PROGRAMA EJECUTIVO BLOCKCHAIN, LAS PALMAS I 2020

PLAN DE NEGOCIO

by

Proyecto presentado por:

Pedro Juanes Notario | Yaiza Santana Santana | Alberto Almenara León
Juan Pedro Dyangani Ose | Michele Piazza Piazza

Tutor

Joaquín López Lérida

CONTENIDO

1 INTRODUCCIÓN

En la actualidad, la mayoría de las organizaciones trabajan con grandes volúmenes de datos no transaccionales que provienen de múltiples y diversas fuentes de origen. El problema aparece cuando este tipo de registros se gestiona mediante diferentes sistemas y se almacenan en distintas Bases de Datos (BBDD, a partir de ahora). Esta circunstancia provoca una gran dispersión de los registros existentes generando elementos duplicados y dando lugar, finalmente, a una información errónea, inconsistente y, a veces, incoherente.

El objetivo de este Trabajo está dedicado a solventar este problema para los datos de las personas, donde la única constante para ellos es que están sometidos a continuas modificaciones lo que puede convertir, rápidamente, en obsoleta la información contenida en los distintos sistemas de información de una organización si no existe en la misma un procedimiento de gestión y tratamiento específico adecuado.

Esta modificación continua de la información de la personas no sólo afecta a las Empresas (clientes), sino que también repercuten en la Administración Pública (ciudadanos) de un país; y si cada una de las modificaciones anteriores no aparecen reflejadas en los registros de datos de estas organizaciones y, además, no se distribuyen a todos sus sistemas y entran a formar parte de los procesos de gestión que dependen de los mismos, llegan a tener importantes consecuencias en forma de sobrecostes operacionales y económicos para todos los participantes en el proceso.

A continuación, citaremos (a modo de argumento ilustrativo) un breve estudio que hemos realizado para conocer la repercusión real que tiene este problema para en la Administración Pública como ejemplo de sector donde la calidad del dato y su gobernanza tienen una repercusión más que trascendente, y donde vamos a centrar nuestro “objetivo cliente” y el desarrollo de nuestra solución.

1.1 LA ADMINISTRACIÓN PÚBLICA

Para las Administraciones Públicas (AAPP, a partir de ahora) los ciudadanos adquieren una importancia capital, ya que a ellos van destinados la mayor parte de los servicios que realiza y, al mismo tiempo, son el origen directo (vía tributos) o indirecto (aportaciones del Estado proporcionales a su población de derecho) de los ingresos que recibe. Por lo tanto, la adecuada gestión de los datos de los ciudadanos que habitan y/o ejercen una actividad en un determinado territorio se convierte en una auténtica prioridad para éstas.

En España, la mayoría de los Ayuntamientos existentes son entidades dependientes de Mancomunidades de Municipios, de Diputaciones/Cabildos o de Comunidades Autónomas uniprovinciales, pero este hecho no implica que sus sistemas de información estén enlazados entre sí porque muchos han sido diseñados diseños ad-hoc a lo largo de estas últimas 4 décadas para cada una de las diferentes administraciones y dentro de las mismas.

Para el caso concreto de una AAPP a nivel de Ayuntamiento, sus sistemas de información suelen formar, en más de un 90% de los casos, un ecosistema heterogéneo, diverso, muy acoplado (es

decir, con alta dependencia entre sistemas y, por tanto, limitada capacidad de cambio) y con varios flujos de entrada de información similar desde y hacia diversos sistemas. Los motivos de esta situación son varios, principalmente la implantación gradual de software de diversos fabricantes, que por motivos comerciales no implementan por defecto soluciones estándares de comunicación a nivel de datos entre ellos. La arquitectura tipo puede resumirse en un front-end donde se ofrecen servicios al ciudadano (Portal del Ciudadano, Sede Electrónica, etc.), un back-end formado básicamente por un horizontal para la gestión de expedientes y documentos, un registro de entrada y salida, y una serie de verticales que van desde la gestión económica financiera, padrón de habitantes, aplicaciones verticales de servicios sociales, gestión tributaria, recaudación y multas, etc. Cada una de esas aplicaciones cuenta con su propia base de datos que, junto con la base de datos de terceros y territorio corporativa, forman los repositorios de datos principales en dicho ecosistema, dándose la situación de que existen varios flujos de entrada de datos de terceros en el sistema, no existiendo preponderancia de unos sobre otros, con la consiguiente situación de redundancia física, lógica, y la repetición y propagación de errores a lo largo de todas estas bases de datos derivada de la incapacidad para determinar la validez del dato, es decir de la ausencia de gobierno del dato. Por lo tanto, es lógico que estas organizaciones tengan, más pronto que tarde, como objetivo prioritario la implantación del “dato único”, que implica, en pocas palabras, que ningún dato que esté en poder de la administración municipal se repita. La información, por lo tanto, se introduciría una sola vez en el sistema, se mantendría depurada y actualizada en todo momento y sería susceptible de ser gestionada o consultada desde cualquier punto o sistema que así lo requiera. Terceros, Territorio y Documento compondrían este sistema troncal de dato único. El objetivo final es que para un determinado tercero, siendo el mismo, pueda verse de forma global desde distintos enfoques, por ejemplo: como ciudadano (desde el Padrón de Habitantes), como contribuyente (desde Gestión Tributaria o Recaudación), como proveedor (desde Contabilidad), como usuario de servicios sociales (desde las aplicaciones de Servicios Sociales), como interlocutor genérico de la administración (Registro General), e incluso como consumidor de servicios externos (tales como agua y saneamiento, desde el concesionario del servicio, etc.), sin límites ni restricciones. Otro objetivo adicional que debe plantearse cualquier AAPP es la reorganización, a nivel de arquitectura de sistemas, de la información para permitir que la misma tenga a su disposición los datos que actualmente están dispersos en diversos sistemas, de manera que sea posible obtener una visión global de los mismos sin depender de sistemas propietarios ni de ninguna otra limitación dando cumplimiento a las obligaciones señaladas en la Ley 39/2015

Debido a las circunstancias anteriormente expuestas y a otros muchos más ejemplos de problemas de esta naturaleza que se podrían exponer, nace la necesidad de almacenar todos los registros críticos personales que maneja una organización en un único repositorio y que éste pase a convertirse en una base común de referencia, capaz de eliminar la redundancia, inconsistencia y duplicidad de las diferentes versiones de los mismos, lo que facilitará y simplificará su intercambio entre los diferentes departamentos, se mejorarán los procesos de

comunicación y aumentará su coherencia, fiabilidad, integridad, exactitud y calidad así como la gestión que se realiza sobre los mismos.

La Gestión de los Datos Maestros se convierte, de esta manera, en la única estrategia efectiva a la hora de afrontar la problemática derivada de la falta de unicidad y coherencia de los registros *inter* e *intra-organizacionales*.

2 MODELOS DE DATOS MAESTROS

Un Modelo de Datos Maestros (MDM, a partir de ahora), consiste en un conjunto de metodologías, procesos y herramientas que definen y gestionan de forma consistente las entidades de datos no transaccionales de una organización o de un conjunto de organizaciones. Su objetivo es recopilar, agregar, identificar, asegurar su calidad y su persistencia, y distribuir los registros en su contexto operativo. Es, objetivamente, una práctica inicialmente metodológica y luego procedimental que permite a una organización almacenar todos sus datos críticos (que en el caso de este trabajo, se va a centrar en los registros de los ciudadanos para una Administración Pública) en un sólo almacén denominado **archivo maestro**, de forma que se obtiene un único punto de referencia común para sus datos más importantes, simplificando además su intercambio entre los distintos departamentos y facilitando el trabajo con múltiples arquitecturas, plataformas y aplicaciones.

Según el glosario de Gartner (2020) el MDM "*es una disciplina basada en la tecnología, donde el negocio y las TI trabajan juntas para asegurar la uniformidad, la precisión, la corresponsabilidad, la consistencia semántica y la responsabilidad de los activos de datos maestros comunes oficiales de la empresa. Los datos maestros son el conjunto coherente y uniforme de identificadores y atributos extendidos que describen las entidades centrales de una empresa como clientes, productos, ciudadanos, proveedores, sitios, jerarquías y planes de cuentas*".

2.1 CÓMO SE GESTIONAN LOS DATOS MAESTROS. ETAPAS DEL PROCESO MDM

Existen múltiples enfoques para gestionar este tipo de registros, pero todos ellos se basan en la construcción de un repositorio único y tienen, fundamentalmente, las mismas exigencias de accesibilidad, disponibilidad, calidad, coherencia, auditoría y seguridad para los datos. Y es que las organizaciones tienen la necesidad, cada vez más imperiosa, de disponer de una plataforma unificada que proporcione servicios de datos compartidos, utilizables en múltiples procesos y en entornos heterogéneos.

Una iniciativa completa de gestión y mantenimiento de datos maestros comprende las siguientes etapas:

- Identificar las fuentes de origen de los datos.
- Identificar los productores y los consumidores de los datos maestros.
- Recopilar y analizar los metadatos para los datos maestros recopilados.
- Determinar los responsables/administradores de los datos maestros.
- Implementar una política de gobierno de datos.
- Desarrollar el modelo de metadatos maestros.
- Escoger una solución o conjunto de soluciones como medio para mejorar la calidad de datos.
- Diseñar la infraestructura necesaria.

- Generar y testear los datos maestros.
- Modificar los sistemas consumidores y productores de información.
- Implementar un proceso de mantenimiento.

Por lo tanto, la integración adecuada de los registros juega un papel fundamental en el marco de una correcta estrategia de gestión de los datos maestros y ésta se basa en los siguientes aspectos:

- La unificación de los datos en un repositorio único, su limpieza, la creación de un registro maestro y su validación en conformidad con la aplicación de las normas de seguridad de cada organización utilizando la metodología denominada "*Record Linkage*" (Dunn, 1946).
- La actualización continua de estos registros gracias a la centralización de las modificaciones que se producen sobre los mismos.
- La disponibilidad de los datos y su divulgación para todos los sistemas gestores de la organización.

2.2 BENEFICIOS DE LOS MODELOS DE DATOS MAESTROS

Podemos resumirlos en 3 cualidades:

- Crear una única fuente confiable de aquellos datos que son claves para una organización.
- Automatizar y centralizar los procesos de calidad del dato y su gobernanza.
- Generar consistencia en los datos y en los procesos desarrollados sobre los mismos a lo largo de las diferentes unidades y sistemas de negocio de una organización.

2.3 LA PROPUESTA MDM MEJORA MEDIANTE LA UTILIZACIÓN DE LA TECNOLOGÍA BLOCKCHAIN.

La implantación y el uso de un MDM no es un requisito imponderable para una organización, por lo menos en las etapas iniciales de su proceso de digitalización, pero cuando empieza a crecer, siempre y cuando se sigan las normas de diseño y el uso de una sola BBDD relacional, éstas normalmente van a garantizar la atomicidad, la consistencia, la independencia y la durabilidad de sus datos; pero cuando lo hace a un nivel donde cada área o departamento maneja su propia herramienta de logística y sistemas de gestión de los datos (CRM, ERP, etc.), las consultas sobre ciertas entidades vitales para el negocio dejan de proceder de una sola fuente.

Y es que el crecimiento de una organización, los procesos de adquisición/fusión o interacción con otras compañías, la migración de datos entre sistemas heredados y las nuevas implementaciones, traen consigo la necesidad de implantar un MDM ya que ésta deja de ser una opción para convertirse en una necesidad.

MDM es un concepto "relativamente nuevo" para un tema que lleva preocupando a las organizaciones durante años: la necesidad de obtener y distribuir datos consistentes a través de los distintos sistemas, BBDD y aplicaciones Departamentales, superando tanto barreras

tecnológicas como organizativas, con el fin de obtener una visión unificada de los datos a través de la organización e incluso de manera extendida a su entorno de negocio. Aunque el problema no es nuevo, los MDM se han convertido de manera muy rápida en el principal foco estratégico para muchas organizaciones a nivel mundial, como evidencia el número cada vez mayor de publicaciones, proveedores de tecnología y estudios por parte de analistas sobre esta disciplina.

Un MDM se puede definir como la disciplina que permite gestionar los datos maestros de una organización (clientes/ciudadanos, proveedores, activos, documentos, etc.) con el fin de crear una única fuente fiable de referencia que mantiene los datos actualizados, con calidad y consistentes. De esta manera, un MDM permite crear una visión completa, actualizada y unificada de los datos maestros en cualquier momento a lo largo de su ciclo de vida.

El MDM es un principio fundacional, que va más allá de la pura integración y estandarización de los datos, para obtener una verdadera fuente de confianza de los mismos y poder, a partir de ahí, mejorar la toma de decisiones, la eficiencia operacional, la satisfacción del cliente y la generación de nuevas fuentes de ingresos y beneficios, entre otras ventajas tecnológicas y de negocios. El principio que acompaña a un MDM es que los datos se pueden desacoplar de las aplicaciones, conciliarlos sintáctica y semánticamente en un HUB de referencia centralizado y ser distribuidos como un servicio a las aplicaciones iniciales o a otras aplicaciones distintas a las que originaron los datos, así como alimentar Data Warehouses (para realizar Business Intelligence) u otros entornos analíticos o transaccionales. Gracias al principio/necesidad del Data as a Service (DaaS), un MDM enlaza con la Arquitectura Orientada a Servicios (SOA), porque se convierte en un auténtico paso necesario antes de abordar SOA ya que uno de sus retos pendientes es la necesidad de resolver primero las inconsistencias existentes en los datos. Con un MDM estamos extrayendo los datos de su “caja negra” y situándolos en el centro de la gestión, e incluso permite a las organizaciones dar un paso más a nivel estratégico y definir a partir de ellos políticas, estándares y procesos que determinen su uso, su desarrollo y su gestión a nivel corporativo en lo que se ha venido a denominar Data Governance tal y como señala Rouse (2007). Por supuesto, como la mayoría de los proyectos importantes, un MDM requiere compromiso a nivel de dirección y la colaboración entre los departamentos tecnológicos y de negocio. Desde nuestro punto de vista llega incluso a suponer abordar retos importantes por parte de las organizaciones, ya que a menudo son las barreras organizativas y políticas (más que las tecnológicas) las más difíciles de superar. Es el momento de pensar en los datos como un activo estratégico a nivel corporativo y como tal, al igual que otros activos económicos, debe ser gestionado para evitar que se deteriore, y es que, aumentando el valor de los datos, aumentará sin duda el valor del negocio.

La implantación del MDM posibilita la resolución de un problema creciente hoy en día: la duplicidad y la inconsistencia de los datos motivada por la multitud de sistemas operacionales con los que se trabajan en las organizaciones. La tecnología Big Data como parte integrante del MDM, aportará un valor innovador disruptivo y diferencial; y su emparejamiento técnico con Blockchain proporciona una combinación de tecnologías sujetas a una evolución continua, abriendo nuevas posibilidades para los sistemas operacionales, dando lugar a una mayor

capacidad en la captura, análisis, confiabilidad y seguridad para los datos contenidos en los mismos, repercutiendo en la mejora de los procesos de gestión y minimizando los costes operativos y económicos de las organizaciones.

Las Organizaciones son cada vez más conscientes de la importancia de la calidad de sus datos y de cómo el Big Data y Blockchain pueden contribuir a ello. El proyecto propone el desarrollo y la implementación de una plataforma Master Data Management (MDM Operacional (Open Source) con arquitectura Big Data y Blockchain que innove con respecto a estrategias y soluciones actuales de MDM, aportando eficacia y eficiencia a la hora de dar soporte estratégico a las organizaciones, combinando el análisis de la información tanto estructurada como no estructurada. Como resultado, las organizaciones tendrán acceso a nuevas formas de automatizar procesos de integración y confianza en master data internos de clientes y, a futuro, incluso en datos externos a su organización (open data, social media, etc.)

Sea como sea, un MDM sumado a la tecnología Blockchain es un verdadero cambio de paradigma que afectará a la manera de abordar iniciativas estratégicas en los próximos años dentro de la AAPP. De hecho, las principales motivaciones para su adopción son y deben ser la mejora del conocimiento y la atención al ciudadano, la gestión unificada del mismo, el cumplimento legislativo y la integración de los procesos derivados de todas las gestiones y requerimientos realizados ya sea de parte, o de oficio por parte de la Administración.

3 MODELO DE NEGOCIO

Nuestro modelo de negocio se basa en la creación de valor añadido a partir del desarrollo de nuestra solución respecto a un MDM tradicional. Para ello, implementaremos tecnología que mejore operativamente las herramientas de este tipo que existen en el mercado y una vez desarrollada e implantada en producción (utilizando metodología de análisis de requerimientos) ofreceremos un servicio de mantenimiento y soporte específico para cada uno de nuestros clientes.

La solución **D.One+** está diseñada para adaptarse a las necesidades particulares de cada cliente con el objetivo de reducir sus costes operativos y mejorar sus procesos de gestión lo que, a medio plazo, se traducirá en una auténtica inversión para la organización en la cual se implemente la plataforma.

Nuestra filosofía de empresa se basa en unos valores perfectamente definidos: personalización de la solución, soporte de calidad y capacitación de nuestro capital humano.

Nuestro beneficio provendrá de la venta e implantación de nuestra solución tecnológica y del soporte técnico que ofreceremos.

Utilizaremos licencias base de tipo *Open Source* (Hadoop, Spark, Hyperledger Fabric, SFTP y LuceneRDD) para desarrollar nuestro MDM **D.One+** lo que permite la no existencia de costes de licenciamiento para nuestra empresa, y que además conozcamos exactamente cuáles son las potencialidades y las limitaciones iniciales para cada una de ellas. Una vez creado el sistema éste se implantará en el Ayuntamiento 1 de Gran Canaria, organización con la que se firmará un convenio de colaboración que permitirá a nuestra empresa probar la solución en un cliente bajo un entorno de producción. Posteriormente, y después de la estabilización de los posibles errores e incidencias que se produzcan, paquetizaremos **D.One+** y lo pondremos a la venta como un producto singular para las organizaciones interesadas.

Los **beneficios** que obtenemos al utilizar licencias *Open Source*:

- *Bajo coste de desarrollo*: podemos adaptar el software a nuestras necesidades específicas cambiando el código fuente. Debido a una menor inversión inicial de capital y a la capacidad de personalizar libremente los paquetes de aplicaciones, es una solución que hemos decidido escoger para desarrollar nuestra solución.
- *Actualizaciones periódicas*: Solución de errores e implementaciones adecuadas con más rapidez que las soluciones propietarias.
- *Alta seguridad*: La actualización permanente del código que mencionamos en el punto anterior realizada por una poderosa comunidad de usuarios activa, permite ir mejorando progresivamente sus niveles de seguridad ante las vulnerabilidades detectadas.
- *Compatibilidad*: Al trabajar con formatos estándar capacita a los usuarios que tienen un determinado hardware para adaptar el código a sus controladores, lo que hace que

finalmente exista una solución muy universal que permite una interoperabilidad más alta entre distintos sistemas.

- *Bajo riesgo por la independencia del fabricante*: La no dependencia de empresas de desarrollo informático tradicionales (expuestas al rigor del mercado) hace que nuestro futuro no dependa de cambios/movimientos a los cuales están continuamente comprometidas y que pueden llegar a traducirse incluso en el abandono de desarrollos y versiones.

Los **riesgos** que asumimos al utilizar licencias *Open Source* son los siguientes:

- Los desarrollos licenciados en *Open Source* están creciendo a gran ritmo, pero aún en ocasiones no están lo suficientemente maduros frente a sus homólogos creados en licencias de software propietario.
- *Absorción por empresas de software propietario*. No se puede asegurar que éstas no exigirán en un futuro licencia para utilizarlo. Ejemplo: Oracle compra MySQL.

En la actualidad existen otras empresas cuyo modelo de negocio están relacionados también con licencias *Open Source*. Como ejemplo de las mismas podemos mencionar: a Red Hat Enterprise Linux y Revolution Analytics.

3.1 PROPUESTA DE VALOR

La propuesta de valor que ofrece nuestra empresa es la ejecución de un sistema único de datos el cual proporcionará a la AAPP una solución a su problemática referida a los datos duplicados existentes en sus distintos sistemas de gestión. El proyecto propone el desarrollo y la implementación, basados en una plataforma Blockchain Open Source de un MDM operacional soportado en una arquitectura Big Data denominada D.One que innove con respecto a estrategias y soluciones actuales, aportando eficacia y eficiencia a la hora de dar soporte estratégico a las organizaciones, combinando el análisis de la información tanto estructurada, como no estructurada. Actualmente, no existe en el mercado una solución con estas características. Como resultado, las organizaciones tendrán acceso a nuevas formas de automatizar procesos de integración y confianza en sus datos maestros internos de personas/clientes y, a futuro, incluso en datos externos a su organización (open data, social media, etc.)

Todo ello proporcionará los beneficios tanto a nivel económico, como de servicio mostrados en la siguiente imagen:

Ilustración 1: Propuesta de Valor. Fuente: Elaboración propia.

Existen 2 aspectos que influyen en la propuesta de valor para nuestro proyecto:

- Cumplimiento del RGPD y del ENS creando la obligación (para los organismos que utilizan datos de terceros) de acatar dicha normativa legal si no quieren ser penalizados con multas de alta cuantía por incumplimiento.
- Utilización de tecnología Blockchain que aporta eficiencia, seguridad, transparencia y accesibilidad.

3.2 SOCIOS CLAVE

El desarrollo de nuestro proyecto es posible, inicialmente, por el acuerdo alcanzado con el Ayuntamiento 1 de Gran Canaria, que nos ha proporcionado sus datos de terceros para que podamos resolver la problemática que presenta (distintos registros para un mismo individuo, vinculados a diferentes sistemas de gestión). Además, la aplicación del RGPD conlleva obligatoriamente, para cada una de las entidades en las cuales se gestionan datos personales, la creación de la figura de Delegado de Protección de Datos, lo cual se convierte en una oportunidad para nuestro modelo de negocio.

Además, debemos establecer acuerdos con proveedores de tecnología TIC, los cuales nos facilitarán el servicio en Cloud para desarrollar nuestra solución y dar servicio a nuestros clientes, que obtendrán de esta manera los siguientes beneficios:

- *Menores costes y aumento de la productividad (automatización de las tareas).*
- *Tecnología puntera para sus necesidades.*
- *Canales de comunicación B2B, con un Acuerdo de Nivel de Servicio (SLA) del 99,99%.*

3.3 RECURSOS CLAVE

Trabajaremos con equipos informáticos de última generación que nos capacitarán para ofrecer de una manera rápida, fiable y robusta la solución D.One a nuestros clientes en cada una de las fases de cada uno de los proyectos.

