedor. En primer lugar, el Administrador de perfiles de datos guía a los proveedores a través de la generación de una clave de cifrado de perfil, que luego se divide con la mitad "*remota*" de la clave almacenada en el Servidor de clave de perfil, para garantizar que no haya un solo punto de falla.

Data-Profile Manager, permite a los proveedores aprovechar sus datos en el intercambio de contratos a través de la coordinación de, los perfiles de los proveedores con el servidor **ProfileKey**, **Smart-Contract Manager**, **Escrow** y el agente de búsqueda del proveedor.

Como se describe en la Sección 5.4, el Agente de búsqueda encuentra dx-contracts que son consistentes con el perfil de datos y las directivas del solicitante. Por ejemplo, el agente de búsqueda devuelve contratos cumplidos publicados por la aplicación **dx-Insights** a un proveedor que busca contratos en la bolsa abierta para adquirir datos DXT.

Para ejecutar el contrato, el administrador de perfil de datos debe aislar los elementos puntuales especificados, descifrarlos con la combinación de claves locales combinadas y Servidor de claves de perfil, formatee los datos como se especifica en el contrato, vuelva a cifrar los datos con la clave pública del solicitante y envíe los datos resultantes al Administrador de contratos inteligentes donde los datos se enrutan a Repo de recopilación de datos del Administrador de solicitudes, o una aplicación de datos externa, o servidor,

como se especifica en el contrato. La ejecución de este contrato se describe en detalle a continuación en la *Sección 5.4*.

4.3 Componentes de intercambio de datos

Como se muestra en la *Figura 7,* el componente central que sustenta el intercambio de datos es el **Smart-Contract Manager** coordina la ejecución y la interacción blockchain de dx-contracts.

Tenga en cuenta que gran parte de la funcionalidad descrita en esta sección como la procedencia del componente Smart-Contract Manager es ejecutada por el contrato inteligente en sí (consulte la Sección 5.1 para conocer la dinámica de la interacción blockchain durante el ciclo de vida de un dx-contract).

Esto requiere la coordinación de los componentes de **blockchain**, Escrow y **Contract-Evaluater** para garantizar la ejecución conforme de los contratos de dx, así como los componentes de **Data-Profile Manager**, **Search Agent, Request Manager** y **Data Applications** del proveedor. Smart-Contract Manager aloja el intercambio dx-contract que internamente se divide en dos categorías:

- El intercambio de solicitud de datos y
- El intercambio de productos de datos.

Cada uno tiene una sección pública 'abierta' y una sección 'dirigida' que solo pueden ver los proveedores con los derechos apropiados (*consulte la Sección 5.5 para obtener detalles sobre esta distinción*).

Smart-Contract Manager proporciona una interfaz al Agente de búsqueda del proveedor de datos que permite el acceso a contratos compatibles que cumplen con las directivas del proveedor y coordina con el Agente de búsqueda y el Perfil de datos del proveedor para verificar que este último cumpla con las especificaciones del solicitante (ver la *Sección 5.4* para los detalles de este proceso).

El Evaluador de Contratos necesita verificar varios aspectos elementales de los contratos inteligentes, tales como:

- La solidez del flujo de control,
- El flujo de datos,
- La asignación de recursos,
- La excepción y
- La gestión de compensación [20], [21].

Por otro lado, Smart-Contract Manager también coordina la promulgación de contratos, por ejemplo, verificar que los datos transferidos sean correctos y luego señalar el lanzamiento del **DXT** al monedero de datos del proveedor. **Smart-Contract Manager** sigue un ciclo de

vida específico [19] que comprende una configuración específica [17], implementación y promulgación [18] y reversión con fase de terminación [22].

5 Dinámica del ecosistema de intercambio de datos

Esta sección describe y formaliza el comportamiento dinámico del ecosistema **DataWallet** descentralizado con diagramas de secuencia UML [27] (de acuerdo con la notación de la *Figura 3b*). Los diagramas de secuencia muestran el intercambio de datos, almacenes de valores (**DXT**) y servicios de datos entre proveedores de datos y solicitantes con los componentes descritos en la Figura 7.

Todos los intercambios descritos a continuación garantizan que :

1. El proveedor de datos sepa exactamente qué datos se intercambian,
2. los datos nunca están disponibles en texto entendible,
3. los datos sólo son recuperables por el solicitante específico que se menciona en el contrato, y

4) el solicitante está obteniendo datos que cumplen con sus especificaciones.

Garantizar estas características requiere agentes de software transparentes y auditables mantenidos por **DataWallet**, contratos inteligentes [23] que garanticen el cumplimiento transparente de todas las partes y el uso juicioso de la criptografía asimétrica entre proveedores de datos, solicitantes de datos y agentes de software mantenidos por DataWallet.

El resto está estructurado de la siguiente manera. La Sección 5.1 discute las transacciones de blockchain que consideramos para lograr la trazabilidad de eventos. La sección 5.2 describe la creación y gestión del perfil de datos por parte de los proveedores de datos. La sección 5.3 muestra una el solicitante de datos crea y publica el contrato inteligente de solicitud de datos que llamamos **dx-contract**.

La Sección 5.4 detalla una transacción de datos simple como las realizadas por **dx-Insights.**

las pequeñas modificaciones del comportamiento dinámico básico requeridas para transacciones de datos simples permiten intercambios de datos y productos más sofisticados. Esta funcionalidad más amplia se muestra en la Sección 5.5, que detalla cómo pueden intercambiarse los productos de datos diseñados para los proveedores como usuarios finales, y finalmente, la Sección 5.6, que describe cómo un proveedor de datos puede proporcionar datos para aumentar uno de sus servicios existentes.

5.1 Operaciones en el Blockchain

Almacenar eventos en un blockchain es costoso en términos de poder de cómputo y tarifas de transacción. En el caso de Ethereum, **PoW** (resolución de enigmas criptográficos para la validación de transacciones) demuestra ser un cuello de botella de rendimiento y

escalabilidad. Por lo tanto, es prudente definir el conjunto mínimo de transacciones que permita la rastreabilidad requerida en el ecosistema **DataWallet**.