Nuestro equipo humano cuenta con profesionales cualificados que darán soporte a cada una de las necesidades de nuestros clientes. Nuestra política es dar valor al negocio mediante la creación de servicios y datos fiables/transparentes para entregar un servicio completo y ad hoc a los requerimientos particulares de cada organización.

3.4 RELACIÓN CON EL CLIENTE

Nuestra principal actividad respecto al cliente estará centrada en el asesoramiento y la consultoría para la AAPP. Nuestra metodología se basa en el siguiente modelo de trabajo:

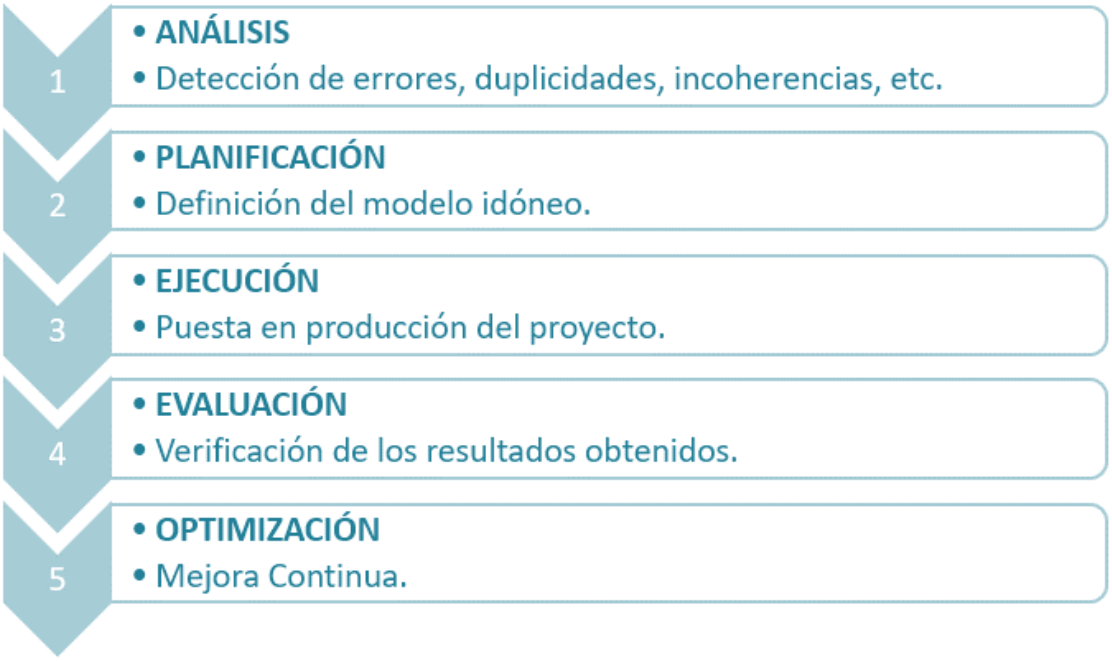


Ilustración 2. Modelo de trabajo D.One. Fuente: Elaboración propia.

Nuestro objetivo es ofrecer un servicio de nuestra plataforma siempre actualizado y, además, mantener en cada fase del proyecto un contacto directo con el cliente que posibilitará una comunicación directa, inmediata, fluida y un mejor servicio siempre personalizado para las circunstancias específicas de cada organización.

3.5 SEGMENTO DE CLIENTE

Nuestro innovadora Plataforma MDM está desarrollada para dar respuesta a un problema creciente en la AAPP, y éste es el segmento de clientes al cual dirigiremos nuestro producto, así como todas las actividades y acciones derivadas del proceso empresarial implícito. En el apartado 5.1 de esta propuesta se desarrollará explícitamente este aspecto.

3.6 CANALES DE DISTRIBUCIÓN

La empresa D.one habilitará una potente red publicitaria que mostrará de manera explícita nuestra solución MDM + Big Data + Blockchain y, además, pondremos a disposición de nuestros potenciales clientes un cuestionario específico de registro donde podrán expresar las circunstancias particulares que presenta su organización respecto al problema que nos ocupa y, a través del cual, podremos iniciar nuestra labor de consultoría particular con datos/situaciones reales. El desarrollo de esta actividad queda reflejada de manera profusa en el apartado 11.1 del presente documento.

3.7 ACTIVIDADES CLAVE

La actividad clave para el desarrollo de la actividad de D.One es la gestión adecuada del dato único. Para ello es necesario dominar los siguientes aspectos: “Conocimiento del Negocio”, “Desarrollo y Conocimiento de la Solución” y “Conocimiento del Modelo de Datos. El desarrollo de cada uno de ellos se detalla en el apartado 6.1 del presente Trabajo.

4 ANÁLISIS EXTERNO

4.1 MERCADO OBJETIVO

Nuestra solución está desarrollada inicialmente en convenio con el Ayuntamiento 1 de Gran Canaria el cual nos facilitará los registros de sus ciudadanos incluidos en distintas BBDD que posee la organización para que procedamos a su procesamiento. El problema que venimos a resolver para esta Ayuntamiento es extrapolable a todas las entidades de la AAPP donde cada Departamento viene a tener su propio registro de personas, por lo que es la solución adecuada para este segmento de cliente (Ayuntamientos, Mancomunidades, Diputaciones/Cabildos, Gobiernos Autonómicos, etc.)

Por lo tanto, esta situación unida a la realidad de que la AAPP está obligada a cumplir con las obligaciones derivadas de la Ley 39/2015 (problema que D.One cubre) nos permite afirmar que existe un gran potencial de éxito para nuestra idea de negocio.

4.2 VARIABLES DE MERCADO Y COMPETIDORES

Las siguientes empresas dominan el mercado mundial del MDM: SAP, Oracle, IBM, Informatica, Stibo Systems, TIBCO, Riversand Technologies, Orchesta Network, EnterWorks, Microsoft, KPMG, Teradata y Software AG. Y las soluciones MDM para estas compañías en el mercado son las siguientes: Informatica MDM, Orchestra Networks EBX, SAP Master Data Governance, IBM MDM Solutions, Reltio Cloud, TIBCO MDM, Riversand MDM, Stibo Systems Master Data Management y Profisee. Estas 2 relaciones no constituyen de forma exhaustiva a todos los proveedores y a todas las ofertas de soluciones MDM existentes, y se puede encontrar más información al respecto en este informe de Reuters .

Tal y como señalamos al inicio de este punto, algunos ejemplos de empresas que han desarrollado herramientas MDM son Stibo System (con su solución STEP, una de las principales en la gestión de datos maestros), IBM (que cuenta con 4 productos de gestión de datos maestros escalables dentro de la familia InfoSphere y donde podemos destacar la existencia de una solución cloud específica para proyectos de Big Data, almacenamientos de datos e iniciativas Hadoop), Informatica (con diferentes soluciones relacionadas con clientes, productos, ciudadanos, etc.), Paradigma (solución ATREO que utiliza tecnología Big Data como parte integrante de su aplicación), e incluso hay asesoras como PowerData que ofrecen la implantación del MDM de Informatica junto a servicio de asesoría y creación de manuales de estándares y buenas prácticas particularizados para cada organización).

A continuación, analizamos el Cuadrante de Gartner de 2020 para este tipo de soluciones:

Ilustración 3. Cuadrante Mágico de Gartner para las soluciones MDM. Fuente (Gartner, 2020).

La información está obtenida en el mes de enero de 2020 y, para asegurar una alineación cercana con los requisitos del comprador, Gartner enfoca su evaluación en múltiples dominios, una variedad de nuevos casos de uso y soporte para diferentes estilos de implementación como parte de los criterios de investigación. Como consecuencia, creemos que la investigación de MDM de Gartner es de alta calidad y se convierte en una guía recurrente de facto por parte de la mayoría de los expertos en tecnología en la actualidad. Con ese trasfondo, veamos ahora las conclusiones clave que se destacan de su estudio:

- En 2020 los MDM siguen ganando importancia entre las organizaciones inmersas en iniciativas comerciales que respaldan una mejora en su experiencia, gobierno, análisis, crecimiento y eficiencia respecto a la gestión de sus clientes.
- Informática es la empresa líder y está posicionada en el cuadrante con el valor más alto para la dimensión '*Capacidad de ejecución*' y más avanzado para la dimensión *'Completitud de visión*'.
- Según Gartner, la tendencia a adquirir soluciones MDM continúa cobrando impulso a medida que aumenta la conciencia de que sus beneficios son cada vez más transformadores a nivel del proceso de negocio.

- El MDM basado en la nube aún permanece en las primeras etapas de madurez. La mayoría (81%) de los 144 encuestados informaron tener una solución en sus instalaciones (cliente-servidor).
- Para el año 2020, la estimación es que el 80% de todos los ingresos de la solución MDM derivará del soporte de los proveedores para todos los requisitos de aquellas organizaciones vinculadas al sector industrial y al dominio de los datos (en comparación con el 15% estimado para el año 2019).
- En el caso de las tecnologías de bases de datos NoSQL, es posible que las tecnologías actualmente disponibles sean inapropiadas para los escenarios de MDM que requieran un alto nivel de integridad transaccional (generalmente provisto únicamente por las bases de datos relacionales).
- Gartner señala que el crecimiento pronosticado para los dispositivos conectados (IoT) y los datos generados por éstos (junto con el valor económico derivado) aumentará la necesidad de gestionar de manera adecuada los datos maestros asociados y el impacto de éstos en las organizaciones (y lo harán de forma muy acelerada).
- Muchos proveedores de soluciones MDM están presentes en el mercado, pero su relevancia no es lo suficientemente importante (en uno o más aspectos) para ser incluidos dentro del Cuadrante Mágico. Es nuestro caso, **D.One** como Start Up, pero si tuviéramos que definirnos dentro del gráfico estaríamos en una posición de propuesta "visionaria" y como solución "competitiva" respecto a las existentes por las características de nuestro desarrollo técnico y nuestra visión operativa.
- Las soluciones MDM analizadas comienzan a apostar por tecnologías Big Data. Así, por ejemplo, algunas de las principales soluciones en el mercado como las de SAP o IBM, ya incorporan Apache Hadoop en su entorno computacional, el cual es utilizado desde muchas de sus soluciones, entre ellas sus sistemas MDM. En el caso de Informatica, líder en el sector, apoya su MDM en una plataforma tecnológica Big Data muy similar a la que planteamos nosotros, ya que cuenta con Hadoop y Spark sobre el entorno Cloud de Amazon AWS, pero ninguna de ellas integra tecnología Blockchain en un entorno de producción, por lo que podemos señalar que nuestra idea de negocio y desarrollo es pionera a nivel mundial en este aspecto y le confiere un halo de unidad con respecto al resto de soluciones.

La situación propuesta para nuestra empresa dentro del Cuadrante Mágico de Gartner de 2020 estaría enmarcada dentro del cuadrante "Visionarios" dado el carácter sumamente innovador de nuestra propuesta tecnológica. Somos una empresa incipiente con una idea innovadora (de ahí nuestro carácter "Visionario"), pero nos encontramos dentro de un mercado dominado por grandes empresas tecnológicas con mucha experiencia en el desarrollo e implantación de soluciones MDM y podemos encontrarnos con el inconveniente de que el desarrollo propuesto sea más complicado de ejecutar y de implantar que lo previsto en este Plan de Negocio y que

los clientes sean más reticentes de elegir nuestra solución que las ya desarrolladas por las empresas que dominan este mercado.

4.3 OBJETIVO Y OPORTUNIDAD DE NEGOCIO

Como ya hemos señalado con anterioridad, un MDM comprende los procesos, la gobernanza, las políticas de seguridad, los estándares de calidad y las herramientas que permiten a las organizaciones gestionar sus datos maestros más valiosos de forma unificada, proporcionando mecanismos para agregar y asegurar su calidad y consistencia, evitando errores y duplicidades. Si a todo ello sumamos las bondades de la BBDD distribuida Blockchain, esperamos obtener como resultado un sistema más completo que los existentes en la actualidad en el mercado.

En la siguiente tabla, realizamos una comparativa real de las soluciones MDM actuales frente a la solución MDM D.One y podemos comprobar cómo las características de nuestra plataforma constituyen por sí mismas un claro objetivo de desarrollo y una inmejorable oportunidad de negocio.

DIMENSIÓN ANALIZADA	SISTEMAS MDM	SOLUCIÓN D.One+
PROCESOS Y FLUJOS DE TRABAJO	Los procesos de extracción, transformación y carga se realizan sobre tecnología Big Data (normalmente Hadoop). Además, el proceso de transformación y deduplicación deben ser ejecutados por un programador de tareas desarrollado al efecto (scheduler) para su automatización.	El proceso de extracción se realiza sobre Hadoop, la transformación de los datos (deduplicación y Record Linkage) sobre Spark y, finalmente, el almacenamiento sobre nodos construidos con tecnología Blockchain (Hyperledger Fabric) respetando el modelamiento inicial del dato y dependiendo únicamente de la disponibilidad de la red.
CALIDAD DE LOS DATOS	Los procesos de transformación y limpieza de los registros están basados en reglas y/o modelos de datos. Finalmente, la decisión de qué registro es el correcto la toma el área	Los procesos de transformación y limpieza de los registros están basados en modelos Machine Learning. La decisión de qué registro es el correcto está tomada mediante un modelo de clasificación

	encargada de la limpieza y calidad de datos.	supervisada de vectores (probabilidad más alta).
REGLAS DEL NEGOCIO	La implementación de un MDM conlleva que se generen modelos basados en el negocio de cada cliente. Al igual que los Data Warehouses, la BBDD creada al efecto del MDM representa también el negocio de una forma lógica.	**D.One+** también está modelado según los requerimientos de los datos de cada cliente. Pero cada bloque de datos creado (al construirse sobre Blockchain) permanece inalterable.
FLUJOS DE DATOS PROCESADOS	Admite grandes flujos de datos provenientes de distintas fuentes validadas. Pero su limitación está relacionada con el hardware utilizado, especialmente por la cantidad de memoria RAM (según IBM InfoSphere)	Al utilizar procesos de paralelización, el flujo de datos que es capaz de procesar **D.One** es prácticamente infinito. La limitación, en este caso, de Spark viene dada por la cantidad de bloques HDFS que sea capaz de absorber cada nodo (64mb por bloque). Sin embargo, no es determinante ya que también utiliza una asignación dinámica de recursos para optimizar el flujo de datos procesados
AUDITORÍA	Mediante metadatos y logs.	Mediante el libro mayor de Blockchain permite verificar cada una de las versiones de cada bloque de datos (histórico).
SOPORTE	Prestado por el desarrollador, empresas	Prestado por el equipo de trabajo de **D.One**

	expertas en alguna solución específica y consultoras.	No hay dependencias de empresas externas a nuestra solución.
ESCALABILIDAD	Modelos escalables en el tiempo basados en la modificación de su diseño original.	No permite la escalabilidad a nivel de bloque. Los históricos son inmutables.
SEGURIDAD	La seguridad es proporcionada por el SGBD, el Sistema Operativo y el Firewall utilizado.	La seguridad del dato está encriptada para cada bloque. Los perfiles de lectura son los establecidos por la red Blockchain. Se utiliza Linux (IPTables) como Sistema Operativo y contamos con un Firewall lógico para cada uno de los nodos de la red creada.
RENDIMIENTO	El rendimiento de InfoSphere MDM de IBM alcanza las 21.000 transacciones por segundo (para ejecuciones de lectura de complejidad media).	Hyperledger Fabric se ejecuta en un servidor con una velocidad de lectura y escritura promedio de 20.000 transacciones por segundo (TPS) y los desarrolladores señalan que pronto llegará a 50.000 TPS.

Tabla 1. Comparativa “MDM tradicionales” vs solución MDM **D.One+**.

La tecnología Big Data + Blockchain, como parte integrante del MDM, aportará solución a los problemas planteados anteriormente en aspectos como:

- La eficiencia del proceso.
- El almacenamiento.
- El análisis de los datos.
- La creación de los registros maestros.
- Su compartición con otras plataformas.

Las Organizaciones son cada vez más conscientes de la importancia de la calidad y gobernanza de sus datos y de cómo esta tecnología puede contribuir a ello. Big Data + Blockchain proporcionan, además, nuevos mecanismos de procesamiento de registros que permiten extraer información valiosa e, incluso, obtener mejores resultados en la estrategia de inteligencia de negocio para la Organización. Por lo tanto, el emparejamiento en este proyecto de **MDM + Big Data + Blockchain** está repleto de retos, ya que partiremos de la combinación de tecnologías ¿incipientes? y sujetas a una evolución continua que requiere tiempo e I+D+i para alcanzar la madurez como producto.

Actualmente, no existe en el mercado una solución que cumpla con el conjunto de todas estas características y el desarrollo del producto posibilitará a las organizaciones el acceso a nuevas formas de automatizar procesos de integración y confianza en datos maestros internos de clientes y, a futuro, incluso en datos externos a su organización (open data, social media, etc.)

El planteamiento de la solución tecnológica de la plataforma **D.One+** innova con respecto a los modelos actuales en 2 aspectos fundamentales:

- **Desde una perspectiva técnica**, la arquitectura Big Data proporciona las **5 V's (Volumen, Velocidad, Variedad, Veracidad y Valor)**, posibilitando la incorporación de registros estructurados y no estructurados; a lo cual se suma el uso dentro de la plataforma de tecnología **Blockchain** cuyas principales ventajas son:
    - *Inmutabilidad de la información*: es prácticamente imposible modificar los registros contenidos en la red, y en caso de suceder podría invalidar la cadena de bloque.
    - *Custodia distribuida*: nadie es dueño del 100% de la red, pues diferentes usuarios almacenan distintos nodos donde cada uno contiene copias actualizadas idénticas de la información.
    - *Resiliencia*: Blockchain es tolerable a los fallos en los nodos, pues si alguna parte de Blockchain falla, toda la red puede continuar trabajando con la última versión disponible de la información.
    - *Confianza*: La tecnología Blockchain funciona según el consenso del contenido de la información, por lo tanto, no necesita un intermediario que brinde confianza sobre el mismo.
- **Desde el punto de vista de negocio**, la solución supone una innovación con respecto a las soluciones MDM propietarias existentes en el mercado, basadas o no en Big Data, y sumando, como nadie lo ha puesto en marcha aún, la tecnología Blockchain para dar respuesta a las exigencias impuestas por el RGPD y la normativa española al respecto de la Ley de Procedimiento de la Administración Pública.

5 ANÁLISIS INTERNO

5.1 PROCESOS ESTRATÉGICOS Y ACTIVIDADES CLAVE DEL PROYECTO

La actividad clave para el desarrollo de la actividad de D.One es la gestión tecnológica adecuada del Dato Único. Para ello, es necesario dominar los siguientes aspectos:

- **Conocimiento del Negocio**: Para llevar a cabo un buen proyecto, todas las personas que intervengan en el mismo deberán conocer los alineamientos, las claves maestras, las actitudes, los procesos y la organización del cliente, además de tener conocimiento sobre sus factores ambientales, legales, operativos y organizativos.
- **Desarrollo y conocimiento de la solución**: La empresa D.One ofrece además de las labores de asesoramiento y consultoría realizadas por su equipo humano, la implementación de una potente solución MDM con las características que desarrollamos en la parte técnica del Trabajo y que ha sido diseñada íntegramente por nuestra empresa y, por tanto, todo su equipo humano es conocedor de su potencial y sus capacidades.
- **Conocimiento del Modelo de Datos**: Para nuestro propósito de negocio, entender desde el inicio de un proyecto cómo están conformados los modelos de datos existentes dentro de la organización objetivo, es fundamental para poder determinar sus problemas de base. Además, podremos estimar directamente las posibles mejoras en su rendimiento, optimizando la información que luego será implementada dentro de la herramienta MDM a través de procesos continuos de Extracción, Transformación y Carga (ETL), conformando en una Base de Datos única el dato maestro que será ofrecido en tiempo real a los Sistemas de Gestión originarios con el objeto de mejorar su gestión de negocio y generar un valor añadido directo. D.One ofrece este sistema como un SaaS, al cual accederán de manera securizada las organizaciones para procesar los datos originales y obtener sus activos digitales únicos.

5.2 DIRECCIÓN ESTRATÉGICA

Tras la obtención de los datos, éstos serán tratados para generar una Base de Datos única (D.One_DB) donde se almacenarán los registros maestros para ser ofrecidos en tiempo real a los distintos Sistemas de Gestión originarios con el objeto de mejorar sus procesos de negocio y generar un valor añadido directo. Se ofrecerá un sistema en la nube, al cual accederán de manera securizada las organizaciones para procesar los datos originales y obtener sus activos digitales únicos. Además, se ofrecerá servicios de asesoramiento y consultoría.

5.3 DIRECCIÓN COMERCIAL

Los servicios de asesoramiento y consultoría proporcionarán contacto directo y personalizado con el cliente tras la implantación del sistema. Pero, además, se creará un Departamento Comercial que llevará a cabo visitas a los actuales y potenciales clientes para recabar feedback sobre la solución implantada. Así mismo, se concretarán visitas con las AAPP que se consideren potenciales clientes con el objetivo de mostrar nuestra solución.

Se dispondrá de un espacio en la web (www.d-one.es) donde se presentará la solución y se mostrarán nuestros principales casos de éxito. Estaremos presentes en las redes sociales adecuadas donde subiremos contenido que explique de forma detallada nuestra solución tecnológica, así como el ofrecimiento de la posibilidad de suscribirse a una newsletter y poder así recibir noticias relacionadas con el sector y nuestra evolución. Por otro lado, dispondremos de un callcenter para la resolución de dudas y problemas técnicos.

Por último, se acudirá a diferentes eventos del sector tecnológico relacionados con nuestra solución para presentar la herramienta, nuestros servicios y nuestras implantaciones.

5.4 INNOVACIÓN Y DISEÑO

La Plataforma estará sujeta a un proceso de mejora permanente debido a la continua entrada de datos desde los diferentes sistemas de gestión de una organización. Se desarrollarán sistemas que conlleven una reducción en los tiempos de carga de la información y en la realización de consultas. Así mismo, las distintas bases de datos serán revisadas para obtener un mejor rendimiento y la interfaz gráfica del sistema y el propio Dashboard de la aplicación será mejorado continuamente.

Con el objetivo de cumplir con la normativa relacionada con la protección de datos, se realizarán todas las modificaciones oportunas.

Por último, se revisarán y se actualizarán los sistemas de seguridad actualizándose con las nuevas mejoras tecnológicas que vayan apareciendo en el mercado.

5.5 PROCESOS DE APOYO DEL PROYECTO

Gestión Financiera

La gestión financiera está destinada a la administración de los recursos financieros, activos y pasivos de la empresa. Con el objetivo de reducir los costes iniciales, los primeros meses no es necesaria la ubicación física del equipo en una oficina porque el trabajo se realizará en remoto, apoyándonos además en una solución construida para su funcionamiento en la nube. El equipo dispondrá de equipos portátiles para poder acceder en todo momento al sistema y a las comunicaciones que se establezcan dando el adecuado servicio a nuestros clientes. Respecto a la parte que concierne al tratamiento contable y fiscal, tal y como comentábamos anteriormente, será realizada completamente por el equipo de D.One que cuenta entre sus integrantes con un experto en el aspecto financiero del proyecto.

Se requerirá una inversión inicial de 10.000 euros para cubrir los costes del primer año de funcionamiento. Los gastos de capital humano, infraestructuras tecnológicas, equipos informáticos, internet, telefonía, publicidad y marketing son los principales gastos que tendrá que soportar la empresa a lo largo de su desarrollo.

Gestión de Personal

Inicialmente no se contratará a personal externo. La incorporación de nuevo capital humano estará supeditada a la necesidad motivada por un número de implantaciones de la solución y procesos de mantenimiento que no podamos asumir los actuales integrantes del equipo.

El equipo de D.One será el responsable de gestionar la empresa. Los diferentes perfiles profesionales (administración de empresa, informáticos y analistas de datos) hacen que queden cubiertas las necesidades de las distintas áreas de la misma y que detallaremos en el Plan de Recursos Humanos.

6 DAFO

En este apartado describimos las Debilidades, Amenazas, Fortalezas y Oportunidades de nuestro modelo de negocio a través de un DAFO, actuando como resumen de los capítulos anteriores a partir del análisis que hemos realizado.

De forma básica, puede afirmarse que un análisis DAFO se centra en los factores internos para reconocer fortalezas y debilidades de un proyecto, lo que a su vez nos permitirá conocer las oportunidades y las amenazas a partir de los factores externos. Se trata de un análisis idóneo y necesario para obtener información crítica a partir de la cual tomar decisiones operativas.

6.1 DEBILIDADES

- Primera aproximación de nuestro equipo en este tipo de proyectos: Inexperiencia de los componentes en los procesos operacionales y en el desarrollo de la arquitectura de la solución tecnológica a desarrollar.
- Empresa de reciente creación: Dificultará, en un principio, aspectos como la confianza por parte del cliente, el acceso a la financiación, los acuerdos de colaboración, etc. La propia 'novedad' de nuestro proyecto empresarial hace que no existan referencias favorables (o desfavorables) dentro de los potenciales clientes.
- Recursos financieros limitados: Posibles inversiones a realizar para el desarrollo de los servicios, que no sean factibles por falta de monetización.

6.2 AMENAZAS

- Resistencia de la Administración Pública a los cambios disruptivos. Los mismos están mejor considerados en las organizaciones empresariales de carácter privado.
- Déficits presupuestarios: recursos económicos por parte de los potenciales clientes que no les permitan abordar este proyecto tecnológico. Actualmente, muchas organizaciones se encuentran inmersas en el proceso de la digitalización externa de sus procesos y están volcando en este proyecto la mayoría de sus recursos financieros.
- Empresas que ofrecen otras soluciones: competidores nacionales e internacionales.
- Desconfianza y desconocimiento por parte de las organizaciones sobre el tema de Blockchain por su encasillamiento en las criptomonedas.
- El propio componente absolutamente innovador hace que la solución desarrollada pueda ser vulnerable a la obsolescencia/desaparición prematura de alguna de las tecnologías utilizadas y que, en un principio, puede parecer la más adecuada.

6.3 FORTALEZAS

- Solución innovadora y disruptiva (MDM + Big Data + Blockchain), que permite obtener grandes beneficios operacionales y una securización y confiabilidad absoluta.

- La solución propuesta resuelve una importante necesidad de las AAPP, dando respuesta a las obligaciones exigidas por la Ley 39/2015, el RGPD y el ENS.
- Servicio personalizado: Nuestro trabajo se basará en generar una relación de proximidad y confianza y se adaptarán a las necesidades específicas de cada cliente.
- Equipo multidisciplinar: Los miembros del proyecto cuentan con perfiles variados, lo que aporta una gran experiencia en diferentes campos: asesoría tecnológica, IT, asesoramiento legal y de negocio.

6.4 OPORTUNIDADES

- Apuesta de las AAPP y las organizaciones empresariales por el impulso de las TIC en el core de su negocio: la unicidad, integridad y confiabilidad absoluta de los datos pertenecientes a sus ciudadanos/clientes.
- Alianzas estratégicas con otras AAPP: Consecución de acuerdos con otras entidades gubernamentales que permitan mejorar y expandir el servicio que ofreceremos a nuestros clientes.
- Pandemia COVID-19. Trámites telemáticos “on fire”. Además de todo el dinero que puede generar este proyecto, innovar en esta industria para la AAPP representa una oportunidad de generar un impacto positivo en la calidad de vida de las personas.
- Plataforma EBSI y caso de uso ESSIF (UE). Identidad Digital Auto Soberana. Proyecto incipiente para el cual no hemos encontrado referencias de utilización en el caso que nos ocupa.

		ANÁLISIS INTERNO	ANÁLISIS EXTERNO	

D	NEGATIVO	• Inexperiencia de los componentes del equipo humano en este tipo de proyectos.
		• Empresa de nueva creación, sin referencias previas:
		  • Falta de confianza,
		  • Acceso a financiación.
		  • Acuerdos de colaboración.
		• Recursos financieros limitados
			• Resistencia de las AAPP a los cambios.
			• Déficits Presupuestario.
			• Empresas que ofrecen otras soluciones MDM.
			• Desconfianza y desconocimiento por parte de los clientes de la tecnología Blockchain.
			• Posible obsolescencia prematura de alguna de las tecnologías empleadas.
				A

F	POSITIVO	• Solución innovadora y disruptiva MDM + Big Data + Blockchain. • La solución propuesta resuelve una importante necesidad de las AAPP, dando respuesta a las obligaciones exigidas por la Ley 39/2015, el RGPD y el ENS. • Servicio personalizado. • Equipo Multidisciplinar. • Alta motivación y capacidad equipo.	• Apuesta de las AAPP por el impulso de las TICS y alianzas estratégicas interadministrativas. • Trámites telemáticos "on fire". Innovar en este sector representa una oportunidad de generar un impacto positivo en la calidad de vida de las personas. • Plataforma EBSI y caso de uso ESSIF. Identidad Digital Auto Soberana.	O

Ilustración 4. DAFO. Fuente: Elaboración propia.

7 PLAN ESTRATÉGICO

D.One nace como Trabajo Final del "*Máster en Business Intelligence y Big Data*" realizado en la EOI en el año 2017 y ahora su diseño, gracias a lo aprendido en el "*Programa Ejecutivo de Blockchain",* se completa desde un punto de vista técnico a nivel de su capa de almacenamiento Blockchain y la integración en el sistema de interoperabilidad con la Plataforma de Intermediación de Datos y el caso de uso ESSIF de la Plataforma EBSI (pasando a denominarse **D.One+**); y éste ha sido el escenario donde hemos afianzado los conocimientos adquiridos durante el mismo, donde avanzamos en una línea de investigación y desarrollo muy atractiva y donde reforzamos la visión de que la innovación, basada en los datos y el buen uso de los mismos otorga a las organizaciones un verdadero factor estratégico diferenciador y de vital importancia en la actualidad.

7.1 VISIÓN, MISIÓN Y VALORES

La visión por parte de **D.One** es el desarrollo y crecimiento constante de nuestra solución y servicios, buscando obtener el mayor valor añadido posible de la tecnología utilizada y el capital humano empleado y puesto a disposición del cliente, generando para el mismo nuevas oportunidades de negocio que satisfagan sus necesidades presentes y futuras (proactividad).

Nuestra misión principal como empresa consultora, es poner a disposición de nuestros clientes un servicio profesional puntero, proporcionando control sobre la gestión completa para el ciclo de vida de todos los proyectos que implantaremos. Queremos ser pioneros en el uso de la tecnología Blockchain y Big Data para dar soluciones MDM íntegras y que generen verdadero valor a nuestros clientes. Somos un equipo altamente cualificado para ofrecer a nuestros clientes el mejor servicio.

Nuestros valores:

- Atención y asesoría personalizada.
- Mantenimiento de las infraestructuras para la prestación adecuada del servicio.
- Cultura de mejora continua dentro de la organización.
- Máxima satisfacción del cliente.
- Aportar valor añadido a las organizaciones mediante la aportación de nuestra solución y nuestros conocimientos.

7.2 OBJETIVOS ESTRATÉGICOS

Los objetivos del proyecto son:

- Plantear un repositorio maestro que centralice los activos de datos de una organización y cree las entidades de datos maestros.
- Definir la captura de flujos de datos para alimentar el MDM.

- Diseñar una solución para la gestión de los datos maestros alineada con la tecnología más adecuada en la actualidad.
- Definir la arquitectura prototipo que integre los distintos componentes del proyecto según avanzan las iteraciones para su desarrollo futuro como Plataforma MDM, creando en un futuro próximo un producto mínimo viable que nos permita validar la parte básica operativa de nuestro proyecto y que será el primer prototipo que presentaremos al mercado para buscar las primeras reacciones de nuestros clientes potenciales, utilizando metodologías Ágiles de Desarrollo (tipo SCRUM).

8 PLAN TECNOLÓGICO

Para desarrollar de una manera adecuada la parte técnica de nuestro proyecto, hemos incluido dentro de este apartado los 3 aspectos (a mayores del MDM que ya desarrollamos en el capítulo 3) que han influido y que deben de influir en la consecución de nuestro propósito final (que es la definición de la arquitectura de nuestra solución), con carácter de revisión y novedad respecto a lo ya definido en el TFM del cual parte este trabajo. De esta manera, se ha desarrollado un capítulo para cada variable con entidad e importancia para conseguir este propósito: **Blockchain**, el **Reglamento General de Protección de Datos** y el **Esquema Nacional de Seguridad** (dada la relevancia legal del proyecto) con el objeto de realizar una revisión de su estado de arte actual y dejar perfectamente definida su variable teórica.

Queremos significar también que en este apartado del proyecto se pretende únicamente introducir al lector de este documento, en las dimensiones antes señaladas siempre desde un punto de vista genérico, de manera que tenga una visión general sobre cada una de las partes teóricas del proyecto ya que en el capítulo relacionado con el desarrollo de la plataforma **D.One+** se objetivarán las cuestiones de detalle.

8.1 BLOCKCHAIN

Hoy en día, las redes informáticas organizacionales son inconsistentes debido a las ineficientes políticas de almacenamiento y las deficiencias que presenta la gobernanza de sus datos, lo que a largo plazo produce transacciones innecesarias entre cada una de las entidades que procesan esos datos con el objetivo de convertirlos en información. Este problema genera sobrecostes a nivel financiero debido a su mantenimiento y, a su vez, problemas de duplicidad e incapacidad para manejar de manera adecuada los datos maestros. Una solución a este problema es **Blockchain**, que proporciona una tecnología Distributed Ledger que permite a cualquier participante de la red ver un único sistema de registro o **Libro Maestro**, generando así integridad, trazabilidad, confiabilidad y seguridad en todos los datos almacenados y no sólo para el último valor de los mismos, sino también para el histórico completo de los valores de cada registro.

Tradicionalmente, las BBDD han sido utilizadas como repositorios de los datos operativos/centrales de las organizaciones para respaldar el procesamiento y el cálculo de las transacciones que se generan como consecuencia de su actividad. Sin embargo, las BBDD rara vez se comparten entre diferentes empresas y/o administraciones debido a una amplia variedad de dificultades tecnológicas y de seguridad. Blockchain es una BBDD distribuida y replicada entre sus nodos que está diseñada para aumentar la transparencia, la seguridad y la integridad de los datos contenidos en la misma. Su arquitectura según James Scheider (2016) puede describirse de la siguiente forma:

- **Una Base de Datos** (con copias replicadas en múltiples ubicaciones o nodos)
- **de transacciones** (entre dos o más partes)

- **divididas en bloques** (con cada bloque que contiene detalles de la transacción)
- **que son validados por toda la red de nodos**, combinando los detalles comunes de la transacción con las firmas digitales de 2 o más partes. Donde la transacción es válida si el resultado de la codificación es el mismo para todos los nodos, de acuerdo a algoritmos de consenso distribuidos
- **y agregada a la cadena de transacciones anteriores** siempre que el bloque esté validado. Si el bloque no es válido, un "consenso" de nodos corregirá el resultado en el nodo no conforme.

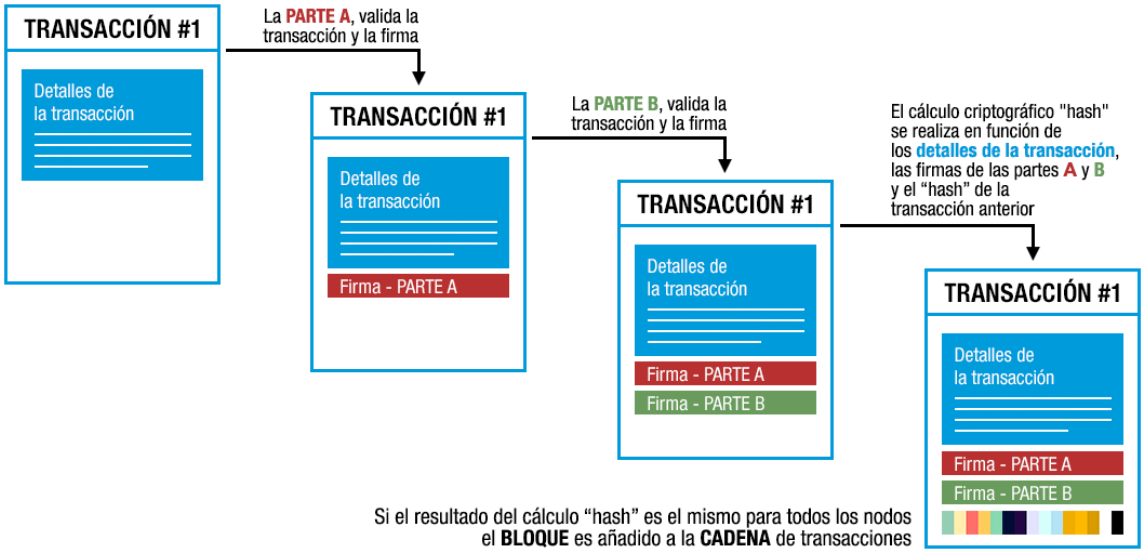


Ilustración 5. Cómo se construye y valida un solo bloque en Blockchain. Fuente: (James Schneider, 2016).

El potencial de Blockchain radica en las propiedades únicas de una BBDD distribuida y en cómo ésta puede mejorar la transparencia, la seguridad y la integridad. Históricamente, la responsabilidad de las BBDD organizacionales creadas como repositorios de registros centrales recaía en su propietario que era el encargado de administrar el acceso, las actualizaciones, limitar la transparencia, la escalabilidad y la capacidad de acceso/manejo de 'elementos externos' a la misma para garantizar que los registros no se manipularan. Una BBDD distribuida que abarcara un conjunto de organizaciones era prácticamente imposible de implantar debido a las limitaciones propias de la tecnología relacional, pero los avances en software, comunicaciones y encriptación ahora permiten hacerlo utilizando soluciones de tipo Blockchain.

Las cadenas de bloques de Blockchain componen un registro central o *Libro Mayor Digital compartido* de transacciones registradas y verificadas a través de una red de participantes en una cadena. A su vez, esta es inviolable y visible entre todos los nodos, donde las variaciones autorizadas o privadas agregan una capa de privilegios para determinar quién puede participar en esa cadena.

La tipología de las redes Blockchain se puede diferenciar en **dos grandes grupos**, por un lado, los que están asociados al **aspecto público o privado**, y, por otro lado, los que se vinculan con la **necesidad o no de tener determinados permisos** de acceso a la red.

En lo relativo al **primer gran grupo**, nos encontramos con los comentados dos posibles tipos, pero se puede incluir un tercero que son los **consorcios**:

- Redes públicas: En este tipo de redes los participantes no tienen por qué conocerse, y cualquier interesado puede unirse a este tipo de redes implantando su propio nodo.
- Redes privadas: En este caso, los interesados deberán **solicitar una autorización** expresa a los gestores de la red, para que se les permita acceder a esta. Dada la necesidad de una acreditación, los usuarios de la misma se conocen, o por lo menos existe una entidad superior que conoce a todos los usuarios.
- Redes consorcio: Estas son **redes privadas que están formadas por más de una organización**, y entre ellas se establece una serie de protocolos y acuerdos, donde el grado de control de la privacidad, es aún más extremo.

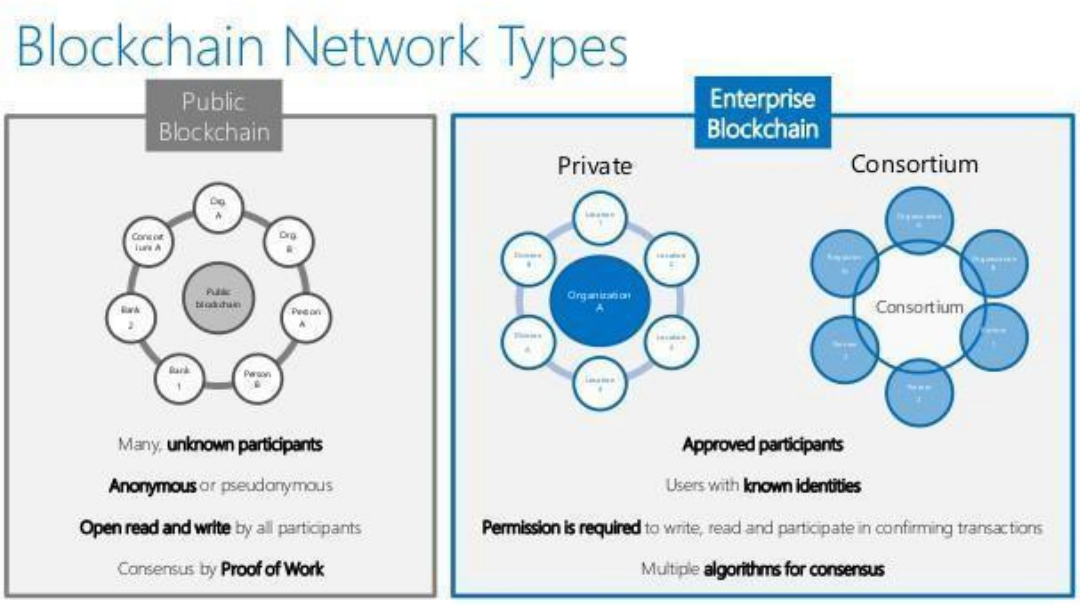


Ilustración 6: Tipos de redes Blockchain. Fuente: Aman Mishra "Unveiling the concept of Blockchain".

Y con respecto al **segundo gran grupo**, **el diferencial serán los permisos de lectura y escritura en la red**, y en función de estos, tenemos **cuatro tipos** de Blockchain:

- No permisionadas: Los interesados en participar en la red, no tienen limitaciones de permisos de lectura y escritura en la red, pudiendo hacer las transacciones que estimen necesarias.
- Permisionadas: Se necesita de permisos de lectura y escritura para poder operar en la red.

- Públicas-permisionadas: Es un modelo de red combinado, que busca unificar lo mejor de las redes públicas y de las privadas. Mejorando los niveles de privacidad y seguridad con la obligatoriedad de tener unos determinados permisos.
- Privada-no permisionadas: Se trata de una red que es accesible a nivel operativo de lectura y escritura, para todos los interesados, pero no todos pueden acceder a la misma.

Casos de Uso de Blockchain

La gran virtud de Blockchain radica en la capacidad de generar bloques de datos inmutables mediante encriptación por hash, lo que genera una capa de seguridad de facto que no es posible en las Bases de Datos relacionales. Además, posee la capacidad (mediante Smart Contracts) de facultar a cada uno de los nodos de la red para aceptar y validar la información a insertar (y/o actualizar) en un nuevo bloque de datos, generando consistencia, unicidad y capacidad de trazabilidad ya que permite registrar cada uno de los movimientos.

Dadas las características principales de Blockchain, esta tecnología se presenta como una solución innovadora e ideal de uso para casos como:

- **Facilitar transacciones seguras y descentralizadas entre muchas partes** debido a que la naturaleza del *Libro Mayor* de Blockchain es efectiva en el manejo de transacciones distribuidas entre un gran número de partes. Además, Blockchain ofrece seguridad para cada transacción debido a la verificación criptográfica realizada por metodología hash. La evolución de los nuevos modelos económicos distribuidos que cubren decenas o incluso cientos de millones de activos (como dispositivos IoT) necesitará de modelos transaccionales distribuidos y seguros para posibilitar y facilitar su desarrollo.
- **Aumento de la transparencia y la eficiencia en las transacciones multiparte**. En cualquier transacción que involucre a dos o más partes, estas operaciones se suelen registrar por separado en cada uno de los sistemas propietarios de esas organizaciones (y de forma independiente de la otra parte) y este hecho provoca la necesidad de crear costosos procesos de conciliación y reconciliación que, incluso actualmente, necesitan una importante intervención manual. Al utilizar tecnología Blockchain, las organizaciones pueden simplificar los procesos de compensación y liquidación, acortando las fechas de operación, y evitando muchos gastos de capital y costes operativos sustanciales.

Respecto a las **AAPP**, el uso de Blockchain puede tener mucha trascendencia en la mejora de los siguientes aspectos:

- **Registros a prueba de manipulaciones**. Los usuarios de una Blockchain pueden reconstruir fácilmente cuándo se produce un cambio en su Libro Maestro, qué información se modificó y en qué parte de la red se originó el cambio (quién).

- **Propiedad digital y transferencia de activos**. Las AAPP podrán prescindir de los documentos en papel y establecer una infraestructura digital *eficiente* para registrar (por ejemplo) la propiedad de los activos mobiliarios e inmobiliarios y proporcionar los medios para transferir fácilmente la información sobre los documentos de venta, escrituras, títulos de propiedad, etc. relacionados con los mismos.

- **Smart Contracts**. Blockchain permite el uso de *código embebido* que posibilita registrar acuerdos entre cada uno de los nodos que participan de ese bloque de datos, lo que asegura de forma automática si un nodo cumple con las condiciones necesarias para ser parte de ese bloque, además de tener la atribución de participar de éste mediante la inclusión de datos. En otras palabras, reemplazan algunas partes de los contratos mediante sentencias escritas en códigos de programación. Esto implica que las reglas son totalmente rígidas y creadas para un uso específico, con lo cual no requieren de más interpretaciones que las propias directrices que describen el código en su programación. Los *Smart Contracts* proporcionan un nivel absoluto de confianza entre cada uno de los nodos que, entre otros aspectos, provoca la reducción drástica de costes, fraudes y automatiza su ejecución ahorrando factores temporales y revisiones innecesarias. Hoy en día, los *Smart Contracts* se han transformado en el core de cada una de las redes de Blockchain, evolucionando con nuevos lenguajes de programación centrados en su creación como son **Solidity en Ethereum** y **Golang en Hyperledger**. Su importancia es tal, que para que un grupo de transacciones sea válido y se genere un nuevo bloque de datos, es necesario que se cumplan todos los Smart Contracts definidos por la misma organización entre todos los nodos que requieren la generación o actualización de nuevos datos.