Role				
rq	pr	cm	number	operation
x	x		1	DataWallet ID created (uid)
x			2	Data Sourced
		x	3	Smart-contract deployed
		x	4	DXT in escrow
	x		5	Data sent
x		x	6	Data received
		x	7	Data checked
x	x		8	DXT received
		x	9	Smart-contract completed

Tabla 2: Transacciones de Blockchain para el ecosistema DataWallet.

La *Tabla 2* enumera las operaciones con las columnas rq, pr y cm que indican los roles de Solicitante de datos y Proveedor de datos, Smart-Contract Manager, respectivamente.

La siguiente columna proporciona una ID de operación a la que nos referiremos en el resto de esta sección para aclarar cómo se usan las operaciones de cadena de bloques.

La columna de la derecha etiqueta brevemente las operaciones. Tenga en cuenta que mantenemos el conjunto de operaciones blockchain mínimas para limitar los costos computacionales y transaccionales.

El conjunto mínimo de operaciones se elige para optimizar el panorama de una utilidad que abarca el costo, el rendimiento, el seguimiento DXT y los cálculos de latencia con respecto a la tecnología actual[16].

Ahora proporcionamos una descripción más completa de cada operación indicada en la *Tabla 2.*

La Operación 1 almacena el id. Único (uid) del proveedor de datos recién creado y las cuentas del solicitante.

La operación 2 se registra cuando una fuente de datos está integrada en el perfil de un proveedor. Tanto el tiempo como la identificación de la fuente están grabados.

[16] a medida que la tecnología mejore, el punto óptimo sobre la compensación de transparencia / eficiencia cambiará

La Operación 3 guarda cada contrato colocado en el intercambio.
La Operación 4 registra el depósito de DXT en el depósito de garantía asociado con el contrato.
Las operaciones 5, 6 y 7 indican cuándo los datos son enviados por el proveedor, recibidos por el solicitante y verificados entre sí para garantizar la fidelidad. La Operación 8 almacena el evento cuando el proveedor ha recibido DXT.
Finalmente, la Operación 9 registra la promulgación y finalización exitosas de un contrato inteligente.

5.2 Crear perfil de datos

Los proveedores de datos crean un perfil de datos local, a través de una aplicación suministrada por **DataWallet,** que está hospedada opcionalmente localmente con el proveedor de datos[17]. La *Figura 8* muestra el protocolo de intercambio de mensajes para la creación de perfiles de datos entre los agentes responsables.

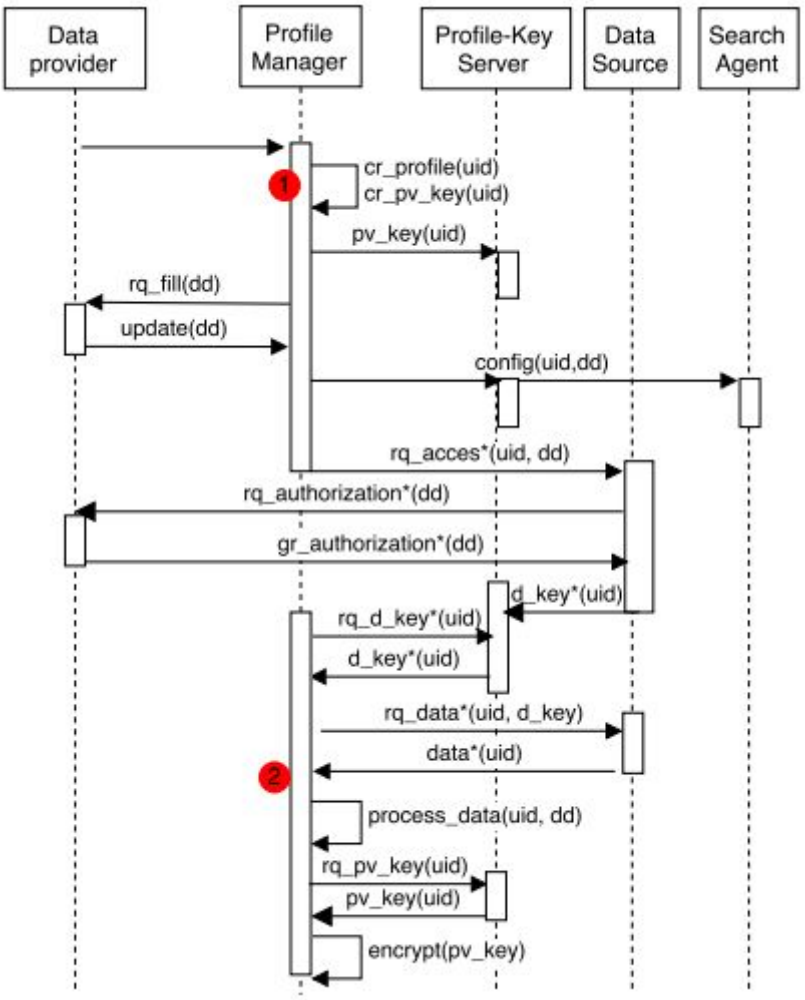

Figura 8: creación del perfil de datos en un diagrama de secuencia UML.

[17] Los detalles de esta sección siguen a los proveedores que optan por un perfil de datos hospedado localmente, pero los proveedores de datos que prefieren utilizar el alojamiento de perfil en la nube de DataWallet también pueden interactuar con el intercambio.

Tenga en cuenta que las etiquetas de mensaje en la *Figura 8* y todos los demás diagramas de secuencia de la *Sección 5* representan pseudocódigos que usamos para facilitar las explicaciones del protocolo.

La creación de un perfil es iniciada por el proveedor de datos que interactúa con su Administrador de perfil de datos. El Administrador ordena la creación del perfil en función de la identificación única del usuario (uid) **cr_profile** (uid) y la creación de una clave de usuario privada **cr_pv_key** (usuario). Luego divide la clave. Transmite la mitad de la clave al servidor de clave de perfil **pv_key** (uid), y almacena la otra mitad localmente.

pNext, **Data-Profile Manager** solicita **rq_fill** (dd) que el proveedor complete la actualización (dd) de sus directivas de datos para controlar/personalizar su perfil de datos y el comportamiento del agente de búsqueda. Las directivas relevantes se transmiten al servidor de clave de perfil y al agente de búsqueda.