- **Esquema Nacional de Interoperabilidad**: Blockchain permite el cumplimiento del ENI, el cual posibilita la realización de principios y derechos de los ciudadanos, singularmente el derecho recogido en el artículo 28 de la Ley 39/2015 (de 1 de octubre del Procedimiento Administrativo Común de las Administraciones Públicas) "*a no aportar documentos elaborados por la Administración o entregados anteriormente por el interesado a cualquier Administración*".

- **Marco Legal de Protección de Datos vigente y Esquema Nacional de Seguridad**: Blockchain permite el cumplimiento del Reglamento General de Protección de Datos. La Blockchain permisionada Hyperledger Fabric permite dar cumplimiento al marco legal de Protección de Datos vigente, así como a los principios relativos a la privacidad, confidencialidad e integridad y al ejercicio de derechos. Tal y como estipula la Disposición adicional primera de la Ley Orgánica 3/2018 de Protección de Datos y Garantía de Derechos Digitales, el Esquema Nacional de Seguridad, incluirá las medidas que deban implantarse en caso de tratamiento de datos personales, para evitar su pérdida, alteración o acceso no autorizado, adaptando los criterios de determinación del riesgo en el tratamiento de los datos a lo establecido en el artículo 32 del Reglamento

(UE) 2016/679. En el caso de las AAPP, la aplicación de las medidas de seguridad estará marcada por los criterios establecidos en el ENS. Por tanto, la elección de la tecnología Blockchain permite el cumplimiento del RGPD y ENS.

Al respecto del objeto de este Trabajo, una tarea importante a realizar por las AAPP es **mantener información confiable sobre individuos, organizaciones, activos y actividades**. Los gobiernos locales, provinciales, autonómicos y nacionales se encargan de mantener registros que incluyen, por ejemplo: fechas de nacimiento y fallecimiento, información sobre el estado civil, licencias comerciales, transferencias de propiedad, actividades delictivas, etc. Con lo cual, administrar y usar estos datos puede ser complicado, incluso para las AAPP más avanzadas: algunos contratos existen sólo en papel y si los cambios deben hacerse en los registros oficiales, los ciudadanos, a menudo, deben presentarse en persona para hacerlo. Las Administraciones tienden a construir sus propios silos de registros y crean protocolos distintos para la gestión de la información, lo que impide que otras partes de la AAPP los utilicen. Y, por supuesto, estos datos deben estar protegidos contra el acceso o la manipulación no autorizados, sin margen de error.

Blockchain simplifica la administración de la información confiable, facilitando a las entidades gubernamentales el acceso y el uso de datos críticos del sector público, al tiempo que mantiene la seguridad de esta información. Tal y como señalamos con anterioridad, Blockchain es un *Libro Maestro Digital Codificado* que se almacena en múltiples computadoras en una red pública o privada y que comprende registros de datos o 'bloques'. Una vez que estos bloques se recopilan en una cadena, no pueden ser modificados o eliminados por un sólo actor; ya que, en su lugar, se verifican y administran mediante la automatización y protocolos de gobernanza de datos compartida.

Hasta ahora los bancos, los proveedores de servicios de pago y las compañías de seguros han mostrado el mayor nivel de interés e inversión en Blockchain, pero creemos que las AAPP tiene mucho que ganar con esta tecnología desplegándola estratégicamente a través de proyectos como el nuestro. Con el tiempo, Blockchain puede ayudar a digitalizar registros existentes y administrarlos dentro de una infraestructura segura permitiendo realizar algunos de estos registros 'inteligentes'. Los departamentos de IT de las agencias gubernamentales pueden crear reglas y algoritmos que permitan, por ejemplo, que los datos contenidos en un Blockchain puedan ser compartidos automáticamente con terceros una vez que se cumplan las condiciones predefinidas. A más largo plazo, la tecnología puede incluso permitir que los propios individuos y empresas obtengan el control directo sobre toda la información que el gobierno mantiene sobre ellos. Este nivel de transparencia podría, a su vez, facilitar a la AAPP la aceptación para la creación de servicios públicos en red de cualquier tipo con la máxima seguridad y operatividad.

En definitiva, para las organizaciones Blockchain tiene las siguientes ventajas:

- Reduce costes: Gastos generales y/o gastos provocados por intermediarios en las transacciones.

- Minora los riesgos respecto a la manipulación de los datos, el fraude y el crimen cibernético. La ciberseguridad es una preocupación común para todas las organizaciones y más aún para los datos personales contenidos en una AAPP. Un MDM debe estar encriptado y trabajar de acuerdo a unas directrices claras de política de protección de datos y la información procedente de Blockchain proporcionará exactitud junto a seguridad, siempre que el conjunto de atributos confidenciales a los que accede esté encriptado y enmascarado. La integración de MDM y Blockchain cubre, por lo tanto, la seguridad de los datos contenidos en el sistema.
- Aumenta la confianza a través de procesos compartidos y el mantenimiento inmutable y resistente a ataques de los registros.

Por qué utilizar Blockchain en un MDM

En la implementación de un MDM la distribución de sus registros maestros suele ser la parte más laboriosa y difícil de gestionar. Es precisamente en este aspecto, donde una Blockchain elimina de facto los costes generales y la falta de confiabilidad de las transacciones autenticadas por los socios de una red P2P que involucren el intercambio de datos y, por lo tanto, puede soportar uno de los grandes desafíos de este tipo de soluciones como es "*la sincronización bidireccional gobernada de los datos maestros entre Blockchain y el MDM*".

McKnight (2018) señala que otro desafío central para un MDM es llegar a la "*versión única de la verdad*", lo cual es difícil de alcanzar incluso con este tipo de soluciones porque, en primer lugar, todos deben aceptar tácitamente el proceso utilizado para crear el registro maestro de una entidad (si bien muchos analistas de MDM hacen todo lo posible para utilizar las reglas de datos de un proceso de gobernanza, éste sigue siendo un proceso sujeto a crítica). El consenso que puede lograr Blockchain es un proxy de gobernanza para esa "*versión única de la verdad*" al lograr el consenso para la confianza, así como la trazabilidad completa de los datos contenidos.

Por lo tanto, parece que la tecnología Blockchain podría abordar los principales desafíos en MDM. Además, podemos señalar lo siguiente:

- *Blockchain vs Base de Datos Relacional*: Con una Base de Datos relacional crear integraciones únicas con cada una de las distintas organizaciones es costoso y no es escalable. Con Blockchain cada vez que se agrega un nodo a la red (una AAPP, en nuestro caso) es un proceso uniforme para todos y lo hace menos costoso. Esto se debe a que, en lugar de crear integraciones, sólo se agregan nodos a los cuales únicamente sus participantes pueden acceder a su contenido siempre con los permisos correspondientes.
- *Confiable y Protegido*: Toda la red Blockchain (y a la vez cada uno de los nodos que la conforman) funciona como un mecanismo de verificación automática (lo que no permite margen de error), mientras que en una solución tradicional el control y la información está en manos de un administrador central.

- *Inmutabilidad*: una vez que se guardan los datos, no se pueden modificar y son inmutables. Eso proporciona un nivel de confianza en la tecnología que no ofrece una base de datos tradicional.

Blockchain, a diferencia de un concentrador centralizado, proporciona una base de datos distribuida que puede almacenar datos totalmente certificados y a perpetuidad ya que al almacenar bloques vinculados y con marca de tiempo esta cadena es inalterable y permanente.

Por todo ello, el desarrollo/evolución de los MDM construidos sobre Blockchain debe comenzar en nichos de negocio que demanden estos rasgos: datos financieros, de seguros y gubernamentales (como es nuestro caso). Blockchain es ahora un vector de disrupción para los MDM y los desarrolladores de estas soluciones, por tanto, deben ser al menos conscientes de la importancia que puede llegar a tener en su negocio la tecnología Blockchain para crear en un futuro cercano la integración entre las dos tecnologías, y que en el caso de aquello que no asuman la necesidad de un cambio de mentalidad, se verán abocados a perder terreno frente a sus competidores.

8.2 REGLAMENTO GENERAL DE PROTECCIÓN DE DATOS Y ESQUEMA NACIONAL DE SEGURIDAD

La solución **D.One+** aglutina diferentes tecnologías y técnicas de sistemas de información que habrá que analizar cautelosamente para poder alinear los objetivos de la misma con el Reglamento General de Protección de Datos (RGPD en adelante) que conforma el marco regulatorio europeo al que nos tendremos de ceñir y dar cumplimiento. El adecuado cumplimiento del RGPD, supone un hito fundamental para poder asegurar la viabilidad de la solución D.One. Por ello, se ha decidido realizar un detenido repaso y análisis de la normativa actual de protección de datos y de los aspectos más relevantes que se han tenido en cuenta desde las primeras fases de este proyecto.

No sólo deben tenerse en cuenta los datos personales que serán tratados en la solución, si no la forma de tratamiento, la plataforma de tratamiento, dónde se alojarán, quién tendrá acceso a los mismos, cómo podrán ejercer sus derechos los interesados, cómo se llevarán a cabo los procesos de anonimización. Éstas son algunas de las cuestiones que se deben abordar para poder alinear la solución D.One con el RGPD.

Partiremos de la base que se realizará el tratamiento de los siguientes **datos personales**:

1. DNI/NIF/CIF/Nº Tarjeta de Residencia/Nº de Pasaporte y sus variantes.
2. Nombre/Razón Social.
3. Apellido 1.
4. Apellido 2.
5. Dirección postal. Compuesta por los siguientes campos:
    a. Código Calle.

b. Tipo Vía.

c. Literal Vía.

d. Número.

e. Letra.

f. Escalera.

g. Bloque.

h. Planta.

i. Puerta.

j. Código Postal.

k. Población.

l. Provincia.

m. País.

6. Dirección electrónica.
7. Teléfono fijo, teléfono móvil.
8. Fecha de Fallecimiento Persona Física.

Además, debemos comentar que la solución **D.One+** se desarrollará utilizando la tecnología "*Open Source*" **Hyperledger Fabric**, para el despliegue de una **red Blockchain privada y permisionada,** y la misma deberá adecuarse también al RGPD por lo tanto deberemos realizar este análisis teniendo en cuenta esta característica.

Estado del Arte

El Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (en adelante, **Reglamento General de Protección de Datos** o RGPD), publicado en el Diario Oficial de la Unión Europea del 04 de mayo de 2016 y en vigor desde el pasado 25 de mayo de 2018, tal y cómo figura en su artículo 1, fija las normas relativas a la protección de las personas físicas en lo que respecta al tratamiento de los datos personales y las normas relativas a la libre circulación de dichos datos. Además, protege los derechos y libertades fundamentales de las personas físicas y, en particular, sus derechos a la protección de los datos personales.

El RGPD es una norma directamente aplicable, que no requiere de regulación interna de transposición ni tampoco de normas de desarrollo o aplicación. Por ello, los responsables deben ante todo asumir que la norma de referencia es el RGPD y no las normas nacionales, como venía sucediendo hasta ahora con la Directiva 95/46.

La norma referente en España en materia de protección de daños se denomina Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (en adelante, LOPDPGDD), publicada en el Boletín Oficial del Estado el pasado 06 de diciembre de 2018 y en vigor desde el 07 de diciembre del mismo año; esta norma derogó, sustituyéndola, **Ley Orgánica de Protección de Datos y Garantía de Derechos Digitales (LOPDGDD)**. Esta ley entró en vigor el 6 de diciembre de 2018, sustituyendo a la antigua Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal. El objetivo de la LOPDGDD es adaptar la legislación española a la normativa europea del RGPD.

Una de las medidas más reiteradas dentro del texto y que el RGPD establece como un requisito legal de cumplimiento es la **protección de datos desde el diseño de un sistema**, señalando que el incumplimiento de esta obligación es sancionable, tal y como se refleja en su artículo 83 (al igual que su correcta aplicación constituye uno de los criterios para baremar la gravedad de una infracción).

En nuestra condición de prestadores de servicios, desde D.One estableceremos una relación de **Encargados del Tratamiento** con aquellas empresas u organizaciones, de carácter público o privado, que asumen el papel de Responsables de Tratamiento. De acuerdo con el artículo 28 del Reglamento, "*cuando se vaya a realizar un tratamiento por cuenta de un responsable del tratamiento, este elegirá únicamente un encargado que ofrezca garantías suficientes para aplicar medidas técnicas y organizativas apropiados, de manera que el tratamiento sea conforme con los requisitos del presente Reglamento y garantice la protección de los derechos del interesado*".

Tal y como estipula su artículo 25.1 "*el Responsable del Tratamiento aplicará, tanto en el momento de determinar los medios de tratamiento como en el momento del propio tratamiento, medidas técnicas y organizativas apropiadas, concebidas para aplicar de forma efectiva los principios de protección de datos*". Si bien el cumplimiento de esta obligación aplica concretamente al Responsable de Tratamiento, a tenor del considerando 78 y lo establecido en el artículo 28 del RGPD, la protección de datos desde el diseño se irradia sobre los actores implicados en el tratamiento de datos personales como son los Proveedores, los Prestadores de servicios y los Desarrolladores de productos y aplicaciones. A éstos, el Responsable ha de alentar a "*que tengan en cuenta el derecho a la protección de datos cuando se desarrollen y diseñen estos productos servicios y aplicaciones*" y para ello, debe seleccionar a aquel "***encargado de tratamiento*** *que permita ofrecer las garantías suficientes para aplicar medidas técnicas y organizativas apropiados*". Por ello, el Responsable del Tratamiento es quien debe ceñirse a la selección de servicios y de encargados capaces de asegurar el cumplimiento de los requisitos del RGPD.

De acuerdo al artículo 37.1 del RGPD, se establecen los casos en los que el Responsable y el Encargado del Tratamiento designarán un Delegado de Protección de Datos (DPD, a partir de ahora). De esta manera, en su apartado b) se especifica uno de los casos en los que se debe nombrar la figura del DPD: "*cuando las actividades principales del responsable o del encargado consistan en operaciones de tratamiento que, en razón de su naturaleza, alcance y/o fines,*

requieran una observación habitual y sistemática de interesados a gran escala". Debido al principio de funcionamiento de la solución D.One, el tratamiento de datos cumple los criterios de tratamiento "*habitual y sistemático*" y en, su caso, podrá ser considerado como de "*gran escala*".

Tal y como lo señala la Agencia Española de Protección de Datos (a partir de ahora, AEPD) "*El artículo 37 se aplica tanto a los Responsables del tratamiento como a los Encargados del tratamiento con respecto a la designación de un Delegado de Protección de Datos (DPD). En función de quién cumpla los criterios de designación obligatoria, en algunos casos solo el Responsable o sólo el Encargado deben designar un DPD, y en otros casos tanto el Responsable como su Encargado deben designar respectivos DPD (que deberán cooperar entre sí).*"

Es importante destacar que, aunque el Responsable cumpla con los criterios de designación obligatoria, su Encargado no está necesariamente obligado a nombrar un DPD. No obstante, puede ser una práctica recomendable. Por ello, acorde al principio de "accountability" y "privacidad desde el diseño y por defecto" tenidos en cuenta desde la concepción de la solución, se considerará conveniente y de buena práctica el nombramiento de un DPD dentro del equipo del proyecto y es que para nosotros también la privacidad desde el diseño y por defecto constituyen una pieza clave y fundamental para garantizar el cumplimiento de la normativa de protección de datos.

Dado el contexto tecnológico en el que se encuadra la solución **D.One+**, cuyo funcionamiento radica en el uso de tecnologías disruptivas y que por su objetivo realizan un continuo tratamiento de datos, el impacto sobre la privacidad de los mismos y la necesidad de cumplimiento del artículo 25 del RGPD se vuelve una necesidad para asegurar la viabilidad del proyecto. En estos términos, es requisito fundamental la implantación de medidas técnicas y organizativas efectivas que aseguren el respeto a los derechos y libertades de las personas en lo que a su tratamiento de datos personales se refiere.

La privacidad desde el diseño y por defecto implica, por lo tanto, dar un enfoque orientado a la gestión del riesgo y a la responsabilidad proactiva que permita establecer los requisitos de privacidad del sistema.

Reglamento General de Protección de Datos y Blockchain

No existe una guía para el cumplimiento normativo del RGPD y la tecnología Blockchain. Para poder conocer el estado legal del RGPD es necesario evaluar cada caso de uso y su aplicación. Sin embargo, las redes privadas permisionadas suponen un menor desafío para su adecuación al Reglamento.

En líneas generales podemos señalar que existen claras tensiones entre el RGPD y Blockchain, donde éstas se pueden englobar en 3 aspectos:

1. **Responsabilidad y roles**. La identificación y las obligaciones del Responsable del Tratamiento y del Encargado del Tratamiento cuando hay múltiples participantes en una red Blockchain involucrados en el tratamiento de los datos contenidos en la misma. El

Participante que tenga el derecho a escribir sobre la cadena de bloques y decida enviar datos para su validación por parte de los mineros, se consideran Responsables de Tratamiento**.** Por tanto, se puede considerar a una participante como **Responsable de Tratamiento** cuando:

a. El Participante es una persona natural y la operación de tratamiento de datos personales está relacionada con una actividad profesional o comercial.

b. El Participante es una persona legal y registra datos personales en la cadena de bloques.

Es necesario tener en cuenta que no todos los Participantes de la red Blockchain serán Responsables de tratamiento. Aquellas personas naturales que introduzcan datos personales en la red Blockchain, que no pertenecen a una actividad profesional o comercial, no son considerados Responsables de Tratamiento. Este sería el caso de los **mineros**, que sólo se encargan de validar las transacciones enviadas por los participantes y no están involucrados en el objeto de dicha transacción.

Si un grupo de participantes decide llevar a cabo tratamientos de datos con un propósito común, se debe identificar de antemano al Responsable de Tratamiento. De otra forma, dichos participantes se consideran **Corresponsables de Tratamiento**, según el artículo 26 del RGPD y deberán, por tanto, determinar sus respectivas responsabilidades para asegurar el cumplimiento normativo.

En cuanto a los ***Smart Contracts***, o cualquier software, el desarrollador del mismo puede simplemente ser un Proveedor de soluciones o, cuando dicho desarrollador participe en el procesamiento, puede ser considerado como Encargado de tratamiento o Responsable de Tratamiento dependiendo de su rol en la determinación de los propósitos del tratamiento.

En Blockchain el **Encargado de Tratamiento**, en términos del RGPD, puede ser:

- El desarrollador del *Smart Contact* que procesa los datos personales en nombre del participante, considerado Responsable de Tratamiento.
- Los mineros que validan la transacción que contiene datos personales en la Blockchain.

2. **Minimización y anonimización de datos.** Anonimizar los datos personales supone un gran debate para Blockchain. La función Hash empleada se considera una técnica de seudonimización. El principio de minimización de datos requiere que los datos tratados sean relevantes y limitados a aquello que es estrictamente necesario acorde a los propósitos por los cuales son objeto de tratamiento. Según el artículo 25 del RGPD, el principio de protección de datos desde el diseño requiere que el Responsable de Tratamiento elija el formato con el menor impacto posible sobre los derechos y libertades de los interesados. Si es necesario subir un dato a Blockchain, es preferible que se haga

en forma de "*Commitment*" criptográfico, y si no es posible, debe ser registrado usando la función Hash con una clave, asegurando un cierto nivel de confidencialidad.

En este punto, se deberá pensar si la solución D.One almacenará los datos "*on-chain*" o por el contrario fuera de la Blockchain y subiendo simplemente una prueba de la existencia de dicho dato.

Si el propósito de tratamiento de datos es justificado y tras una **Evaluación de Impacto de Protección de Datos** (en adelante, EIPD) se prueba que el nivel de riesgo residual es aceptable, los datos personales pueden ser, excepcionalmente, almacenados en la Blockchain.

La **CNIL** (Comisión Nacional de la informática y las Libertades) la autoridad de control en materia de protección de datos en Francia realiza una serie de aclaraciones respecto a la minimización de datos:

- Dado que los identificadores de los participantes, como pueden ser las claves públicas, son propiedades inherentes al funcionamiento de Blockchain y por tanto no es posible minimizarlas más.
- Con respecto a datos personales adicionales, en relación al cumplimiento normativo, el CNIL recomienda aplicar soluciones en que los datos sean tratados fuera de la Blockchain o bien, en los casos en que sean almacenados en la Blockchain, en el siguiente orden de preferencia:
    - "Commitment" del dato.
    - Hash generado a partir una función hash con clave sobre el dato.
    - Texto cifrado del dato.
- Si ninguna de las soluciones mencionadas previamente se puede implementar, y cuando el propósito del tratamiento sea justificado, y la EIPD haya probado que el nivel de riesgo residual es aceptable, entonces el dato podrá almacenarse, bien usando la función hash sin clave o en ausencia de otras posibilidades en texto claro.

3. **Derechos y obligaciones**. El capítulo III del RGPD (art. 12 a 23) nos introduce en la materia de los derechos que la normativa vigente otorga a los interesados, en referencia a sus datos personales: derechos tan variados como el de acceso (art. 15), de rectificación (art. 16), de supresión o "derecho al olvido" (art. 17), de limitación del tratamiento o "marcado de datos" (art. 18), a la portabilidad de los datos (art. 20) y de oposición (art. 21), además del derecho a ser informado de la violación de seguridad de los datos personales (art. 34);para ello, el art. 12 de la LOPDPGGD establece que "los derechos reconocidos en los artículos 15 a 22 del RGPD podrán ejercerse directamente o por medio de representante legal o voluntario […] el encargado podrá tramitar, por cuenta del responsable, las solicitudes de ejercicio formuladas por los afectados de sus derechos si así se estableciera en el contrato o acto jurídico que les vincule". Una de las

finalidades contempladas en los “considerando” del RGPD, por el que este texto ha sido aprobado, es devolver a los interesados el control de la información personal y reforzar el ejercicio de derechos de los mismos sobre quienes procesan sus datos. Además de la minimización de riesgos, como se ha comentado previamente, el formato con el que se registran los datos en la Blockchain puede facilitar el ejercicio de los derechos de los interesados. A pesar de que el ejercicio de ciertos derechos no supone un problema, aplicar el **derecho de supresión, de rectificación o de oposició**n en una Blockchain es preferible considerarlos habiendo hecho una profunda reflexión. Es técnicamente imposible garantizar la solicitud de supresión realizada por un interesado una vez el dato se haya registrado en la Blockchain. Sin embargo, cuando el dato registrado es un “Commitment”, hash generado por una función hash con clave o texto cifrado obtenido por los diferentes algoritmos y criptografías, el Responsable de Tratamiento puede hacer el dato prácticamente inaccesible y por tanto, aproximarse a los efectos que tendría la supresión del dato.