Por ejemplo, los tipos/categorías de solicitantes que el proveedor desea considerar (ya sea por exclusión a través de una lista negra o por inclusión a través de una lista blanca), se transfieren al agente de búsqueda. A continuación, el administrador solicita acceso a los puntos de datos que el proveedor especificó desde cada origen de datos externo **rq_acces** * (dd) [18]

Para la mayoría de los orígenes de datos externos, esto requiere que el proveedor autorice explícitamente el intercambio de los puntos de datos especificados **rq_authorization** (dd). Cuando el proveedor concede esta autorización **gr_authorization** (dd), las claves de datos resultantes (en algunos casos, las claves API que proporcionan acceso programático a los datos) se transfieren al servidor de claves de perfil.

Para construir luego el perfil de datos intercalados del proveedor, el Administrador de perfiles de datos solicita los datos para cada fuente de datos especificada en las Directivas de datos del proveedor.

Esto requiere solicitar la clave de cada fuente de datos **rq_d_key** * (uid) y recibirla **d_key** * (uid). Con la clave de datos, el administrador solicita los datos del origen de datos externo **rq_data** * (uid, d_key). Después de recibir todos los datos de datos * (uid) y cotejarlo **process_data** (uid, dd), el administrador encripta el cifrado de perfil resultante (**pv_key**), que requiere solicitar y recibir la mitad remota de la clave privada **rq_pv_key** (uid) / **pv_key** (uid).

Finalmente, la finalización del perfil se almacena en la cadena de bloques a la Operación 1 de la *Tabla 2*. Específicamente, la marca de tiempo y el uid del perfil se almacenan (como se visualizó con el círculo rojo 1 visto en la *Figura 8*).

[18] Tenga en cuenta que las funciones que son repetibles se indican con un asterisco, p. fn * (param). En esta instancia, se solicitará acceso para cada fuente de datos especificada en las directivas de datos del proveedor.

5.3 Creando el dx-Contract

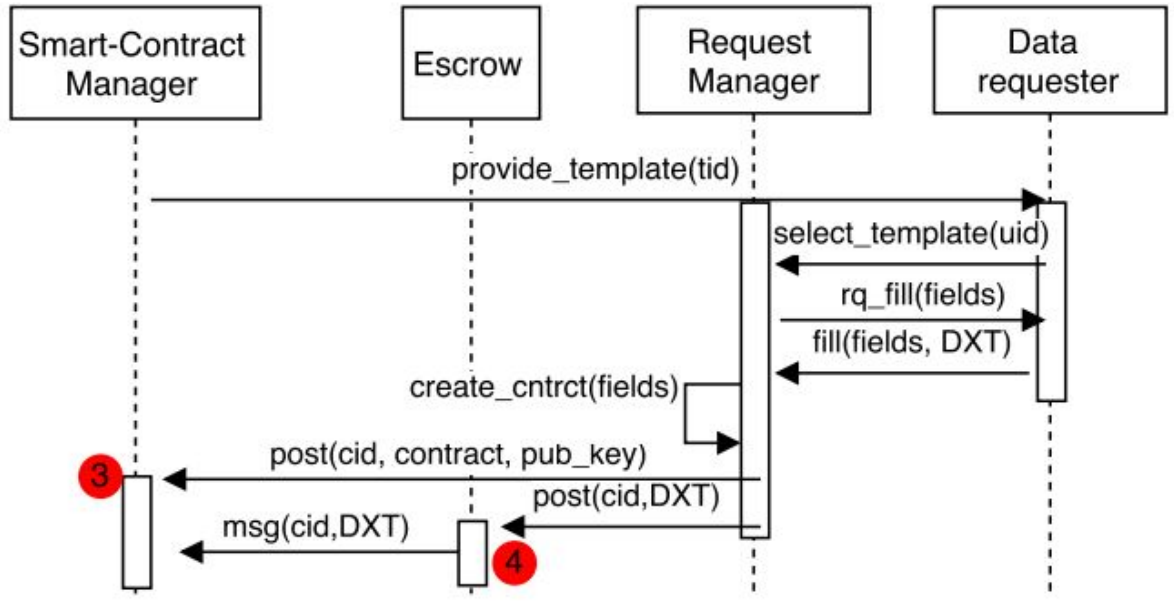

*Figura 9: diagrama de secuencia UML de la creación de solicitud de datos, o **dx-Contract***

Los solicitantes de datos pueden publicar contratos en el intercambio de datos abierto a través de su A**dministrador de solicitudes**, que coordina con el **Administrador de contratos inteligentes** y el **Fideicomiso** como se describe en la *Sección 4.1*.

Como se ve en la *Figura 9,* para crear un contrato, el solicitante primero selecciona una plantilla **select_template** (uid) del grupo disponible en Smart-Contract Manager. El Administrador de solicitudes asegura que todos los campos requeridos (consulte la *Tabla 1*) se completen **rq_fill** (campos). El Administrador de solicitudes luego crea el contrato **create_cntrct** (fields) y lo envía al **Smart-Contract Manager** junto con el id del contrato (cid) y la clave pública del proveedor (**pub_key**) post (cid, contract, pub_key).

Esta publicación del contrato se guarda en la cadena de bloques (*Operación 3*).

El Administrador de solicitudes también publica el DXT necesario en la cuenta de depósito de garantía asociada con la publicación de identificación del contrato (cid, DXT). Esta asignación de DXT se registra en el blockchain (*Operación 4*) y se comunica al contrato en el Administrador de contratos inteligentes.

5.4 Intercambio de datos simple

Una vez que un contrato de dx (vea la *Sección 4.1*) se coloca en el intercambio abierto, es visible por el agente de búsqueda de los proveedores de datos. En la *Figura 10*, describimos el proceso por el cual un proveedor encuentra y ejecuta una solicitud de datos. Si bien es probable que los usuarios prefieran que el agente se ejecute en modo **por lotes**, examinamos la obtención secuencial y la evaluación de los contratos para mayor claridad.

Para iniciar una búsqueda de una solicitud de datos simple, en la central abierta, un proveedor de datos ordena a su agente de búsqueda que inspeccione los contratos publicados en la cuenta abierta **rq_open_contracts** (uid). Para que una solicitud sea viable para un proveedor de datos, el proveedor debe poseer todos los puntos de datos especificados y ajustarse a la información demográfica especificada (opcionalmente).