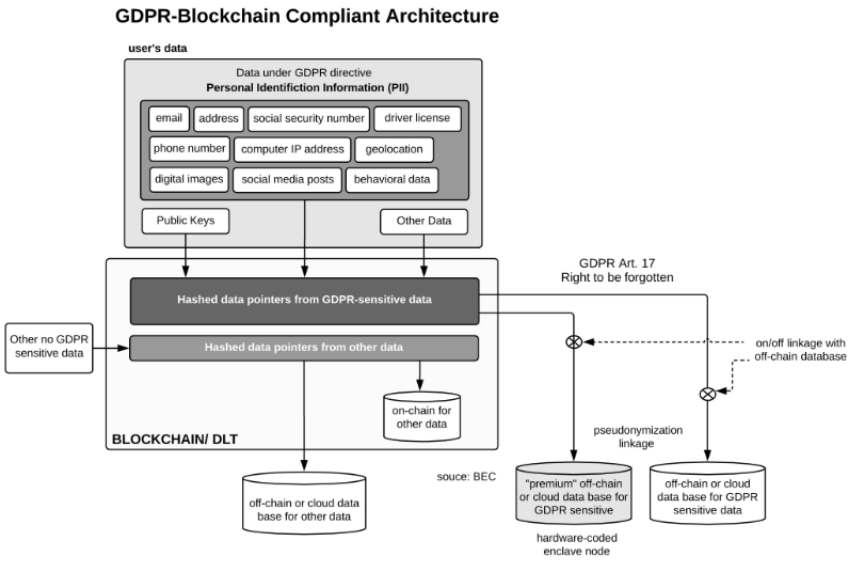


Ilustración 7. Arquitectura tipo para el cumplimiento RGPD-Blockchain.

4. **Encaje del RGPD en Blockchain.** Dentro de la solución creada con el proyecto **D.One+**, en línea con las especificaciones técnicas de la arquitectura Hyperledger Fabric versión 2.2 y cuyas características serán descritas más detalladamente en el apartado “Estructura de Hyperledger Fabric”, se ha optado por aplicar la solución denominada “**Colecciones de datos privados**”, implementada de forma nativa en HLF desde su versión 1.2, según el cual dichas colecciones están compuestas por datos privados, compartidos entre las organizaciones y almacenados en una base de datos privada[1] de la organización, también llamada en ocasiones “SideDB”, y el *hash* de dichos datos privados, accesible para todos los participantes del canal. El hash representa la

[1] *On Blockchains and the General Data Protection Regulation*, de Luis-Daniel Ibáñez, Kieron O’Hara, y Elena Simperl, Universidad de Southampton.

evidencia y la inmutabilidad de la transacción, evitando que dicha información personal pueda ser manipulada ya que la manipulación modificaría necesariamente también dicho hash. En caso de necesidad, una organización podría compartir los datos privados con un tercero que necesite comprobar la validez de dichos datos, y solo necesitaría calcular el hash de los datos recibidos y comprobar la coincidencia. Es preciso apuntar que los datos privados **no viajan**[2] con el resto de los datos de la transacción y, por ende, no son visibles por el servicio de ordenación ("*Ordering Service"*). De especial relevancia, en términos de cumplimiento de las previsiones del RGPD en materia de derechos de los interesados, es que con la solución empleada en **D.One+** uno de los derechos que se verán reforzados será el derecho a la supresión (o "derecho al olvido" del art. 17 RGPD), a través del cual los interesado pueden solicitar el borrado de sus datos personales almacenados: la solicitud legitima del interesado a la supresión de sus datos personales se llevará a cabo a través del **borrado manual de los mismos de la base de datos privada**, en cualquier momento.

Esquema Nacional de Seguridad

Tal y como se ha explicado al principio de la documentación, el objetivo de **D.One+** es la creación de un repositorio de registros maestros de personas para la AAPP como base común de referencia que unifique, facilite y simplifique su intercambio entre los diferentes sistemas de gestión, aumentando la coherencia, confiabilidad, integridad, exactitud y calidad de los datos. De esta manera, con **D.One+**, no sólo se solventa el problema de la existencia de registros duplicados e inconsistentes, sino que supone una plataforma que da cumplimiento a normativas vigentes para las AAPP que garantizan derechos fundamentales a los ciudadanos.

La Ley 11/2007, de 22 de junio, de Acceso Electrónico de los Ciudadanos a los Servicios Públicos (en adelante, LAECSP) derogada por la Ley 39/2015, reconoce a las personas su derecho a relacionarse electrónicamente con las AAPP, así como la obligación de éstas a garantizar ese derecho. La Ley 11/2007, en su artículo 42.2 crea el Esquema Nacional de Seguridad (en adelante, ENS), que establece las condiciones necesarias de confianza en el uso de los medios electrónicos, a través de medidas para garantizar la seguridad de los sistemas, los datos, las comunicaciones y los servicios electrónicos. Posteriormente, la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, recoge el Esquema Nacional de Seguridad en su artículo 156, apartado 2 en similares términos.

El Real Decreto 3/2010, en cumplimiento a lo que dispuso en su momento la LAECSP y de lo que ha recogido el texto de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (LRJSP), regula una de las piezas fundamentales que vertebran lo que se ha dado en llamar la Administración Electrónica: la seguridad de los sistemas de información del Sector Público.

[2] De acuerdo con las recomendaciones contenidas en el dossier "Blockchain and the GDPR" publicado por el *European Union Blockchain Observatory and Forum* en su web euBlockchainforum.eu

Los ciudadanos, profesionales y empresas españolas, beneficiarios últimos de lo exigido en el ENS, han observado en estos años una significativa mejora en la disponibilidad de los servicios prestados por vía electrónica, constatando el esfuerzo de los poderes públicos por dotar al tratamiento de la información de las debidas garantías de seguridad y legalidad que cualquier acto de las Administraciones Públicas exige. Constituyendo una norma de obligado cumplimiento, y ante el importante volumen de entidades que conforman el sector público, conviene precisar el ámbito de aplicación del ENS muy especialmente tras la entrada en vigor de la Ley 39/2015, de 1 de octubre, de Procedimiento Administrativo Común de las Administraciones Públicas y de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

El Centro Criptológico Nacional (en adelante, CCN) afirma que el ENS es de aplicación, entre otros, para los siguientes casos:

- Sistemas de Información accesibles electrónicamente por los ciudadanos.
- Sistemas de información para recabar información y estado del procedimiento administrativos.
- Sistemas de Información para el cumplimiento de deberes.

El MDM desarrollado por D.One, es una solución que ofrece tanto a las AAPP vinculadas o dependientes, como al conjunto de la ciudadanía, un sistema de información que permite el acceso a una serie de datos estipulados y al conocimiento del estado de un procedimiento determinado, cumpliendo con la normativa.

Además, gracias a su desarrollo tecnológico, **D.One+** permite la interrelación entre AAPP. El CCN, también establece como ámbito de aplicación las relaciones entre las distintas Administraciones Públicas.

El CCN recuerda la "*responsabilidad de la Administración contratante de la suscripción del correspondiente contrato de prestación del servicio, que deberá contener todas las estipulaciones necesarias para dar cumplimiento a lo dispuesto en el ENS, en virtud de la naturaleza del servicio prestado*".

Con lo descrito anteriormente, queda convenientemente clarificado que **D.One+** entra dentro del ámbito de aplicación del ENS. Y, por tanto, la implementación de esta solución teniendo en cuenta desde su concepción e inicio del diseño el ENS, facilita significativamente el plan operativo y la puesta en marcha de **D.One+**.

La implantación de **D.One+**, tendrá como primer cliente el **Ayuntamiento 1 de Gran Canaria** (municipio situado en el sur de Gran Canaria) y cuyas cifras oficiales de población resultantes de la revisión del Padrón municipal a 1 de enero, según el INE, ascienden a un total de **53.443 habitantes**.

De acuerdo con el principio de proporcionalidad y para facilitar la conformidad con el ENS a determinadas entidades, se podrá implementar perfiles de cumplimiento específicos que comprenderán un conjunto de medidas de seguridad que resulten de aplicación para una categoría concreta de seguridad.

Por ello, tras la realización del estudio de las necesidades de seguridad, recursos y tras un análisis de riesgos contemplando las vulnerabilidades y amenazas a las que se ven expuestas las Entidades Locales y, particularmente, los Ayuntamientos, con el objetivo de garantizar la máxima seguridad de los sistemas de información, se da cumplimiento al mandato impuesto al CCN validando el siguiente Perfil de Cumplimiento Específico para Entidades Locales que permita la implantación del ENS en las mismas, con necesidades de categoría **MEDIA.**

Este Perfil de Cumplimiento Específico podrá ser de aplicación a todos aquellos Ayuntamientos cuya población sea mayor de 20.000 y menor de 75.000 habitantes, estableciéndose los siguientes supuestos (donde el segundo de ellos compete a **D.One+**):

- Servicios alojados en el Ayuntamiento, donde el Perfil de Cumplimiento Específico será de aplicación al sistema de información del Ayuntamiento.
- **Servicios del Ayuntamiento externalizados en la modalidad Software como Servicio (SaaS)**. Que, en este caso, será necesario que el sistema de información que soporta los servicios externalizados de la conformidad en **categoría MEDIA**, siendo de aplicación el Perfil de Cumplimiento Específico al Sistema de Información del Ayuntamiento desde el que se accede a los servicios.

La declaración es un conjunto de medidas que son de aplicación para el cumplimiento del ENS. Dicho conjunto, depende de los niveles asociados a las dimensiones de seguridad. Se determina, por lo tanto, que para garantizar la seguridad en los sistemas a los que hace referencia este Perfil de Cumplimiento Específico, la relación de medidas que son de aplicación y la exigencia en el nivel de seguridad de cada medida aplicada, es la que se indica en la siguiente tabla. Que de las 75 medidas de seguridad definidas en el Anexo II del RD 3/2010, aplican un total de 58, aproximadamente, para los **servicios del Ayuntamiento externalizados**:

SISTEMA AYUNTAMIENTO	SISTEMA EXTERNALIZADO
Marco Organizativo (4): [org.1] Política de seguridad [org.2] Normativa de seguridad [org.3] Procedimientos de seguridad [org.4] Proceso de autorización **Marco Operacional (26):** [op.pl] Planificación [op.pl.1] Análisis de riesgos [op.pl.2] Arquitectura de seguridad [op.pl.3] Adquisición de nuevos componentes [op.pl.4] Dimensionamiento / Gestión de capacidades [op.acc] Control de acceso	**Marco Organizativo (4):** [org.1] Política de seguridad [org.2] Normativa de seguridad [org.3] Procedimientos de seguridad [org.4] Proceso de autorización **Marco Operacional (25):** [op.pl] Planificación [op.pl.1] Análisis de riesgos [op.pl.2] Arquitectura de seguridad [op.pl.3] Adquisición de nuevos componentes [op.acc] Control de acceso [op.acc.1] Identificación [op.acc.2] Requisitos de acceso

SISTEMA AYUNTAMIENTO	SISTEMA EXTERNALIZADO
[op.acc.1] Identificación [op.acc.2] Requisitos de acceso [op.acc.3] Segregación de funciones y tareas [op.acc.4] Proceso de gestión de derechos de acceso [op.acc.5] Mecanismo de autenticación [op.acc.6] Acceso local (local logon) [op.acc.7] Acceso remoto (remote login) [op.exp] Explotación [op.exp.1] Inventario de activos [op.exp.2] Configuración de seguridad [op.exp.3] Gestión de la configuración [op.exp.4] Mantenimiento [op.exp.5] Gestión de cambios [op.exp.6] Protección frente a código dañino [op.exp.7] Gestión de incidentes [op.exp.8] Registro de la actividad de los usuarios [op.exp.9] Registro de la gestión de incidentes [op.exp.11]Protección de claves criptográficas [op.ext] Servicios externos [op.ext.1] Contratación y acuerdos de nivel de servicio [op.ext.2] Gestión diaria [op.com] Continuidad del servicio [op.com.1] Análisis de impacto [op.mon] Monitorización del sistema [op.mon.1] Detección de intrusión [op.mon.2] Sistema de métricas **Medidas de Protección (34)** [mp.if] Protección de las instalaciones e infraestructuras [mp.if.1] Áreas separadas y con control de acceso [mp.if.2] Identificación de las personas [mp.if.3] Acondicionamiento de los locales [mp.if.4] Energía eléctrica [mp.if.5] Protección frente a incendios [mp.if.6] Protección frente a inundaciones [mp.if.7] Registro de entrada y salida de equipamiento [mp.per] Gestión del personal [mp.per.1] Caracterización del puesto de trabajo [mp.per.2] Deberes y obligaciones [mp.per.3] Concienciación [mp.per.4] Formación [mp.eq] Protección de los equipos [mp.eq.1] Puesto de trabajo despejado [mp.eq.2] Bloqueo de puesto de trabajo [mp.eq.3] Protección de equipos portátiles	[op.acc.4] Proceso de gestión de derechos de acceso [op.acc.5] Mecanismo de autenticación [op.acc.6] Acceso local (local logon) [op.acc.7] Acceso remoto (remote login) [op.exp] Explotación [op.exp.1] Inventario de activos [op.exp.2] Configuración de seguridad [op.exp.3] Gestión de la configuración [op.exp.4] Mantenimiento [op.exp.5] Gestión de cambios [op.exp.6] Protección frente a código dañino [op.exp.7] Gestión de incidentes [op.exp.8] Registro de la actividad de los usuarios [op.exp.9] Registro de la gestión de incidentes [op.exp.11] Protección de claves criptográficas [op.ext] Servicios externos [op.ext.1] Contratación y acuerdos de nivel de servicio [op.ext.2] Gestión diaria [op.com] Continuidad del servicio [op.com.1] Análisis de impacto [op.mon] Monitorización del sistema [op.mon.1] Detección de intrusión [op.mon.2] Sistema de métricas **Medidas de Protección (29)** [mp.if] Protección de las instalaciones e infraestructuras [mp.if.1] Áreas separadas y con control de acceso [mp.if.2] Identificación de las personas [mp.if.3] Acondicionamiento de los locales [mp.if.4] Energía eléctrica [mp.if.5] Protección frente a incendios [mp.if.7] Registro de entrada y salida de equipamiento [mp.per] Gestión del personal [mp.per.1] Caracterización del puesto de trabajo [mp.per.2] Deberes y obligaciones [mp.per.3] Concienciación [mp.per.4] Formación [mp.eq] Protección de los equipos [mp.eq.1] Puesto de trabajo despejado [mp.eq.2] Bloqueo de puesto de trabajo [mp.eq.3] Protección de equipos portátiles

SISTEMA AYUNTAMIENTO	SISTEMA EXTERNALIZADO
[mp.com] Protección de las comunicaciones [mp.com.1] Perímetro seguro [mp.com.2] Protección de la confidencialidad [mp.com.3] Protección de la autenticidad y de la integridad [mp.com.4] Segregación de redes [mp.si] Protección de los soportes de información [mp.si.1] Etiquetado [mp.si.2] Criptografía [mp.si.3] Custodia [mp.si.4] Transporte [mp.si.5] Borrado y destrucción [mp.sw] Protección de las aplicaciones informáticas [mp.sw.1] Desarrollo [mp.sw.2] Aceptación y puesta en servicio [mp.info] Protección de la información [mp.info.1] Datos de carácter personal [mp.info.2] Calificación de la información [mp.info.4] Firma electrónica [mp.info.5] Sellos de tiempo [mp.info.6] Limpieza de documentos [mp.info.9] Copias de seguridad (backup) [mp.s] Protección de los servicios [mp.si.1] Protección del correo electrónico [mp.s.2] Protección de servicios y aplicaciones web	[mp.com] Protección de las comunicaciones [mp.com.1] Perímetro seguro [mp.com.2] Protección de la confidencialidad [mp.com.3] Protección de la autenticidad y de la integridad [mp.com.4] Segregación de redes [mp.si] Protección de los soportes de información [mp.si.1] Etiquetado [mp.si.2] Criptografía [mp.si.3] Custodia [mp.si.4] Transporte [mp.si.5] Borrado y destrucción [mp.info] Protección de la información [mp.info.1] Datos de carácter personal [mp.info.2] Calificación de la información [mp.info.4] Firma electrónica [mp.info.5] Sellos de tiempo [mp.info.6] Limpieza de documentos [mp.info.9] Copias de seguridad (backup) [mp.s] Protección de los servicios [mp.si.1] Protección del correo electrónico

Ilustración 8. Medidas de seguridad definidas en el Anexo II del RD 3/2010 aplicables al Ayuntamiento 1 de Gran Canaria.

Una vez conocidas las medidas al respecto del ENS que nos ofrecen un conocimiento del grado de cumplimiento del Anexo II del Real Decreto ENS, Informe de Insuficiencias, se elaboraría un Plan de mejora de la seguridad donde se propondrá un Plan de medidas de seguridad para subsanar las desviaciones de cumplimiento de lo dispuesto en el ENS, y donde se proponen una serie de tareas. La responsabilidad de su ejecución y la provisión de recursos recae sobre el Ayuntamiento 1 de Gran Canaria, ya sean propios o mediante externalización. El Responsable de Seguridad (RSEG) se encargará de la supervisión de su ejecución. Las tareas por realizar se organizan en tres grupos:

1. **Tareas prioritarias.**
2. **Tareas de implementación del ENS.**
3. **Tareas Periódicas, que competen al CSI y los responsables de la AAPP** correspondiente.

El proceso de implementación del ENS, según la guía CCN-STIC-883, se eleva a **4 trimestres**, es decir, a un año de duración. **D.One+**, tendrá en consideración aquellas medidas de índole operativas, organizativas o de seguridad que impliquen a un **servicio externalizado**, de manera

que aseguramos el correcto encaje de nuestra solución en la adecuación de la AAPP correspondiente al ENS.

Tareas Prioritarias	Control/Cumplimiento	Responsable	Trimestre
En caso de servicios externalizados: completar con certificados de Conformidad ENS de los servicios subcontratados por el proveedor. Para la gestión diaria completar con los Informes/herramientas seguimiento SLA proporcionadas por el proveedor.	Contratación y acuerdos de nivel de servicio [op.ext.1] Gestión diaria [op.ext.2]	CSI	T1,T2.
En caso de servicios externalizados: completar con los procedimientos documentados proporcionados por el proveedor de configuración de roles/perfiles de acceso a los servicios.	Identificación [op.acc.1]	RSIS	T1,T2
En caso de servicios externalizados: completar con los procedimientos documentados proporcionados por el proveedor de configuración de roles/perfiles de acceso a los servicios.	Requisitos de acceso [op.acc.2] Proceso de gestión de los derechos de acceso [op.acc.4]	RSIS, CSI	T1,T2
En caso de servicios externalizados: completar con procedimientos documentados proporcionados	mp.info.9 Copias de seguridad (back up)	RSIS	T1

por el proveedor de la política de copias de seguridad y de restauración.			
En caso de servicios externalizados: completar con los procedimientos documentados de coordinación con el Ayuntamiento para la gestión incidentes y de comunicación de los mismos a las autoridades de control.	Gestión de incidentes [op.exp.7] Registro de la gestión de incidentes [op.exp.9] mp.mon.1 Detección de Intrusión	RSEG	T1,T2
En caso de servicios externalizados: completar con procedimientos documentados de configuración de los registros de actividad de los usuarios a los servicios Información/plataforma de visualización proporcionada por el proveedor de los accesos de los administradores del sistema que soporta los servicios (en caso de que se hayan requerido).	Registros de la actividad de los usuarios [op.exp.8] Segregación de funciones y tareas [op.acc.3]	RSIS	T2,T3

En caso de que se encargue a terceros desarrollos de software, solicitar que se utilicen metodologías de desarrollo seguro y que satisfagan los requisitos necesarios para cumplir con el ENS.	Desarrollo [mp.sw.1] Formación [mp.per.4]	CSI	T3,T4
En caso de servicios externalizados: recopilar los procedimientos documentados proporcionados por el proveedor de coordinación con el Ayuntamiento para la realización de pruebas de aceptación y puesta en servicio. Informes resultados pruebas y plan de acción y los Informes proporcionados por el proveedor con resultados de las inspecciones periódicas realizadas y plan de acción	Aceptación y puesta en servicio [mp.sw.2] Protección de los servicios y aplicaciones web [mp.s.2]	RSIS	T3,T4
En caso de servicios externalizados: completar con procedimientos documentados proporcionados por el proveedor de los mecanismos de autenticación de acceso a los servicios	Mecanismo de autenticación [op.acc.5] Acceso remoto (Remote Login) [op.acc.7]	RSIS	T1,T2
En caso de servicios externalizados: completar con procedimientos documentados proporcionados por el proveedor de configuración de los requisitos	Acceso local (local logon) [op.acc.6]	RSIS	T1,T2

del control: limitación de intentos de acceso, aviso de obligaciones, información sobre el último acceso.			

Tabla 2. Tareas Prioritarias ENS.

TAREAS DE IMPLEMENTACION DEL ENS	CONTROL	RESPONSABLE	T1	T2	T3	T4
MARCO OPERACIONAL - PLANIFICACIÓN						
ARQUITECTURA DE SEGURIDAD- Recopilar, organizar, completar y mantener actualizada documentación sobre: áreas y puntos de acceso, del sistema, líneas de defensa, identificación y autenticación, controles técnicos, relaciones con terceros, para que formen parte del SGSENS. **En caso de servicios externalizados**: completar con la documentación proporcionada por el proveedor sobre las comunicaciones con el Ayuntamiento, y con otros sistemas interconectados comunicaciones con el Ayuntamiento, y con otros sistemas interconectados	op.pl.2	RSIS		X	X	
DIMENSIONAMIENTO Y GESTIÓN DE LA CAPACIDAD - (en caso que aplique) - Implantar un procedimiento para la realización de un estudio de estos parámetros antes de la entrada en producción de nuevos elementos. **En caso de servicios externalizados**: completar con la documentación regular proporcionada por el proveedor sobre los recursos disponibles y consumidos	op.pl.4	RSIS		X		

MARCO OPERACIONAL – EXPLOTACIÓN						
MANTENIMIENTO - Documentar todas las acciones de mantenimiento (físico y lógico). Registrar estas acciones y sus resultados. Desarrollar un procedimiento para analizar, prioridad la aplicación de actualizaciones de seguridad, parches, mejoras, etc. **En caso de servicios externalizados**: completar con procedimientos documentados de coordinación con el Ayuntamiento para realizar acciones de mantenimiento sobre el sistema.	op.exp.4	RSIS			X	
GESTIÓN DE CAMBIOS – Desarrollar un procedimiento de Gestión de cambios. **En caso de servicios externalizados:** completar con procedimientos documentados de coordinación con el Ayuntamiento para realizar cambios sobre el sistema que soporta los servicios	op.exp.5	RSIS			X	
PROTECCIÓN DE LAS CLAVES CRIPTOGRÁFICAS- Documentar las medias de seguridad implementadas para garantizar la protección de las claves criptográficas durante todo su ciclo de vida. Para sistemas de categoría media asegurará la utilización de programas evaluados o dispositivos criptográficos evaluados que empleen algoritmos acreditados por el CCN. **En caso de servicios externalizados:** completar con procedimientos documentados de protección de las claves criptográficas del Ayuntamiento que se encuentren alojadas en el sistema que soporta los servicios	op.exp.11	RSIS				X

MARCO OPERACIONAL – SERVICIOS EXTERNOS						
CONTRATACIÓN Y ACUERDOS DE NIVEL DE SERVICIO - (incluido en medidas priorizadas)	op.ext.1					
GESTIÓN DIARIA - (incluido en medidas priorizadas)	op.ext.2					
MARCO DE PROTECCIÓN – PROTECCIÓN DE LAS COMUNICACIONES						
PROTECCIÓN DE LA CONFIDENCIALIDAD- Realizar un procedimiento que describa la forma en la cual se protege la confidencialidad de la información cuanto esta discurre por redes fuera del propio dominio de seguridad. **En caso de servicios externalizados**: completar con documentos proporcionados por el proveedor con información sobre los mecanismos de cifrado implementados en las comunicaciones.	mp.com.2	RSIS				X
PROTECCIÓN DE LA AUTENTICIDAD Y DE LA INTEGRIDAD - Realizar un procedimiento/norma que establezca la necesidad de utilizar redes privadas virtuales para garantizar la autenticidad y la integridad de la información antes de su intercambio. **En caso de servicios externalizados**: completar con documentos proporcionados por el proveedor con información sobre los mecanismos implementados para proteger la autenticidad y de la integridad	mp.com.3	RSIS				X

SEGREGACIÓN DE REDES – (incluido en medidas priorizadas)	mp.com.4					
MARCO OPERACIONAL – PROTECCIÓN DE LA INFORMACIÓN						
DATOS DE CARÁCTER PERSONAL – Desarrollar las acciones de seguridad necesarias para llevar a cabo la implantación de la normativa de protección de datos (RAT, designación DPD, Análisis de Riegos RGPD, Evaluación de Impacto, contratos de encargado del tratamiento, alinear medidas de seguridad con las del ENS). **En caso de servicios externalizados:** recopilar documentos /plataformas online, proporcionados por el proveedor, con evidencias de cumplimiento de la normativa de protección de datos.	mp.info.1	CSI	X	X	X	X
FIRMA ELECTRÓNICA – - (en caso que aplique) Desarrollar, aprobar y dar publicidad a la Política de Firma Electrónica. Realizar un procedimiento que recoja los requisitos que deben cumplir los mecanismos de firma electrónica. **En caso de servicios externalizados:** completar con documentos proporcionados por el proveedor con información sobre las medidas de protección de la firma implementadas.	mp.info.4	CSI			X	X

SELLOS DE TIEMPO –- (en caso de que aplique) Realizar un procedimiento que recoja los requisitos que deben cumplir los mecanismos de sello electrónico **En caso de servicios externalizados**: recopilar documentos proporcionados por el proveedor con información sobre las medidas de seguridad implementadas para proteger el sello de tiempo	mp.info.6					

Tabla 3. Tareas de Implementación del ENS.

9 SOLUCIÓN TÉCNICA

La solución técnica **D.One+** está desarrollada en su totalidad en infraestructura Cloud de tipo Privado, permitiendo su escalabilidad en el tiempo mediante la contratación de más y mejores servicios con el proveedor Microsoft a través de componentes Azure, lo que ya se considera un estándar funcional para este tipo de soluciones. Sin embargo, lo que está implementado dentro de cada uno de nuestros nodos es lo que verdaderamente diferencia a **D.One+** como una solución novedosa respecto a lo existente en el mercado a nivel de MDM. Además, todo el desarrollo está realizado sobre tecnología en licencia Open Source que, como ya se ha comentado en el capítulo relacionado con el Modelo de Negocio, permite un ahorro sustancial en licenciamiento y la no dependencia de otros proveedores de servicios.

A continuación, explicaremos pormenorizadamente las características de la arquitectura en 3 capas de la solución **D.One+**:

1. Capa de Datos y Fuentes Externas *(Extracción).*
2. Capa de Limpieza y Record Linkage *(Transformación).*
3. Capa de Blockchain y Aplicación *(Carga).*

Independientemente de lo señalado en cada una de estas capas internamente se produce un proceso ETL, que definiremos a continuación:

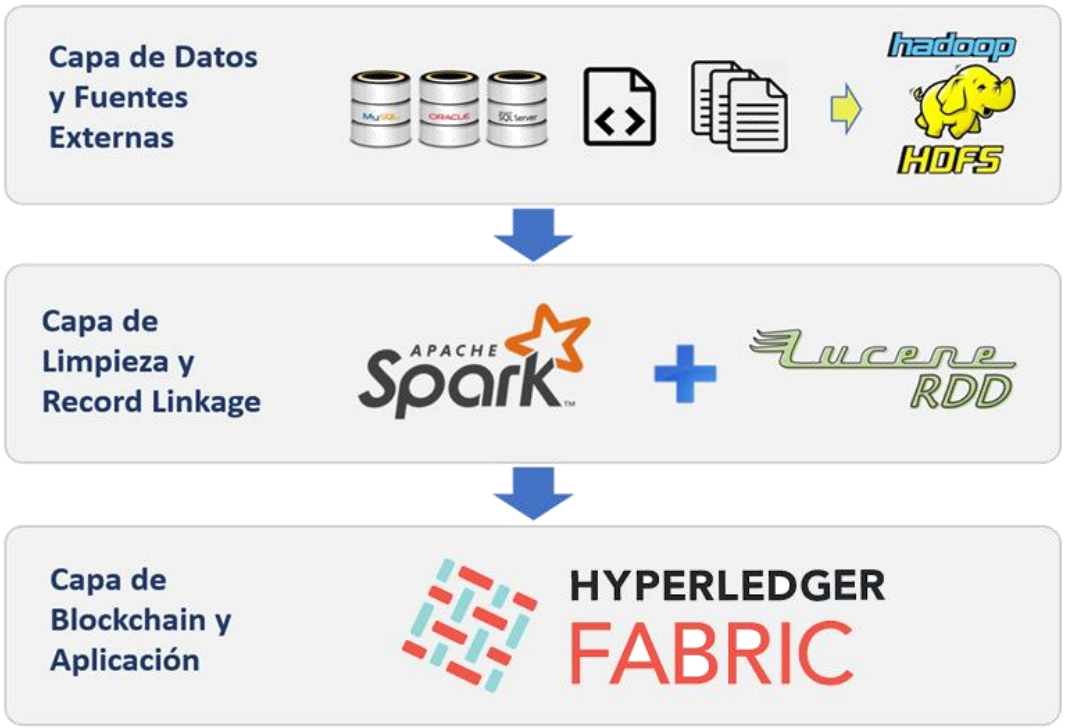


Ilustración 9. Esquema de 3 capas de la arquitectura d.one. fuente: elaboración propia.

9.1 ARQUITECTURA DEL PROYECTO

Fuentes de datos externas

Para el proyecto **D.One+** se han identificado varias fuentes de datos externas que proveerán los datos de los ciudadanos necesarios para sus casos de uso; sistemas gestores internos de las administraciones públicas participantes en la infraestructura de **D.One+**, la Plataforma de Intermediación de Datos (PID); y de cara al futuro, la implementación dentro del mismo del caso de uso ESSIF de la red EBSI (red Blockchain desarrollada por la Comisión Europea para brindar servicios públicos transfronterizos en la Unión Europea, la cual estará desplegada y disponible a partir de 2020/2021).

Más adelante, veremos cómo según el proceso a ejecutar y el campo de datos a obtener, se priorizará una fuente externa de los datos u otra.

Plataforma de Intermediación de Datos

En la política de hacer más sencilla la relación del ciudadano con la Administración General del Estado la Ley 39/2015, de 2 de octubre, de Procedimiento Administrativo Común establece, en su artículo 28.2, el derecho del ciudadano a no aportar los datos y documentos que obren en poder de las Administraciones Públicas, las cuales utilizarán medios electrónicos para recabar dicha información siempre que, en el caso de datos de carácter personal, se cuente con el consentimiento de los interesados.

Sin embargo, en los procedimientos administrativos, ha sido habitual la petición de documentos acreditativos de identidad y del lugar de residencia a efectos de verificar estos datos personales. Con los Reales Decretos 522/2006 y RD 523/2006, de 28 de abril, se suprime la necesidad de aportar estos documentos en todos los procedimientos de la AGE y de sus organismos públicos vinculados o dependientes. No obstante, la verificación de estos datos sigue siendo esencial para la tramitación de los procedimientos.

Los servicios de verificación y consulta de datos de la Plataforma de Intermediación de Datos (PID, a partir de ahora) permiten que cualquier organismo de la Administración, pueda consultar o verificar dichos datos, sin necesidad de solicitar la aportación de los correspondientes documentos acreditativos, permitiendo así hacer efectiva esta supresión.

El objetivo de los servicios de verificación de datos es hacer posible la validación, por medios electrónicos de información que obra en poder de la AAPP. Con estos servicios se pretende:

- Dar cumplimiento a los derechos reconocidos en el artículo 28.2 de la Ley 39/2015 del Procedimiento Administrativo Común de las Administraciones.
- Hacer más cómodo para el ciudadano el inicio de los trámites, evitando que tenga que adjuntar a la solicitud documentos que acrediten su identidad y su empadronamiento.
- Simplificar la tramitación de los procedimientos administrativos.
- Reducir el volumen de papel gestionado en la Administración.

Las consultas a los servicios de verificación de datos necesarios para este proyecto se pueden realizar de forma automatizada desde una aplicación de gestión de un trámite, adaptadas para invocar los Web Services proporcionados por el servicio. Actualmente en SVD se pueden verificar los siguientes tipos de datos:

Servicios de Verificación y Consulta de Datos de Identidad (SVDI)

A través de la PID se pone a disposición de los organismos públicos 2 servicios que vamos a utilizar respecto al proceso de identificación de una persona:

- Verificación de Datos de Identidad: Este servicio permite confirmar o verificar que un determinado conjunto de datos corresponde al número de identificación introducido por el usuario.
- Consulta de Datos de Identidad: Este servicio permite obtener los datos de un determinado documento de identidad a partir del número de identificación del mismo (DNI/NIE o Número de Soporte). Gracias a este servicio, obtendremos los datos exactos del **DNI, Nombre y Apellidos (según Censo),** campos necesarios para componer con absoluta exactitud estos datos en el registro maestro de una persona.

La validación se realiza contra las Bases de Datos del Organismo que los custodia: la Dirección General de Policía (DGP).

En ambos casos, la petición que recibe la PID debe venir firmada por el organismo que realiza la consulta a través del sello de órgano de la entidad consultora.

Servicio de Verificación de Datos de Residencia (SVDR)

Estos servicios son los encargados de consultar al INE los datos de empadronamiento de un ciudadano para aquellos organismos que requieran de un certificado de empadronamiento de un ciudadano. Devuelve la dirección que consta a fecha actual. Para la verificación de los datos de residencia existen 3 tipos de servicios, siendo de interés para nuestro proyecto el *Servicio de Consulta de Datos de Residencia con Fecha de la última Variación. Dicho servicio i*ncluye sobre el servicio de Datos de residencia Extendido, la fecha de la última variación padronal válida (Alta o modificación). No incluye información histórica, ni relación de convivencia. Es decir, no es posible a través de este servicio saber los datos de anteriores domicilios, ni las personas que conviven en el mismo domicilio que él. La referencia temporal incluida, si permite conocer la fecha en la que el ciudadano realizó el último padrón, y en la mayoría de los casos, si se cumplen los requisitos de tiempo de residencia en un domicilio o zona geográfica (Región, provincia, municipio). Gracias a este servicio, obtendremos los **mejores datos del ciudadano referentes a su domicilio**, campos también necesarios para los registros maestros.

Cuando se inicia cualquier proceso hacia la PID, podemos observar dos tipos de roles y actores en esta llamada de servicios [https://administracionelectronica.gob.es/ctt/scsp#.X57YNoj7RPY]:

- La plataforma de Intermediación de Datos con un rol de cedente y emisor.
- La Administración Pública, a través de la plataforma Big Data que veremos más adelante con un rol de cesionario.

Para que puedan interactuar los diferentes actores mencionados, se hará uso de la plataforma de intermediación MHAP para el intercambio de información. Para ello se hará uso de la librería SCSPv3 la cual permite integrar en las aplicaciones de *Backoffice* de los distintos organismos públicos procesos de consulta de los datos ofrecidos por emisores que cumplan con las especificaciones SCSP como todos los disponibles en la plataforma de intermediación.

Los requisitos para que los organismos puedan solicitar cualquier servicio de consulta nombrado a la PID son los siguientes:

- Acceso a la Red SARA. En el caso de Comunidades Autónomas o Ayuntamientos es necesario la firma de un convenio marco que se proporciona.
- Solicitar el alta de la aplicación /usuario mediante el formulario de 'Alta en el servicio'. Descarga de los formularios en el apartado Documentos.
- Los formularios tienen que remitirlos firmados electrónicamente a Soporte para su tramitación, mediante el formulario Web de apertura de solicitudes de soporte técnico, de la Plataforma de Intermediación. Es recomendable remitirlos en formato electrónico editable o que permita copiar la información del mismo, sobre todo en caso de que el documento firmado remitido sea en papel (excepcionalmente) o escaneado.

Caso de uso ESSIF de la infraestructura EBSI

La Infraestructura Europea de Servicios *Blockchain* (EBSI), es un conjunto de recursos tecnológicos y servicios basados en tecnología *Blockchain* que ofrecen a las personas y empresas de la Unión Europea una potente red de nodos distribuidos en sus diferentes estados miembros, que les permite la notarización de documentos, la gestión de certificados, una **identidad digital soberana en el ámbito europeo** y finalmente, una transferencia fiable de información.

El caso de uso de notarización de documentos permite verificar la integridad de los datos en los procesos automatizados, agilizando así la burocracia a la que nos enfrentamos día a día.

La gestión de certificados hace posible que el usuario sea propietario de sus credenciales educativas, posibilitando al mismo disponer de sus títulos, certificados, etc. en tiempo real y que éstos sean siempre accesibles y verificables por diferentes actores de la red.

Por otro lado, la identidad digital soberana presenta infinidad de mejoras en los procesos implantados en esta red de nodos *Blockchain* puesto que permite al ciudadano poder gestionar su propia identidad, siendo el propietario de los datos, y haciendo uso de éstos para identificarse en los diferentes procesos que se lo requiera y sólo con los datos que sean necesarios. Con esto se pretende eliminar la duplicidad de los datos del usuario, poder acceder a la versión de los

datos más actualizada, y que terceros no empleen la información de los usuarios sin su consentimiento y con fines ajenos a los propuestos, entre otros.

El último caso de uso aprovecha esta tecnología para compartir datos de forma segura entre diferentes actores de la UE y así contribuir a la transparencia de los datos y la simplificación de la burocracia nuevamente.

Así mismo, teniendo en cuenta el papel tan importante que tienen las administraciones públicas frente a la Unión Europea, y cómo adoptando soluciones tecnológicas basadas en esta infraestructura hacemos posible una rápida adopción de la tecnología *Blockchain* y la autoidentidad soberana (SSI) que empodera al ciudadano, la UE conseguiría así liderar la constitución de esta infraestructura, alcanzando los siguientes objetivos:

- **Unificación de soluciones**: la colaboración entre los diferentes actores del caso de uso ESSIF y el consenso sobre el diseño de la ingeniería SSI permite una rápida adopción de este tipo de infraestructuras de datos distribuidos, facilitando así transacciones eficientes y confiables entre actores en una única solución y válida para todos los procesos existentes en la red de uso.
- **Interoperabilidad de datos en la UE**: creando acuerdos y políticas de gobernanza comunes, los ciudadanos y entidades podrán acceder y verificar la información disponible, haciendo posible el plan de interoperabilidad de la UE. Este plan persigue garantizar la gobernanza, la coordinación y la puesta en común de iniciativas de interoperabilidad, desarrollar soluciones de interoperabilidad organizativa, la inclusión de las partes interesadas y concienciación de los ciudadanos sobre la interoperabilidad, desarrollar, mantener y fomentar los capacitadores clave de interoperabilidad y por último, desarrollar, mantener y promover instrumentos que apoyen la interoperabilidad.
- **Mayores beneficios**: arquitectura más sencilla para la identidad digital, que potencia la seguridad (jurídica y tecnológica) de los datos y la privacidad de los datos centrada en el usuario, es decir, el usuario es el propietario de sus datos. La red de nodos está regulada desde las instituciones públicas y desde las de derecho privado ya que los procesos que ruedan sobre dicha red son legalmente vinculantes.
- **Mejora de la reputación de la UE y de las AAPP**: la seguridad de los datos que se encuentran alojados en los nodos de esta red Blockchain podría ser un problema existente, corriendo el riesgo del robo de éstos por parte de agentes externos. Diversos nodos de la red de estudio pertenecen a universidades, administraciones públicas, etc que llevan a cabo tareas de ciberseguridad, por lo que complicaría un ciberataque de la red. Así mismo, uno de los principios básicos de los estados miembros como comunidad, es ver cómo se defienden ante un ataque que viene de fuera. Esta confianza para los datos genera un enorme beneficio para los ciudadanos, mejorando así la imagen y funcionalidad de la UE y de las AAPP.

Este proyecto nace por la necesidad de la implantación de la GDPR, la mejora en la educación, la aplicación de *Blockchain* en las esferas públicas y privadas, así como la implantación de un modelo de identificación que cumpla con el marco legal y regulatorio europeo y finalmente, contribuir en la investigación e innovación de la tecnología *Blockchain*.

La gran ventaja de este proyecto es la mejora de los procesos que vienen implícita con la tecnología *Blockchain*, la seguridad y trazabilidad de los datos que ofrece, y la colaboración entre los estados miembros de la UE y las AAPP presentes en ellas. Permitirá el desarrollo de nuevos modelos de negocio y nuevas aplicaciones en entornos públicos y privados que conecten a la red de la infraestructura EBSI y hacer uso de los servicios anteriormente descritos.

Es por ello por lo que se propondrá la integración a largo plazo de esta fuente de información en nuestro proyecto ya que, para el despliegue de esta infraestructura, primeramente, en 2019, se seleccionaron los casos de uso que harían posible este gran sistema de la información, para así ir desplegando los demás casos de uso a lo largo del 2020 y 2021. El caso ESSIF de la infraestructura EBSI será un pilar fundamental para facilitar los datos de identificación más fiables del usuario de una manera segura, fidedigna y verificada.

Ilustración 10. Estado actual de la infraestructura EBSI.

Sistemas Gestores Internos de las Administraciones Públicas.

Los datos de identificación del ciudadano que se obtendrán de los sistemas gestores internos mencionados podrán tener múltiples formatos como bases de datos específicas, datasets de tipo semiestructurado, etc.

En los siguientes capítulos se explica cómo se abordarán los procesos de extracción, transformación y carga que recaen sobre éstos para obtener los mejores datos del individuo y cómo finalmente estos datos forman parte del registro maestro almacenado en los nodos de la red Blockchain del proyecto.

9.2 DESCRIPCIÓN DEL PRODUCTO/SERVICIO

Para soportar una red Blockchain/Hyperledger Fabric + Arquitectura Big Data cumpliendo las características que hemos descrito en nuestro trabajo, es necesario plantear una arquitectura prototipo pensada para el caso de uso de la organización con la que hemos firmado el convenio de desarrollo y colaboración (Ayuntamiento 1 de Gran Canaria), asegurando la disponibilidad en servidores contratados, tanto para el servicio de *Paralelización*, como para la red de *Blockchain*.

Capa de Datos

Es la encargada de la extracción, estandarización y centralización de todos los registros provenientes de las distintas fuentes de datos que el cliente desea masterizar independientemente de su formato (BBDD, archivos de texto planos, etiquetados, etc.) Por lo tanto, para el correcto desarrollo de esta tarea, el proceso de análisis y consultoría es esencial, ya que desde este momento debemos cumplir con todas las políticas y estándares de seguridad exigidas tanto a la organización, como a D.One,

Dependiendo de la naturaleza de la fuente de los datos, la extracción se realizará mediante sistemas/conectores diferentes:

Para el caso de datasets contenidos en archivos de texto plano etiquetados (tipo JSON) y ubicados en los sistemas gestores internos de las administraciones públicas, su extracción se realizará vía protocolo SFTP para depositarlos en los Data Node, realizar el proceso de estandarización necesario y dejarlos disponibles para su procesamiento en la Capa de Limpieza y Record Linkage.

Para las conexiones remotas a Bases de Datos, primero se realizará un análisis de las políticas de seguridad de la empresa, del cual dependerá la estrategia que adoptará el equipo D.One para extraer los datos. Las alternativas iniciales propuestas son las siguientes:

1. Uso de Apache Sqoop/Flume o Kafka para conexión directa con RDBMS.
2. Conexión mediante VPN para la extracción de datos.
3. Procesos de tipo batch que generen archivos de texto plano y su remisión vía SFTP a **D.One**.
4. Procesos de extracción mediante MapReduce.

Cada uno de los proyectos de implantación de la solución D.One contará con procesos ETL que envíen la información al NameNode, que estará diseñado en base al ecosistema Hadoop (Spark) ya sea con Apache Flume o Apache Sqoop. Para casos más especiales (como es la extracción en tiempo real/streaming) utilizaremos Apache Kafka.

A continuación, mostramos el diagrama funcional de los ETL en comunicación con el ecosistema HADOOP/SPARK, donde se puede observar conceptualmente su alcance.