Como se ve en la *Figura 8,* los proveedores pueden, a través de sus directivas de datos, otorgar al agente de búsqueda información relevante para garantizar que todos los contratos devueltos satisfagan ambas condiciones.

Concederle al agente de búsqueda los datos relevantes le permite realizar comprobaciones de viabilidad de contratos en el entorno de gestión de contratos optimizado. La *figura 10* supone, por lo tanto, que el agente de búsqueda tiene toda la información necesaria. De lo contrario, se requieren comunicación y coordinación adicionales entre el agente de búsqueda, el contrato dx, el administrador de perfiles del proveedor y el servidor de claves de perfil para descifrar los puntos de datos relevantes y luego asegurarse de que se ajusten a las especificaciones del contrato dx.

Un ejemplo de este proceso es dx-Insights publicando un contrato a la *Sección 4.1* y *Tabla 1.* Por ejemplo, dx-Insights podría estar buscando los programas de televisión y artistas musicales favoritos de personas entre 25 y 33 años que viven en ciudades en el Costa este. Un proveedor de datos que estaba buscando intercambiar algunos datos para DXT podría encontrar rápidamente dichos contratos que sean viables para ellos si le han dado a su agente de búsqueda la información relevante. En este caso, esa sería la lista de campos disponibles en su perfil de datos, que incluyen programas favoritos de televisión y músicos, así como los valores de los parámetros demográficos clave, como la edad y la ubicación.

Una vez que se ha determinado que una solicitud es viable, el proveedor de datos tiene la opción de dar su consentimiento al contrato (contratos). Si el proveedor acepta el contrato, su administrador de perfiles selecciona y descifra los datos acordados.

Esto se logra solicitando la clave de perfil privada **rq_pv_key** (uid) desde el servidor **ProfileKey** y combinándola con la mitad localmente almacenada. A continuación, los datos se someten a hash y se envían al hash de datos de contrato (cid) como referencia para la fidelidad de la transferencia de datos.

El Administrador de perfiles luego encripta los datos con el cifrado de clave pública del solicitante (**req_pub_key**) (que se proporciona con el contrato). En este punto, si el contrato lo requiere, el proveedor apuesta DXT en el Depósito en depósito + (dxt). Finalmente, los datos cifrados se envían al puerto especificado del Administrador de solicitudes alojado u otro sistema compatible.

Una vez que el Administrador de solicitudes recibe los datos, los descifra con el descifrado de la clave privada correspondiente (**req_pv_key**)[19]. Luego envía un hash de los datos descifrados al hash de datos del Administrador de contratos (cid).

Esto hace que el Gerente de Contrato emplee el Evaluador de Contrato para garantizar que los hashes de datos enviados y recibidos coincidan y que se cumplan todas las demás condiciones del contrato. Si el contrato se cumple resultado (cid), el Administrador de contratos instruye a Escrow para que libere la versión de DXT especificada (uid, cid) al proveedor de datos. Finalmente, la transacción se agrega a los registros de los proveedores y de los solicitantes.

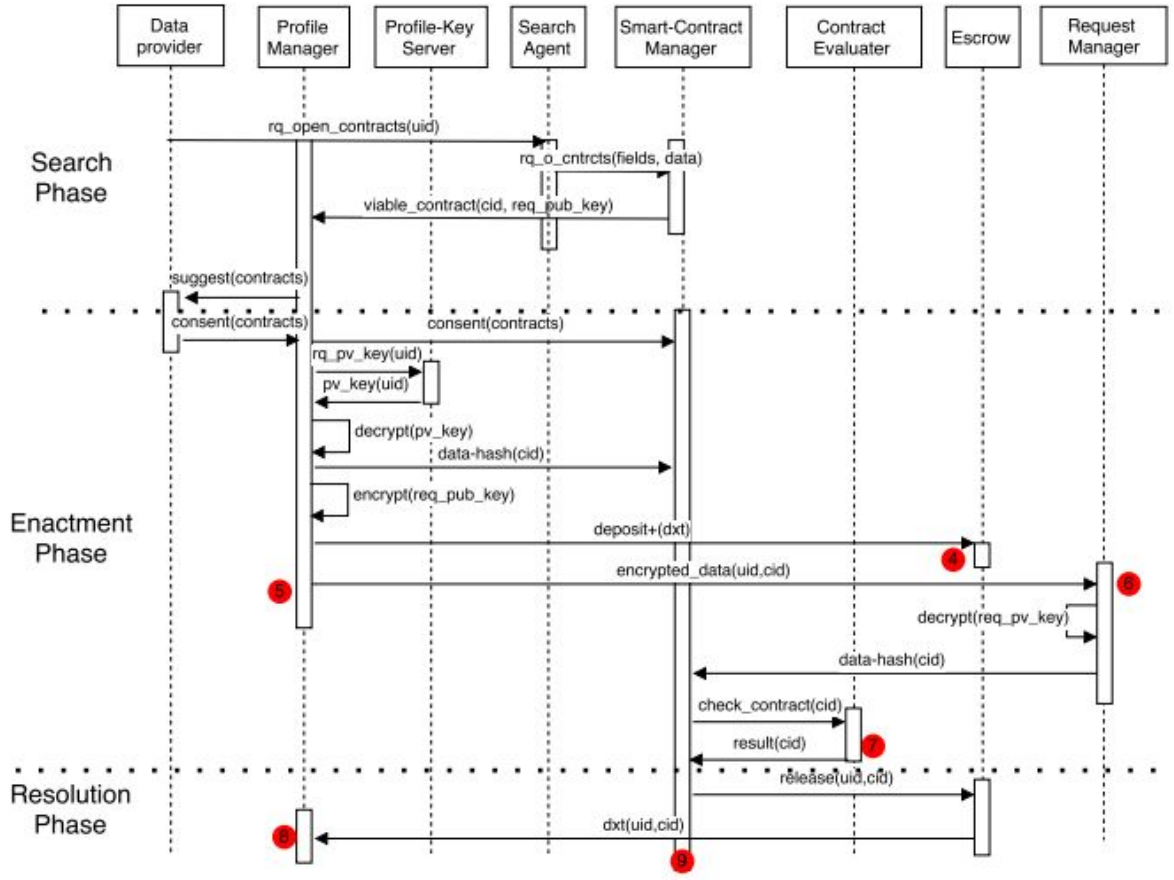

Figura 10: diagrama de secuencia UML de la ejecución de una solicitud de datos simple.

Las transacciones en el blockchain se indican en la *Figura 10* con los puntos rojos numerados que corresponden a las ID de transacción de la *Tabla 2*.

Por lo tanto, cuando el proveedor apuesta **DXT** (si es necesario), este se almacena en la 4. También se registra cuando los datos se envían desde el Administrador de perfiles del proveedor (*Operación 5)* y cuando el Administrador de solicitudes los recibe (*Operación 6*) .

[19] El solicitante de datos (y solo el solicitante de datos) puede descifrar los datos que reciben dado el esquema de criptografía asimétrica empleado.

La *Operación 7* registra cuando el Evaluador de Contratos ha validado el intercambio de datos y otros términos del contrato antes de la promulgación del contrato por parte de **Smart-Contract Manager.**

La *operación 8* indica cuando el proveedor recibe el **DXT**. Finalmente, la *Operación 9* termina la promulgando del contrato.

Las siguientes dos secciones describen cómo alteraciones menores a la dinámica de intercambio descritas aquí, pueden permitir una gran cantidad de intercambios de datos y productos.

5.5 Data-Product Exchange

En lugar de solicitar datos para ofrecer un producto externo como **dx-Insights** o realizar investigaciones, algunas aplicaciones en **App Exchange** son servicios para los mismos proveedores de datos. La dinámica de dicho intercambio de productos de datos comparte mucho con el intercambio previo de datos puntuales de la *Figura 10*, pero opcionalmente requiere que el productor de datos intercambie tokens **DXT** además de sus datos, y debe garantizar la entrega del servicio.

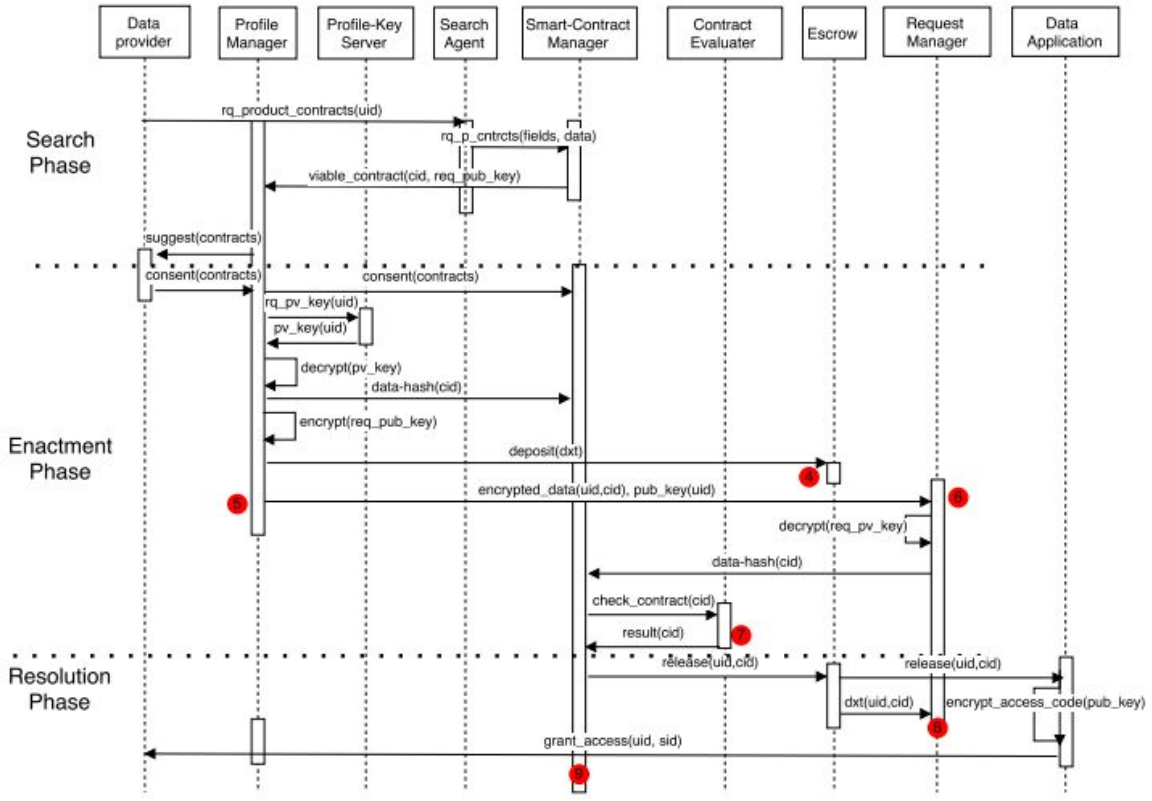

Figura 11: Experiencia reflexiva del monedero de datos en un diagrama de secuencia UML.

Considere en el ejemplo de la *Figura 11* a un desarrollador que ha diseñado una aplicación que visita sitios web oscuros para ver si los correos electrónicos y contraseñas de un usuario se han visto comprometidos.

Deben especificar un contrato como se describió anteriormente y, como se describe en la *Sección 4.1* y *Tabla 1*, adicionalmente deben establecer la categoría de producto de una lista enumerada, describir de manera concisa el producto en el contrato, opcionalmente referenciando recursos adicionales a través de un enlace externo, describir el mecanismo de entrega del producto, indicar la cantidad de DXT requerida para el producto y opcionalmente puede utilizar un contrato inteligente para el método de entrega.

Por ejemplo, puede requerirse una verificación de contraseña de acceso del producto para que se libere el token. En el caso de la *Figura 11*, el desarrollador utiliza la clave pública del proveedor para encriptar la contraseña y transmitirla al proveedor de datos y, opcionalmente, requiere una validación de hash.
Tenga en cuenta que el solicitante debe enviar una pequeña cantidad de **DXT** para facilitar el ecosistema de administración de contratos.

Después de la finalización del contrato del producto de datos, se carga en el intercambio de productos de datos, donde es visible para los proveedores de datos. Tenga en cuenta que los productos de datos, también pueden estar orientados a proveedores específicos que utilizan, por ejemplo**, códigos qr,** tal como se describe en la siguiente sección.

Por ejemplo, los proveedores que luego buscan en la categoría de seguridad de datos de aplicaciones, ven la aplicación y (*después de los controles de viabilidad discutidos anteriormente*) eligen cambiar por el producto.