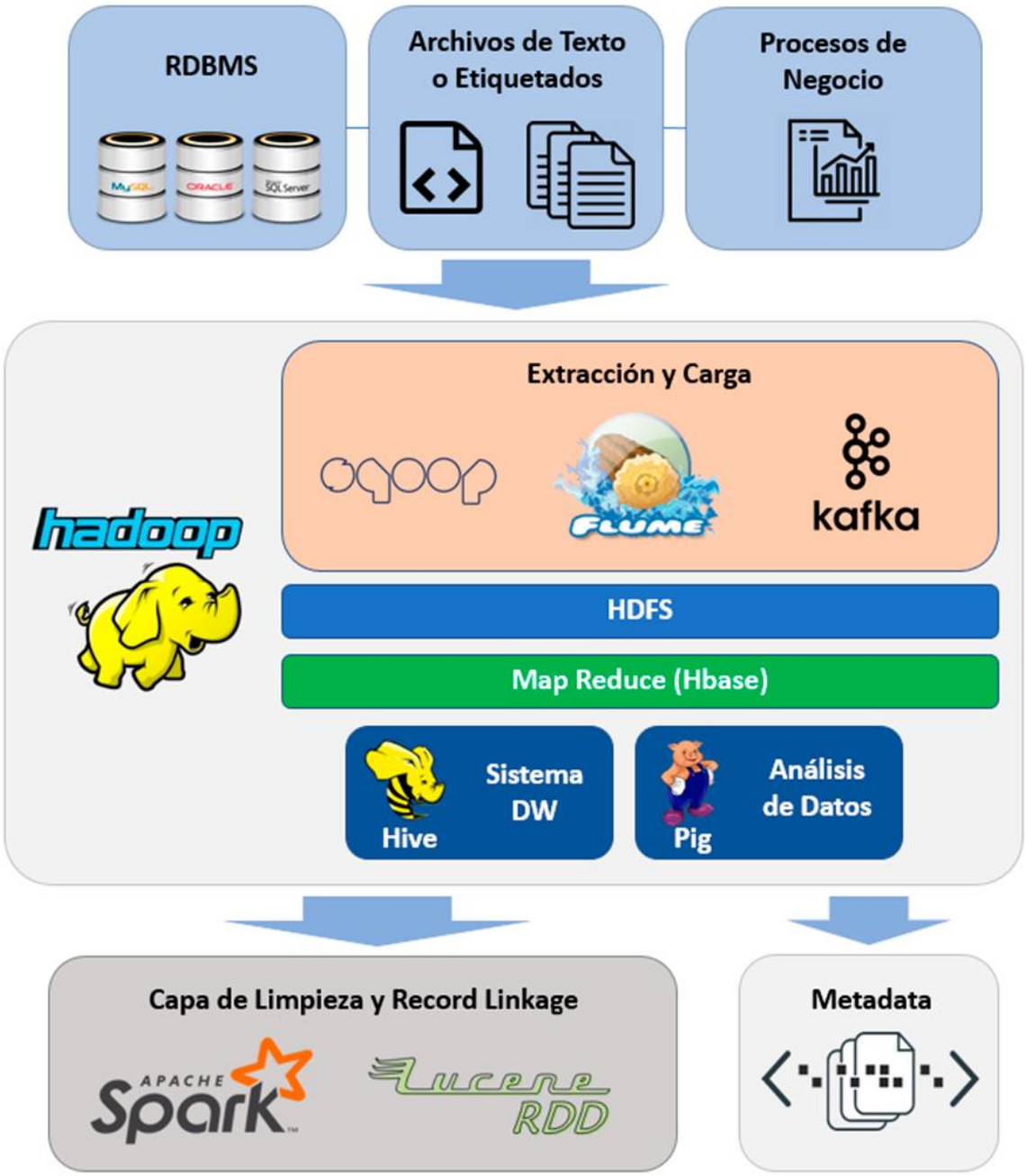


Ilustración 11. Estructura funcional de la Capa de Datos y Fuentes Externas. Fuente: Elaboración propia.

Finalmente, tenemos que señalar que este punto del proyecto es crítico, ya que tratamos directamente con los datos más sensibles de una organización y la solución que se adoptará siempre se establecerá en consenso con el cliente, asegurando (con la máxima calidad) que los registros que se almacenan y transforman en nuestro sistema, están seguros desde su extracción hasta su almacenamiento en la Capa Blockchain.

9.3 CAPA DE LIMPIEZA Y RECORD LINKAGE

Esta capa es la encargada de realizar el proceso CORE de la solución D.One y para ella es de vital importancia el procesamiento y la centralización de los datos que se ha producido en la anterior etapa, ya que apoyándose en éstos es ahora cuando se realiza la Transformación dentro

del proceso ETL. Ya que una vez que tenemos los registros en la estructura SPARK comienzan esos procesos de transformación.

La primera estandarización del dato implica su limpieza 'higiénica', que se realiza mediante la normalización de los registros que tienen espacios, puntos o caracteres especiales y que no afectan a su semántica, con el objeto de excluirlos y/o reemplazarlos, con el fin de mejorar su legibilidad de cara al proceso de Record Linkage.

El siguiente proceso que se realiza es la deduplicación del dato. El cual evalúa cuáles son los registros que están duplicados y los normaliza a un único registro. Esta tarea es fundamental para crear el dato único y se constituye como la primera fase para encontrar ese Dato Maestro que genera la aplicación D.One. En el siguiente diagrama explicativo, se visualiza el funcionamiento de la deduplicación.

Ilustración 12. Proceso de Deduplicación. Fuente: Blogbuzzllc.com

Después de la deduplicación, se ejecuta el proceso de Record Linkage que es el encargado de realizar la estandarización estadística (utilizando Machine Learning) que posteriormente permitirá la indexación de los datos provenientes de distintas fuentes y el proceso que a continuación detallamos: la evaluación generada por el proceso de Record Linkage se realiza mediante clasificación de vectores utilizando aprendizaje automático (SVM). Cabe señalar que este proceso está basado en un algoritmo de aprendizaje supervisado, que genera como resultado tuplas para registros que: son iguales, o son parecidos, o no tienen ninguna relación. Para una mejor comprensión del proceso que nos ocupa, a continuación, mostramos su diagrama funcional.

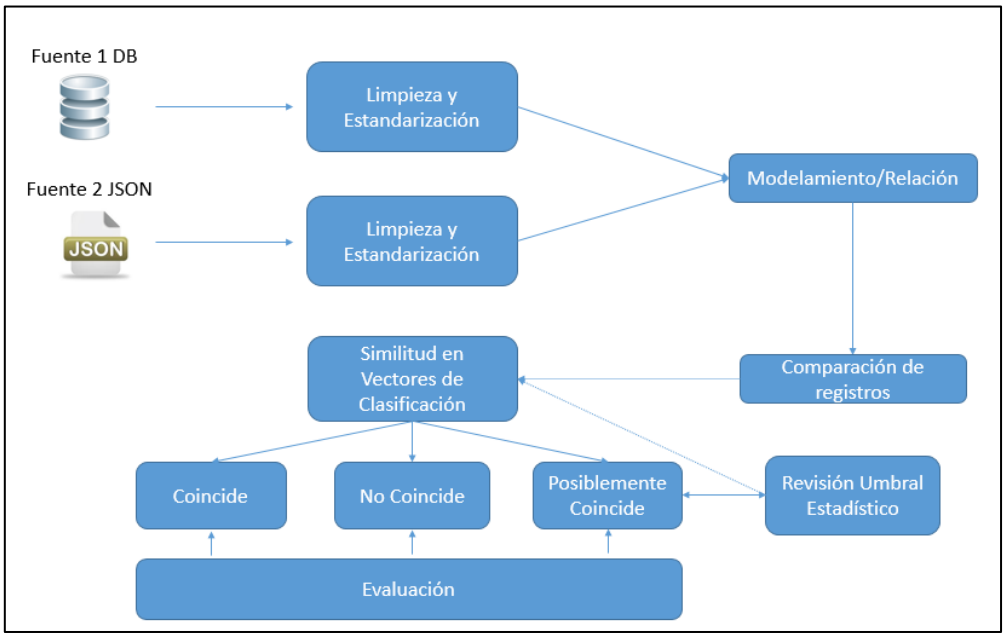


Ilustración 13.: Diagrama de funcionamiento del Record Linkage. Fuente: Elaboración propia.

Como vemos en la imagen, el Record Linkage es capaz de determinar los registros que coinciden, los que no coinciden y los que probablemente coinciden (tal y como señalábamos antes). Todo está paquetizado y sirve, entre otros fines, para llegar a un dato más confiable y generar una BBDD estandarizada, limpia, deduplicada y propensa a almacenar el mejor registro recopilado en comparación a una tabla o archivo sin procesar.

Cómo logramos obtener el mejor registro en D.One+

Para desarrollar la arquitectura **D.One+**, se ha tenido en cuenta la necesidad de que el entorno de computación tenga una capacidad de procesamiento escalable ya que la solución tiene que ser válida tanto para pequeñas organizaciones, como para grandes corporaciones.

Por lo tanto, la Capa de Limpieza y Record Linkage se ejecuta en un entorno HADOOP + SPARK (ecosistema Big Data) que optimiza el uso de recursos, paralelizando tareas que pueden ser costosas en una simple consulta tanto a nivel de código, como a nivel de recursos de máquina. Este entorno está desarrollado sobre componentes Azure de Microsoft que permite variar el número de máquinas en el clúster cuando existen picos por necesidades de computación.

Por lo tanto, nos vemos en la necesidad de encontrar y aplicar una librería o conjunto de librerías que permita efectuar el proceso de Record Linkage en un entorno paralelizado y que, además, permita el uso de MapReduce sobre un sistema de archivos distribuidos de HADOOP HDFS. Y aquí es donde hemos utilizado LuceneRDD, librería basada en Lucene y optimizada para su funcionamiento paralelizado sobre SPARK (y que además realiza el proceso de Record Linkage).

El origen de la misma es Lucene, una API Open Source desarrollada originalmente por Doug Cutting en lenguaje Java en 1999, e incorporada a la “Apache Software Foundation” en 2001. Su principal objetivo era la búsqueda de información y la indexación de texto. Incluso ha sido usado

en la implementación de motores de búsqueda web, los cuales necesitan comparar datos con la cadena de caracteres escrita por el usuario antes de realizar una búsqueda exhaustiva de un término concreto en Internet.

De hecho, puede decirse que Lucene está en el núcleo y origen del propio Apache Hadoop, ya que fue la necesidad de utilizar Nutch, un motor de búsquedas web basado en Lucene de forma paralelizada, lo que llevó al propio Doug Cutting a desarrollar HADOOP.

Con la evolución del procesamiento masivo de registros y, por tanto, el uso de Big Data, se han desarrollado módulos de procesamiento que aplican algoritmos complejos como es el Record Linkage, que deriva de un proyecto GitHub llamado LuceneRDD. Este módulo combina las capacidades de SPARK RDD con las de Lucene.

El desarrollo de LuceneRDD no está centrado únicamente en realizar Record Linkage, sino que tiene otras operaciones que permiten otros tipos de procesamiento de datos. Ejemplos de operaciones de LuceneRDD:

1. *Term Query*: Búsqueda de un término exacto.
2. *Fuzzy Query*: Búsqueda de un término difuso.
3. *Phrase Query*: Búsqueda de expresiones.
4. *Prefix Query*: Búsqueda de un prefijo.
5. *Query Parser*: Consulta para búsquedas analíticas.
6. *Faceted Search*: Búsqueda por facetas.
7. *Record Linkage*: Proceso objeto de nuestro análisis.
8. *Circle Search*: Búsqueda dentro de un radio determinado.
9. *Bbox Search*: Búsqueda enlazada de coordenadas.
10. *Spatial Linkage*: Búsqueda espacial de coordenadas.

Al estar LuceneRDD construido con la librería Lucene sobre estructuras de datos RDD (Resilient Distributed Datasets), tenemos la capacidad de utilizar su potencial con procesamiento en paralelo sobre SPARK.

SPARK trabaja con RDD's y el primer RDD que se crea/genera proviene de la carga de los registros externos procesada en la Capa de Datos y Fuentes Externas almacenados en HDFS. Este RDD se divide en particiones que son enviadas a los SPARK Executors para proceder a realizar todos los procesos propios de esta capa.

Cabe señalar que los RDD son inmutables, por lo cual, si se aplica una transformación sobre uno de ellos se generará un nuevo RDD. Pero, sin embargo, no actúan como una copia de los datos que contiene el anterior 'datasets', sino como una referencia sobre éste.

Lo señalado con anterioridad, sumado a la aplicación de la función de SPARK denominada Lazy Evaluation hace que un nuevo RDD sólo se genere en el momento que en SPARK se ejecuta una función concreta (como puede ser un proceso Reduce).

Teniendo en cuenta este pequeño marco teórico, podemos deducir que LuceneRDD proporciona 3 aspectos fundamentales para el core de la solución D.One. La primera es la indexación y búsqueda de datos, la segunda es el Record Linkage y la tercera es la capacidad de procesar datos en estructuras RDD que permiten que el modelo sea evaluado mediante procesamiento en paralelo y genere tantos RDD como sea necesario para poder determinar cuál es el mejor registro existente entre las distintas fuentes de datos.

Posteriormente, se ejecuta un Reduce en el mismo SPARK lo que terminará por generar el dato maestro que estábamos buscando. De forma gráfica el proceso de Paralelización sería:

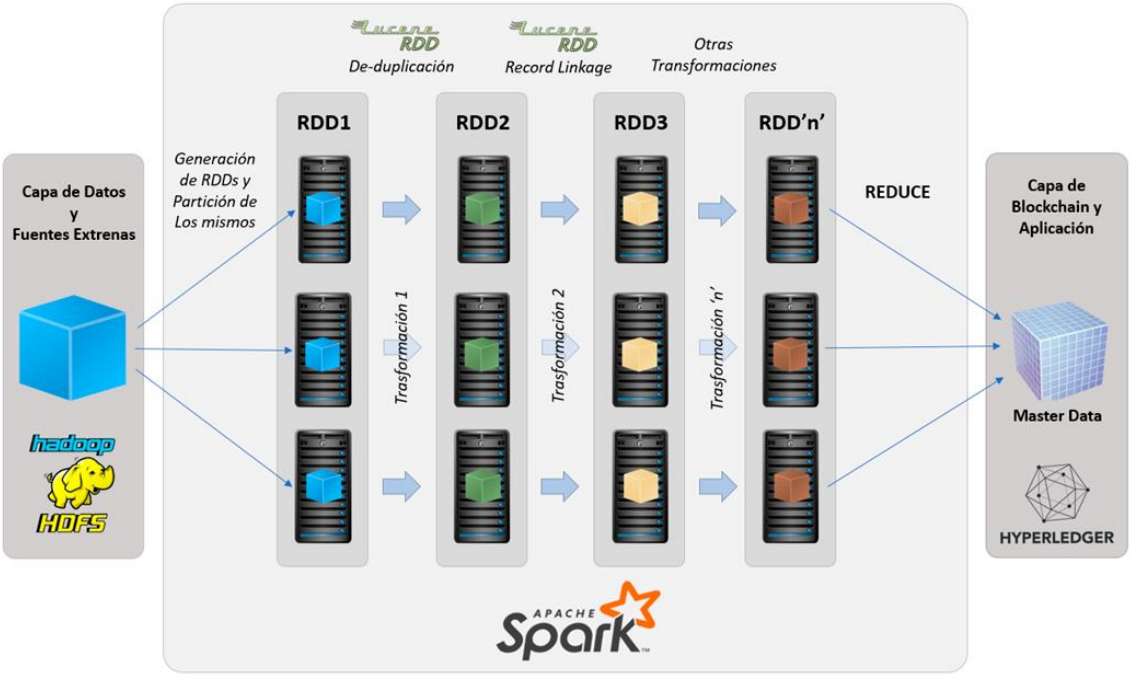


Ilustración 14. Proceso de transformación y paralelización mediante RDD's en Spark. Fuente: Elaboración propia.

9.4 CAPA DE BLOCKCHAIN Y APLICACIÓN

La última capa definida de nuestro proceso será la de Blockchain y Aplicación. En ella se securizarán los datos maestros antes calculados y se generará un libro mayor que contendrá cada uno de los bloques, logrando la trazabilidad perfecta para todos ellos.

Tecnologías aplicadas en el proyecto

1. **Justificación de la elección de Hyperledger Fabric como Blockchain referencia para el proyecto D.One+**

   Tras la descripción de la estructura de la que se compone el proyecto en estudio, donde ya se ha planteado como está organizado un MDM y su relación con BIG DATA, a partir de este punto, nos centraremos, de manera detallada, en la parte del proyecto que está vinculada

con la tercera pata de **D.One+**, es decir, **Blockchain**, y en particular, en **Hyperledger Fabric (a partir de aquí también lo denominaremos HLF)**, poniendo el foco en analizar la elección de esta tecnología de **código abierto** de carácter **colaborativo**, y que está orientada por defecto a las **redes privadas**.

En la actualidad, nos encontramos con diferentes tipologías de soluciones tecnológicas asociadas con DLT o en particular con Blockchain, que podrían ser base y referencia para la generación de la estructura de cadena de bloques que ya hemos comentado en este documento que, entendemos como necesaria para completar con éxito las necesidades tecnológicas del proyecto.

2. Por qué hemos elegido Hyperledger Fabric y no otra plataforma Blockchain

Partiendo de los primeros comentarios que se han esbozado en este apartado del proyecto, donde se han significado **tres elementos clave** que nos pueden servir de camino para hacer una primera diferenciación con respecto a otras plataformas Blockchain, **a los que hay que sumar un cuarto elemento clave, que está asociado con las iniciativas a nivel de Blockchain que se están queriendo implantar desde la Unión Europea**.

A continuación, se hace énfasis en estas cuatro primeras claves, las cuales se explican en detalle en los siguientes tres puntos:

- **Recomendación de la EBSI**

  La **EBSI o European Blockchain Services Infrastructure**, *es "la Infraestructura Europea de Servicios Blockchain (EBSI) es una red de nodos distribuidos en toda Europa que brindará servicios públicos transfronterizos. En última instancia, la tecnología Blockchain mejorará la forma en que los ciudadanos, los gobiernos y las empresas interactúan*", según se explica en el portal de la UE "ec.europa.eu", habilitado para exponer este ambicioso proyecto de la Unión Europea.

  Dentro de la arquitectura prevista para esta infraestructura, la **EBSI establece que las dos Blockchain recomendable son *Hyperledger Fabric* e *Hyperledger Besu***, y cumpliendo con esa recomendación de este Servicio Europeo, hemos decidido elegir *Hyperledger Fabric* para la capa de *Blockchain* de **D.One+**.

- **Código abierto**

  Según se ha señalado en diferentes partes de este documento, el proyecto está cimentado en soluciones de código abierto u *Open Source*, y en el caso de esta parte del proyecto se mantiene el mismo criterio, cumpliendo esa exigencia la elección de HLF, que es una cadena de bloques de código abierto.

  Este tipo de propuesta asociada al software libre es uno de los axiomas que permite la expansión del ecosistema Blockchain, logrando que el desarrollo de estas plataformas de código abierto no sólo se haya popularizado en el ámbito de los desarrolladores, sino

que grandes corporaciones, como, por ejemplo, IBM, se han implicado de manera sustancial en los procesos de mejora de este tipo de redes.

Además, la accesibilidad al código fuente de cualquier interesado en la materia, hace que en todo momento exista una **completa transparencia**, y que asociada a esta Blockchain haya una comunidad muy activa, que está marcada por la colaboración entre todos sus integrantes, como veremos en el punto 2.

- **Impulso colaborativo**

Una de las premisas determinantes de Blockchain, es la búsqueda y necesidad de la descentralización, con la generación de nodos, que lleven el registro de todas las transacciones efectuadas en la red, por ello, es evidente que para que eso sea posible **debe existir colaboración** o un impulso colaborativo entre todos los actores o agentes participantes.

Según se ha expuesto con anterioridad, la facilidad de acceso al código hace que ese impulso colaborativo sea fundamental para la creación de nuevas herramientas vinculadas con las redes de Blockchain, y al mismo tiempo**, permiten soluciones que mejoran las plataformas a todos los niveles**, desde el punto de vista de la seguridad hasta su implementación en posibles casos de uso.

Finalmente, otro de los argumentos que llama a este impulso, es la **formación de alianzas**, incluso entre empresas competidoras o administraciones de distintos foros, comunidades o países.

- **Redes privadas**

Es el tercer punto clave para decidirnos por la red HLF. Para ello, se parte del estudio de Hyperledger que realizó durante 2015, donde vieron que **existía la necesidad de establecer soluciones a nivel de redes privadas, con una reestructuración de la composición de los nodos, dando una jerarquía diferenciada a los mismos**, de tal manera que, a diferencia de las redes públicas, que en cada nodo constan de toda la información de manera distribuida en cada peer y un rol único para cada uno, en el caso de estas nuevas redes privadas, se diferenciarán los nodos en función de un rol predeterminado.

La necesidad de este planteamiento nace del hecho de que **las empresas empiezan a ver las ventajas y beneficios que puede tener el uso de Blockchain, pero al mismo tiempo sienten que al tratarse de redes públicas**, **existía un considerable riesgo a nivel de privacidad de la información sensible que estas compañías manejan**.

Este riesgo se minimiza enormemente con la aparición de las redes privadas o públicas permisionadas. Además, al no ser necesario que todos los *peers* tengan toda la información, se **mejora notablemente el rendimiento** de estas redes, disminuyendo los

tiempos por transacción entre las organizaciones incluidas dentro de estas redes privadas.

Esta tipología de red privada permite una **completa privacidad** en las transacciones entre los distintos nodos de la red, con la inclusión de un concepto nuevo, denominado **Canal o Ledger**, del que hablaremos en detalle más adelante, y que implica la creación de espacios privados de transferencia de información, que solo pueden ver aquellos que estén autorizados en ese canal.

Este concepto, se complementa con otro, las **Colecciones**, que nos permiten establecer otro nivel de privacidad dentro del canal de nuestra organización, que al igual que el canal, explicaremos en detalle con posterioridad.

Todo ello, confiere a esta red un enorme grado de **confidencialidad** y por ende mayor **seguridad** en los datos transferidos.

Tras esta visión más generalista fijada en estos primeros elementos claves, se puede profundizar un poco más en esta **plataforma de arquitectura modular**, teniendo presente otras características de esta Blockchain que desglosamos a continuación:

a) **Confidencialidad**.

Para la obtención de esta característica de la red, se utilizan los mencionados canales, que como ya se ha indicado, **permiten generar espacios o subredes entre participantes de una misma red**, impidiendo la visibilidad de los datos o las transacciones a otros miembros de la red que estén fuera del canal.

b) **Resiliencia**.

Dado que la red está formada por varios nodos, esto permite que en el caso de que se produzca algún fallo o caída del sistema de alguno de los *peers*, la información no se puede perder dado que siempre quedará indemne en los otros nodos de la red.

Por eso será clave, que en la organización exista más de un nodo, por lo tanto, lo ideal es partir de dos. Y **lo recomendable, sería tener tres nodos** para asignarle a cada uno diferentes funciones.

c) **Flexibilidad**.

Debido a que se pueden utilizar en muchos y varios tipos de negocio, desde actividades vinculadas con la banca hasta servicios de trazabilidad relacionados con el proceso de transporte de un producto “X”, desde su origen hasta su destino final.