El administrador del perfil de datos del proveedor ejecuta la transferencia de datos como se describió anteriormente. Sin embargo, hay pasos adicionales opcionales para colocar el **DXT** especificado en custodia y proporcionar una clave pública para la codificación del código de acceso. Continuando con nuestro ejemplo, el token se colocará en custodia hasta que los datos y el acceso al producto hayan sido verificados (*a través de la comparación de hash en el contrato inteligente*).

5.6 Mejora en el Intercambio de Experiencia

Al dirigirse específicamente a los clientes, los desarrolladores y las empresas pueden aumentar la experiencia de sus usuarios con sus productos.

Por ejemplo, un servicio de transmisión de música puede mejorar su algoritmo de recomendación a través de datos personales del usuario. La dinámica de dicho intercambio de experiencia aumenta el desarrollo del intercambio de servicios de datos anterior, pero requiere derechos especiales de acceso específico al contrato.

Un contrato de mejora en el intercambio de experiencia debe especificar adicionalmente los derechos de acceso.

La manera más simple de restringir el acceso es asumir (*razonablemente*) que solo las personas expresamente invitadas tienen acceso al hash único del contacto aprovechando la "seguridad mediante el anonimato" del intercambio específico. Si es necesario, se pueden tomar medidas adicionales, *por ejemplo*, exigir a los proveedores de datos que suministren un código de acceso.

Este hash se puede comunicar al proveedor de datos a través del **código qr** o la entrada manual en el administrador de perfiles del proveedor.
El agente de búsqueda del proveedor de datos puede acceder a los contratos específicos a los que se ha notificado, y luego de verificar la viabilidad del contrato, le da a los proveedores la opción de consentir y participar en el contrato.

La dinámica de esta interacción es paralela a la del intercambio de servicios de datos, aunque con una complejidad adicional externa al intercambio. El desarrollador de la aplicación de datos debe asegurarse de que la experiencia aumentada se base en los datos del proveedor y se otorgue a su cuenta asociada.

6. Conclusión

Este documento técnico presenta el ecosistema **DataWallet**, una aplicación de intercambio de datos basada en la tecnología blockchain que permite a los productores de datos recuperar los datos que crean en línea de aquellos que los utilizan para su propio beneficio.

Al proporcionar a los productores de datos un control detallado de sus datos, **DataWallet** resuelve varios problemas existentes actualmente, no solo para los productores de datos, sino también para los consumidores de datos.

- En primer lugar, el problema del silo se supera al poder combinar datos de manera determinista en los jardines amurallados de Internet.
- En segundo lugar, la recopilación de estos conjuntos de datos de una manera inteligente a través de contratos inteligentes resuelve el problema de la calidad de los datos.
- En tercer lugar, el problema de ética se resuelve al garantizar que cada elemento de datos fue expresamente consentido para ser compartido por las personas que lo produjeron.

En consecuencia, este documento técnico presentó requisitos que resuelven estos tres problemas para los consumidores de datos, así como la injusticia de monetizar los datos de las personas a sus espaldas. Luego, el documento generó la arquitectura estática del ecosistema **DataWallet** a partir de estos requisitos y posteriormente describió cómo su comportamiento dinámico permite el intercambio de datos, sin intermediarios, utilizando contratos inteligentes basados en cadenas de bloques.

Los requisitos de **DataWallet** muestran que la principal propuesta de valor del ecosistema es facilitar el intercambio de datos seguro y consensuado[20].

En la *Sección 3,* la propuesta de valor se refinó jerárquicamente con sub objetivos funcionales que se asignaron a las partes interesadas, a el proveedor de datos y el solicitante. El primero controla las funciones de **DataWallet** de refinación de gestionar los perfiles de datos y gestionar los contratos disponibles. Este último controla las funciones para crear solicitudes de datos y ofrecer productos de datos.

En la *Sección 4,* derivamos una arquitectura estática para el ecosistema **DataWallet,** a partir de los modelos objetivos, que dio como resultado tres conjuntos coherentes de componentes distribuidos.

El grupo de componentes del proveedor de datos, es un administrador de perfiles de datos que coordina la seguridad mediante un servidor de claves de perfil, la creación y el mantenimiento de un perfil de datos a través de fuentes de datos externas y un agente de búsqueda para interactuar con los intercambios de contratos inteligentes.

El intercambio de datos impulsado por la tecnología blockchain está compuesto por un administrador de contrato inteligente que coordina con un fideicomiso de garantía de DXT y un evaluador de contrato.

[20] Los objetivos de calidad del sistema, deben realizarse en la implementación de la arquitectura, patrones y estilos de software, que están fuera de foco en este documento.

Finalmente, el grupo de componentes del solicitante de datos se centra en un componente solicitante-administrador que interactúa con el depósito en garantía, el administrador de contrato inteligente y las aplicaciones de datos del proveedor.

Además de estos tres grupos de componentes, la plantilla de contrato de intercambio de datos es un elemento esencial del ecosistema de **DataWallet,** estático que comprende campos de datos obligatorios y opcionales, y especificaciones para operacionalizar acciones para solicitudes de datos y productos.

En la *Sección 5,* el caso en ejecución se usó para mostrar varios protocolos de interacción entre los componentes de la arquitectura. Se delineó el conjunto mínimo de ocho operaciones de cadena de bloques que permiten la trazabilidad de eventos de protocolo dinámicos inmutables.

En consecuencia, mostramos cuatro casos de ejecución del protocolo dinámico para crear perfiles de datos, realizando una transacción de datos puntuales para el proveedor que consume música adicional, intercambio de datos y productos donde el proveedor de datos verifica, si ha sido pirateado y un intercambio de experiencia aumentada donde el el solicitante de monedero de datos inicia una campaña de medios.

Tenga en cuenta que los cuatro protocolos muestran en qué puntos del tiempo el blockchain almacena eventos de transacción al ejecutar las ocho operaciones detectadas.

El sistema resultante se desarrollará por etapas, como se describe en el *Apéndice A.3.*

Sin embargo, cada etapa del desarrollo representa un producto viable, que aborda problemas urgentes, tanto para los proveedores de datos como para los consumidores de datos, y culmina en un intercambio de datos totalmente desintermediado y basado en el consentimiento.

El sistema resultante tiene el potencial de nutrir la próxima generación de productos de datos personales y experiencias personalizadas basadas en inteligencia artificial, permitiendo finalmente que los usuarios de Internet hagan que sus datos funcionen para ellos.

Referencias

[1] C.C. Aggarwal y C.X. Zhai. Minería de datos de texto. Springer Science & Business Media, 2012.

[2] I. Alqassem. Marco de requisitos de privacidad y seguridad para internet de las cosas (iot). En los Acompañamientos de la 36ª Conferencia Internacional de Ingeniería de Software, páginas 739-741. ACM, 2014.

[3] O. Arias, J. Wurm, K. Hoang e Y. Jin. Privacidad y seguridad en internet de cosas y dispositivos portátiles. Transacciones IEEE en sistemas de computación a múltiples escalas, 1 (2): 99-109, 2015.

[4] R. Baheti y H. Gill. Sistemas ciber físicos. El impacto de la tecnología de control, 12: 161-166, 2011.

[5] K. Balasubramanian, A. Gokhale, G. Karsai, J. Sztipanovits y S. Neema. Desarrollar aplicaciones utilizando entornos de diseño impulsados por modelos. Computer, 39 (2): 33-40, 2006.

[6] D. Bell. Conceptos básicos de Uml: el diagrama de componentes. IBM Global Services, 2004.

[7] I. Bentov, A. Gabizon y A. Mizrahi. Cryptocurrencies without proof of work, páginas 142-157. Springer Berlin Heidelberg, Berlín, Heidelberg, 2016.

[8] K. Bhargavan, A. Delignat-Lavaud, C. Fournet, A. Gollamudi, G. Gonthier, N. Kobeissi, N. Kulatova, A. Rastogi, T. Sibut-Pinote, N. Swamy y S. Zanella-Béguelin. Verificación formal de contratos inteligentes: documento corto.

En Actas del Taller ACM 2016 sobre Lenguajes y Análisis de Programación para Seguridad, PLAS '16, páginas 91-96, Nueva York, NY, EE. UU., 2016. ACM.

[9] D.M. Blei. Modelos probabilísticos de temas. Commun. ACM, 55 (4): 77-84, abril de 2012.

[10] L. Chung, B.A. Nixon,. Yu y J. Mylopoulos. Requisitos no funcionales en ingeniería de software, volumen 5. Springer Science & Business Media, 2012.

[11] P. Dai, N. Mahi, J. Earls y A. Norta. Protocolos de transferencia de valores de contrato inteligente en una plataforma de aplicaciones móviles distribuidas.
URL: https://qtum.org/uploads/files/cf6d69348ca50dd985b60425ccf282f3.pdf, 2017.

[12] N. Dalvi y D. Suciu. Gestión de datos probabilísticos: fundamentos y desafíos. En Actas del Vigésimo sexto Simposio ACM SIGMOD-SIGACTSIGART sobre Principios de Sistemas de Bases de Datos, PODS '07, páginas 1-12, Nueva York, NY, EUA, 2007. ACM.

[13] Zyskind G., NathanO., Y Pentland A. Descentralización de la privacidad: uso de blockchain para proteger los datos personales. En 2015 Talleres de seguridad y privacidad de IEEE, páginas 180-184, mayo de 2015.

[14] G. Kotonya y I. Sommerville. Ingeniería de requisitos: procesos y técnicas. Wiley Publishing, 1998.

[15] N. Kshetri. Problemas de privacidad y seguridad en la computación en la nube: el papel de las instituciones y la evolución institucional. Política de telecomunicaciones, 37 (4): 372-386, 2013.

[16] L. Luu, V. Narayanan, C. Zheng, K. Baweja, S. Gilbert y P. Saxena. Un protocolo de sharding seguro para blockchain abierto. En Actas de la Conferencia ACM SIGSAC 2016 sobre Seguridad de Computadoras y Comunicaciones, CCS '16, páginas 17-30, Nueva York, NY, EE. UU., 2016. ACM.

[17] A. Norta. Creación de colaboraciones Smart-Contracting para Organizaciones Autónomas Descentralizadas, páginas 3-17. Springer International Publishing, Cham, 2015.

[18] A. Norta. Establecimiento de infraestructuras de gobernanza distribuidas para promulgar colaboraciones entre organizaciones, páginas 24-35. Springer Berlin Heidelberg, Berlín, Heidelberg, 2016.

[19] A. Norta. Diseño de una capa de aplicación de contrato inteligente para transacciones de organizaciones autónomas descentralizadas, páginas 595-604. Springer Singapore, Singapur, 2017.

[20] A. Norta, P. Grefen y N.C Narendra. Una arquitectura de referencia para gestionar procesos comerciales dinámicos entre organizaciones. Ingeniería de Datos y Conocimiento, 91 (0): 52 - 89, 2014.

[21] A. Norta, L. Ma, Y. Duan, A. Rull, M. Kõlvart y K. Taveter. Características del idioma coreográfico contractual hacia la colaboración comercial entre organizaciones. Revista de Servicios de Internet y Aplicaciones, 6 (1): 1-23, 2015.

[22] A. Norta, A. B. Othman y K. Taveter. Ciclos de vida de resolución de conflictos para la colaboración de organizaciones autónomas descentralizadas gobernadas. En Actas de la 2ª Conferencia Internacional 2015 sobre Gobernabilidad Electrónica y Sociedad Abierta: Desafíos en Eurasia, EGOSE '15, páginas 244-257, Nueva York, NY, EE. UU., 2015. ACM.

[23] A. Norta, A. Vedeshin, H. Rand, S. Tobies, A. Rull, M. Poola y T. Rull. Gestión de contrato con apoyo del agente autoconsciente en blockchain para responsabilidad legal.
URL: http://whitepaper.agrello.org/Agrello_Self-Aware_Whitepaper.pdf, 2017.

[24] B.S. Panikkar, S. Nair, P. Brody y V. Pureswaran. ADEPT: Perspectiva de un profesional de IoT, 2014.

[25] S. Prince, P. Li, Y. Fu, U. Mohammed y J. Elder. Modelos probabilísticos para la inferencia sobre la identidad. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34 (1): 144-157, enero de 2012.

[26] R. Ragunathan, I. Lee, L. Sha y J. Stankovic. Sistemas ciber físicos: la próxima revolución informática. En Actas de la 47.a conferencia de automatización del diseño, DAC '10, páginas 731-736, Nueva York, NY, EE. UU., 2010. ACM.

[27] J. Rumbaugh, I. Jacobson y G. Booch. Manual de referencia del lenguaje de modelado unificado, The (2da Edición). Pearson Higher Education, 2004.

[28] S. Sicari, A. Rizzardi, L. A. Grieco y A. Coen-Porisini. Seguridad, privacidad y confianza en internet de las cosas: el camino por delante. Computer Networks, 76: 146 - 164, 2015.

[29] L. Sterling y K. Taveter. El arte del modelado orientado a agentes. MIT Press, 2009.

[30] A. S. Tanenbaum y M. Van Steen. Sistemas distribuidos: principios y paradigmas. Prentice-Hall, 2007.

[31] M. Vukolić. The Quest for Scalable Blockchain Fabric: Prueba de trabajo vs. BFT Replication, páginas 112-125. Springer International Publishing, Cham, 2016.

[32] R.M. Weber y B.D. Horn. Rompiendo vulnerabilidades de seguridad. Journal of Financial Service Professionals, 71 (1): 50-54, 2017.

[33] G. Wood. Ethereum: Un libro de transacciones generalizadas, descentralizado y seguro. Ethereum Project Yellow Paper, 2014.

The Apéndice

Hoja de ruta. Desarrollo del ecosistema

A continuación, se describe la hoja de ruta para el desarrollo del mercado transparente de datos opcionales.

Hay tres versiones principales del ecosistema del mercado:

- Nuestra implementación actual, v.10,
- El ecosistema basado en contrato inteligente v.20 y
- El v3.0 almacenado descentralizado.

La figura *XYZ*[21] muestra la línea de tiempo y el calendario de implementación, mientras que el resto de esta sección describe el sistema en cada versión de hito.

Es importante tener en cuenta que la línea de tiempo de implementación actual se basa en una serie de suposiciones. Existen tanto supuestos tecnológicos como supuestos de demanda / negocio de la comunidad. Las suposiciones de la demanda de la comunidad, por ejemplo, que existe una demanda sustancial de una solución completamente descentralizada al nivel de **Milestone 3**, considerados en nuestros objetivos de desarrollo a largo plazo. Sin embargo, estos planes solo reflejan nuestra investigación y experiencia en el ecosistema actual.

La investigación posterior y la participación de la comunidad pueden apuntar a datos más deseables que almacenan, gestionan e intercambian soluciones. Los supuestos tecnológicos incluyen el surgimiento de implementaciones escalables de algunas soluciones actualmente no escalables y preguntas abiertas de investigación.

Nos hemos esforzado por incluir esta incertidumbre en nuestra progresión de hitos, pero la aparición de las tecnologías necesarias puede afectar la línea de tiempo propuesta. Por lo tanto, las circunstancias exógenas pueden afectar esta hoja de ruta, pero las desviaciones significativas se comunicarán y explicarán claramente a la comunidad.

A.1 v1.0 Implementación actual

En la implementación actual de **Data Exchange**, los proveedores de datos gestionan sus perfiles a través de un dispositivo móvil. Los perfiles de datos se recopilan en una carpeta central del **Blockchain Wallet** para la base de datos de Intercambio de datos y todas las ventas de datos están mediadas por **dx-Insights.**

A.1.1 Sistema de Solicitante de Datos

No hay un Mercado de productos de datos abierto. Las empresas interesadas en acceder a los datos de los proveedores lo hacen a través de **dx-Insights.**

A.1.2 Sistema de proveedor de datos

Data-Profile Manager es la aplicación móvil **DataWallet**, que está disponible para iOS y Android. Se comunica con un servidor de claves y una base de datos centralizada alojada en **DataWallet**. Las posibles solicitudes de datos se publican directamente en la aplicación móvil.

[21] a discutir con James

A.1.3 Sistema de intercambio de datos

Todos los intercambios de datos se originan en **dx-Insights** y se publican en la aplicación móvil.

A.2 v2.0 Smart-Contract Based

El ecosistema basado en **Smart-contract** incorpora el compromiso de un **DataWallet** transparente, el intercambio de datos optin directamente en la arquitectura. En esta etapa, la fidelidad del sistema no depende de una parte confiable (**DataWallet**). Las garantías del sistema son estructurales y verificables por los miembros interesados de la comunidad. Todo **PII** está encriptado con una clave privada que solo los proveedores de datos poseen. Por lo tanto, todas las transacciones de datos requieren el consentimiento expreso de los proveedores de datos.

El Mercado de productos de datos está abierto de manera tal que, los Proveedores de datos no están limitados a interactuar con **dx-Insights**; Los Solicitantes de Datos pueden publicar productos en el mercado. El intercambio de datos está abierto: Todas las solicitudes de datos se ejecutan mediante contratos inteligentes en el blockchain.

A.2.1 Sistema de Solicitante de Datos

Data Product Marketplace está abierto. Los solicitantes pueden crear solicitudes de datos a través de la interfaz en línea o programáticamente con la API.

A.2.2 Sistema de proveedor de datos

Todos los datos almacenados en la base de datos centralizada alojada en **DataWallet** se cifran con una clave privada que solo los proveedores de datos poseen. Los intercambios se pueden buscar desde la aplicación móvil del proveedor

A.2.3 Sistema de intercambio de datos

El Intercambio de datos está abierto: Todas las solicitudes de datos se ejecutan mediante contratos inteligentes en la cadena de bloques que coordina el perfil de datos del proveedor y el **Request Manager** del solicitante.

A.3 v3.0 Perfiles de datos totalmente descentralizados

El Ecosistema con perfiles de datos completamente descentralizados es la realización final de la billetera auto-soberana. Los proveedores interesados pueden almacenar sus perfiles en los sistemas que elijan e interactuar directamente con el Intercambio de datos.

A.3.1 Sistema de proveedor de datos

Los perfiles de datos se pueden crear localmente usando código abierto y auditable. La estructura del perfil de datos y los agentes de búsqueda también tienen un origen abierto y son personalizables (*siempre que proporcionen la interfaz necesaria para el intercambio de contratos inteligentes*).