Además, esta capacidad también es tremendamente flexible en lo relativo a la parcela de privacidad y confidencialidad, como ya se ha indicado con anterioridad, con los ejemplos de los canales y las colecciones, creándose libros mayores por cada canal, y también para el caso de la asignación de permisos, que hace que si no existen esos permisos podemos transformar la red de Fabric en una red pública.

Sin dejar de olvidar, que en esta red también se puede establecer un mecanismo de consenso, en función de los protocolos establecidos por los integrantes de las organizaciones que forman la red.

Por lo tanto, de todo esto, se entiende que otra de las grandes virtudes que tiene HLF es su capacidad de moldeo a nivel de diseño, para ajustar la red en función del caso de uso que esté en estudio.

d) **Escalabilidad**.

Es la cuarta característica a resaltar, y que tiene una relación directa con la flexibilidad de la red seleccionada.

Según **Alex Preukschat** en su artículo *"Hyperledger: la Blockchain privada que todos tenemos que conocer"* para la publicación **El Economista** de 29 de enero de 2018, la escalabilidad es**: "…la capacidad que tendría el sistema de poder añadir más nodos en función de la demanda del mismo.".**

Teniendo en cuenta esta premisa, el diseño de esta Blockchain, logra que la escalabilidad no sea una barrera, y que cualquier entidad empresarial vea a HLF como la solución perfecta para plantearse introducir a su corporación en el mundo de las DLT.

Además, otra de sus propiedades a nivel de escalabilidad es que se pueden distribuir sus nodos en multitud de localizaciones geográficas, y, por ejemplo, asociar ese posicionamiento de uno o varios nodos a distintas organizaciones.

Por supuesto, todas estas características expuestas, vendrán aparejadas con las virtudes innatas en toda Blockchain, que son la **descentralización** y la **inmutabilidad** de la red.

Como se puede comprobar en las características mencionadas, **HLF responde al ya comentado obligado cumplimiento del Esquema Nacional de la Seguridad por parte de las Administraciones Públicas**, en relación a la capacidad de las redes o de los sistemas de información de resistir, con un determinado nivel de confianza, los accidentes o acciones ilícitas o malintencionadas que comprometan la disponibilidad, autenticidad, integridad y confidencialidad de los datos almacenados o transmitidos y de los servicios que dichas redes y sistemas ofrecen o hacen accesibles. Así mismo, **la tipología de red escogida cumpliría las normas técnicas de interoperabilidad entre las Administraciones Públicas y el ciudadano que se establece en el Esquema Nacional de Interoperabilidad**.

Estructura de Hyperledger Fabric

Ya hemos comprobado que HLF es la elección adecuada para el proyecto **D.One+**, y conociendo este hecho, vamos a ver dónde se puede conectar esta plataforma con el resto de las tecnologías del proyecto, para de esta forma extraer los datos que se estimen necesarios para complementar y mejorar el dato maestro de **D.One+**.

En este punto debemos de hacer un inciso, ya que, para la implantación de la tecnología Blockchain, se opta por unirse con la asociación sin ánimo de lucro, **ALASTRIA**, cuyo objetivo principal, según se indica en su folleto publicitario es *"la promoción de tecnologías descentralizada/Blockchain"*, y que se ajusta a las necesidades de este proyecto.

Por ello, y dado que en la actualidad la **Red H**, que es como denomina Alastria a su red bajo HLF, ya ha emigrado a la **versión 2.2 de Fabric**, **será esta versión la que se utilice para el desarrollo de la capa de Blockchain del proyecto**.

Antes de comenzar con el análisis del diseño que hemos decidido para la estructura de nuestra red bajo HLF, hay que detallar varios **componentes que son fundamentales para la conformación y entendimiento de la Blockchain** que se integra dentro de nuestro proyecto.

En primer lugar, que **HLF está formada por dos partes**, es decir, que la estructura del libro mayor o Ledger de esta plataforma está formada por dos estructuras, la **Blockchain** y el **World State**, que intrínsecamente se relacionan, y que en caso del segundo deberá ser el punto de extracción de los datos solicitados en las consultas, dada su compatibilidad con DBMS y a su vez, contar con herramientas que pueden trabajar con formato JSON.

Estas herramientas son ***CouchDB*** y ***LevelDB*****,** que son **herramientas de código abierto**, que según se establece en el apartado del hyperledger-fabric.readthedocs.io que explica sus particularidades, *LevelDB* es una base de datos de estado de clave - valor interna al nodo, mientras que *CouchDB* es otra base de datos de estado que según se indica en el documento *"es una alternativa externa"*.

Además, en el caso de *CouchDB*, según este mismo documento, es capaz de: *"…almacenar datos en formato JSON, emitir consultas enriquecidas contra sus datos y utilizar índices para respaldar sus consultas"*.

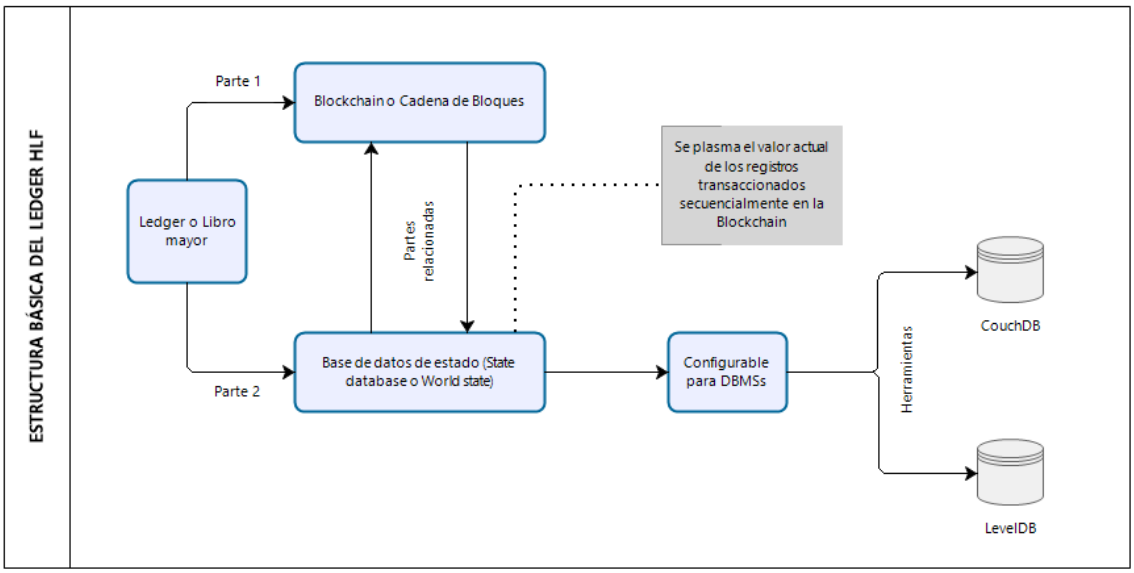


Ilustración 15. Estructura Básica del Libro Contable de HLF. Fuente: Elaboración propia.

Definida la estructura básica de Fabric, toca detallar los componentes que forman parte de la misma, y que nos permitirán realizar el diseño de nuestra solución.

1. **CA (Autoridad Certificadora)**

   Según explica **María Teresa Nieto** en su artículo *"How to generate Hyperledger Fabric certificates using Certificate Authorities"*, para el portal **medium.com** de 29 de diciembre de 2019, una CA es**: "una entidad de confianza responsable de emitir y revocar un tipo específico de certificado mediante firma digital.".**

   Partiendo de esta definición, será necesario establecer cuál será la entidad certificadora de la red que se desee diseñar, y en este sentido existirán dos opciones:

   a) La utilización de la CA de que consta *Hyperledger Fabric*.

   b) Que la propia organización que desea diseñar la red, posea su propia autoridad certificadora.

En la imagen se muestra las operaciones que están vinculadas con una CA, y ya comentadas, desde el momento en que se realiza la solicitud del certificado y su posterior emisión, y el repositorio de los certificados revocados.

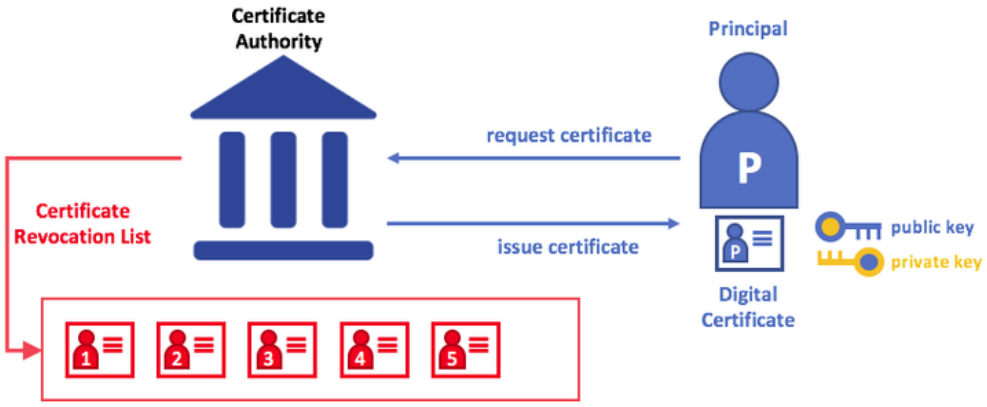


Ilustración 16. Acciones que ejecuta una CA. Fuente: Viblo.asia "Conceptos básicos en Hyperledger Fabric"

2. **MSP (Proveedor de Servicios de Membresía)**

   Se trata de una configuración del sistema de la red, que hace que las transacciones que se producen en la plataforma tengan el nivel de confianza que ha de tener HLF, para lograr los estándares de confidencialidad y de seguridad necesarios para el adecuado funcionamiento de la red.

   Para ello, en el MSP se encuentran las claves públicas de los pares de claves (pública/privada) de los integrantes u organizaciones que interactúan en Fabric, de tal manera que cuando se realiza una transacción de A a B, tras el firmado de la transacción por parte del primero, el MSP realizará una comprobación de la clave pública del sujeto

A, verificando que se trata del par público de las claves de este sujeto, y que ha de coincidir de manera inequívoca para que se pueda ejecutar la transacción.

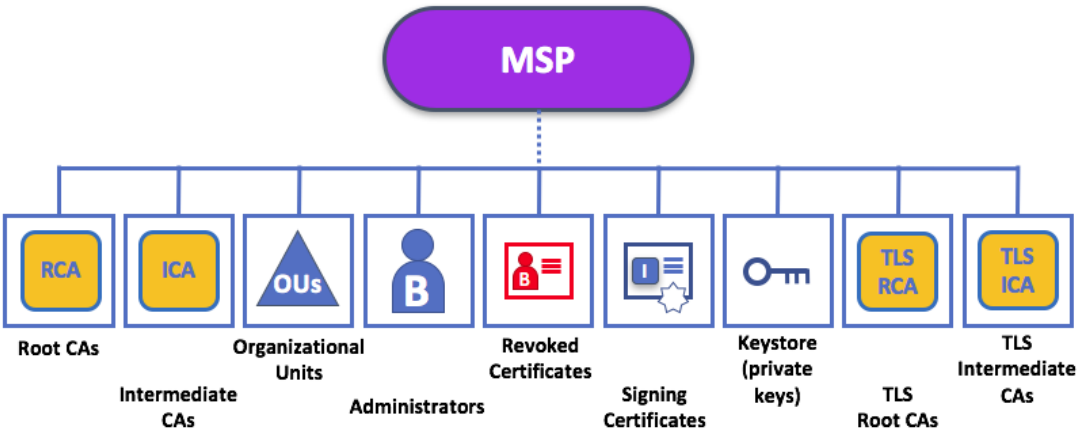


Ilustración 17. Estructura de la capacidad de identificación de MSP. Fuente: Hyperledger-fabric.readthedocs.io.

Por lo tanto, el MSP es el **mecanismo que activa los permisos para que un miembro de la organización o la red pueda realizar transacciones en la plataforma**, es decir, que se le puede denominar como el **AUTORIZADOR**.

Esta facultad de autorizar no sólo se ciñe a la consideración de ser o no ser miembro de la organización, sino que **también establece el rol que tendrá el sujeto en la red**.

Según se muestra en la imagen X, además de autorizar, el MSP tiene la capacidad de identificar certificados RCA y ICA, reconocer las unidades organizacionales, a los administradores, a los certificados revocados o firmados, la clave privada y el aseguramiento de las comunicaciones entre pares mediante los TLS.

3. **Peers o nodos**

Son los componentes que **tienen la función de mantener la Blockchain**, y de conservar la información o los datos de las transacciones que se van incluyendo en los bloques que componen la cadena.

La vigorosidad de la Blockchain estará en función del número de *peers*, por lo tanto, **a mayor número de nodos, mayor capacidad de resiliencia** de la cadena de bloques.

En HLF **no existe un único tipo de peers**, como ya se ha planteado en párrafos anteriores, y **su denominación estará en función del ROL que desempeñen**, siendo estos los siguientes:

a) **Orderer**: Es el nodo que tiene la función de generar el consenso que poseen los algoritmos de consenso de otras redes, además es el receptor, distribuidor y asegurador del orden de las transacciones. Y también es el formador de los bloques.

Por lo tanto, es el que remite a los *leader peers* de cada organización los bloques con las transacciones, como vemos en la siguiente imagen.

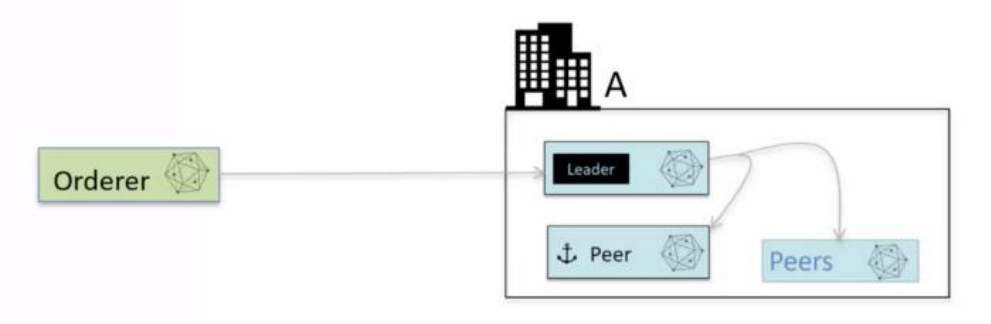


Ilustración 18. Relación entre peers o nodos de una organización y el nodo orderer. Fuente: "Hyperledger Fabric — Conceptos y Tipos de Nodos"

En ella se muestra a la organización o sujeto A, con tres tipos de nodos: Leader, Anchor y otros Peers y el enlace de esta con el nodo Orderer [JD1].

b) **Leader**: Este tipo de nodo es el que recibe los bloques del nodo Orderer y los transmite a cada uno de los nodos de la organización.

Este *peer* tiene la particularidad de que su liderazgo está vinculado con el canal, pero no de manera exclusiva, por lo tanto, **esta tipología de nodos puede ser líder en más de un canal**.

c) **Anchor**: Es el *peer* que recibe la cadena de bloques desde el nodo *Leader* y será el que se comunique de manera bidireccional con los *Anchors* de las otras organizaciones de la red; siendo este, **el único nodo conocido fuera de la organización**.

Esta comunicación se realiza mediante el **protocolo *Gossip***, siendo *Gossip* el modus operandi que tienen los pares de la red entre las organizaciones participantes, para transferirse los datos del canal y los *ledgers* donde se incluyen los bloques y sus correspondientes transacciones.

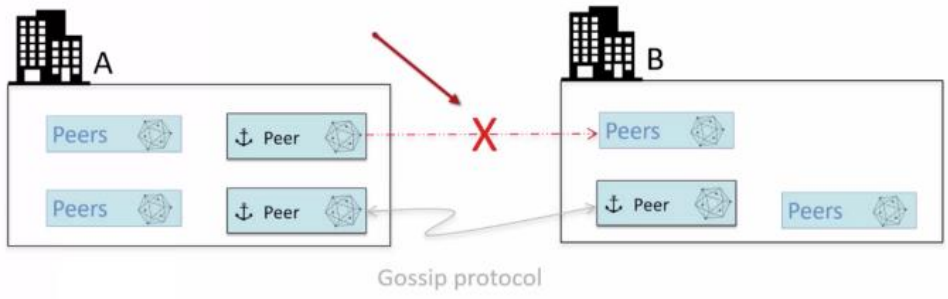


Ilustración 19. Relación entre peers *o nodos de una organización y el nodo* orderer. *Fuente: "Hyperledger Fabric — Conceptos y Tipos de Nodos"*

Este tipo de nodos es básico para la organización de la red y **deberá existir por lo menos una unidad de estos en cada organización**.

d) **Endorsers:**

Partiendo de la definición de los *Endorsers Peer* del Glosario de Telefónica para su espacio dedicado a Blockchain, este tipo de nodos tienen la función de *"simular el resultado de la transacción"*.

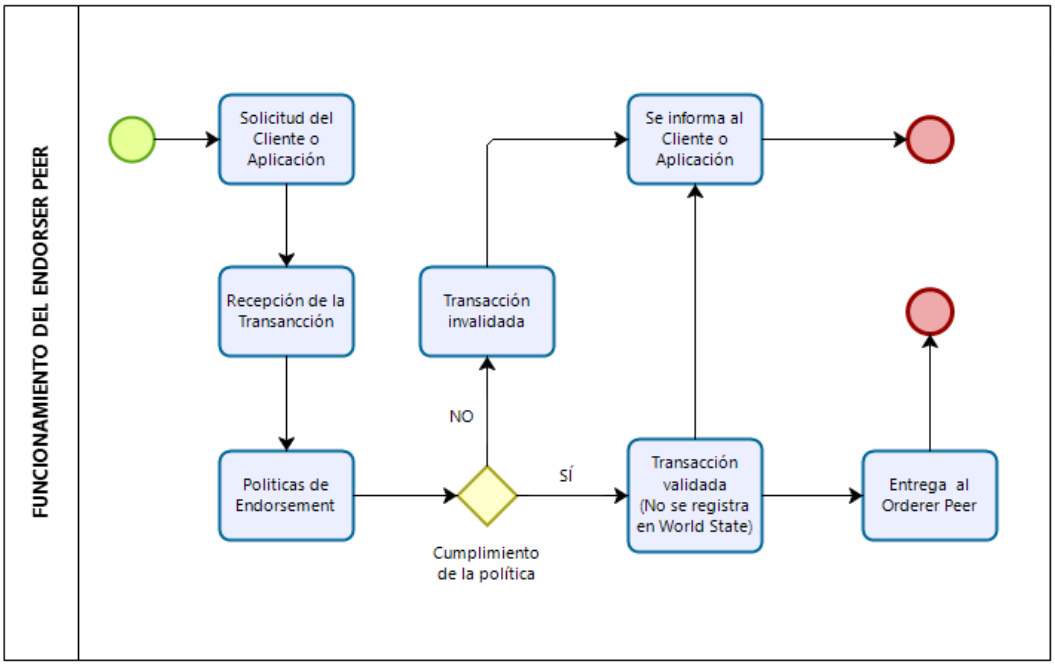


Ilustración 20. Funcionamiento del Endorser Peer. Fuente: Elaboración propia.

Como se muestra en el diagrama de procesos, solicitada la transacción por parte de un Cliente o una Aplicación, se hace la recepción de esta por los *Endorser* y se realiza la simulación bajo la política de ***Endorsement***, y que deberá cumplirse para que la transacción sea validada y transmitida o entregada al *Orderer Peer*.

Se debe destacar que, previo a la entrega a los *Orderer*, estos nodos no dejan ningún tipo de registro en el *World State*.

Es cierto que, para ser puristas, **el Rol o función *Committer***, que es aquella que **verifica las propuestas de transacciones** tras su simulación y **da validez al resultado** de estas antes de grabarlas en la Blockchain, puede estar asociado a un nodo independiente, pero para este proyecto, este Rol lo asume también el *Endorser Peer*, siendo esta una posibilidad que admite la red de HLF.

El proceso continuará en estos nodos de servicio de ordenación, que serán los destinados a solucionar el anclaje de los bloques y derivar esa información a los de *Leaders peers*, como se indicó en el apartado específico de estos nodos.

e) **Regular**:

De los posibles nodos que pueden ser parte de la estructura de una organización **son los *peers* más básicos**, y a diferencia de los *anchors peer* no son conocidos por otras organizaciones que pertenezcan al canal.

Estos nodos tienen la función de llevar el mantenimiento actualizado del *ledger* o libro contable y de sincronizarse con el resto de nodos.

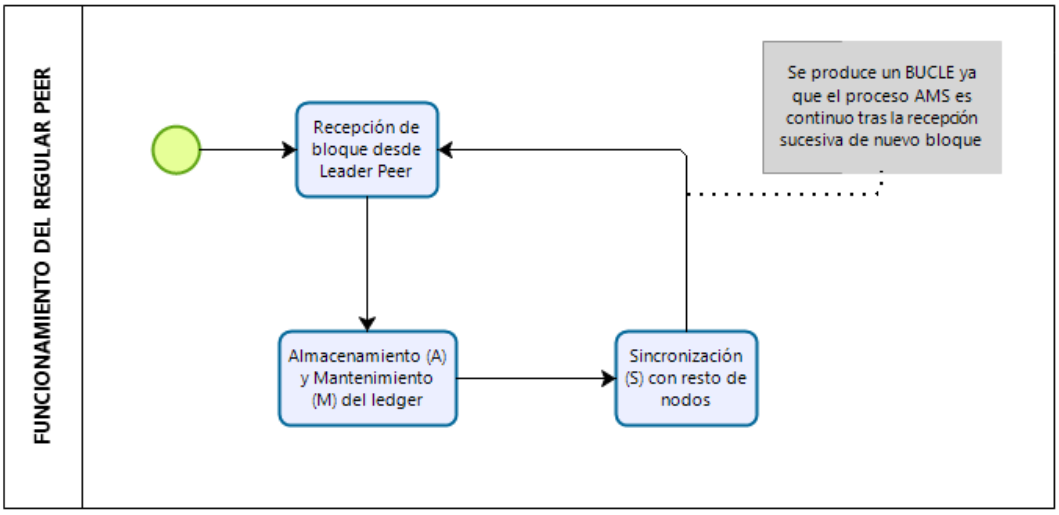


Ilustración 21. Funcionamiento del Regular Peer. Fuente: Elaboración propia.

A los procesos de **Almacenamiento, Mantenimiento y Sincronización**, se le ha dado un acrónimo con las iniciales de estas tres actividades, que es **AMS**, y así se ha destacado en el esquema del funcionamiento de este nodo.

f) **Cliente**:

Se podría definir como el **nodo "punto de partida" u origen**, ya que será el que solicite y envíe la transacción que ha de recibir el *Endorser Peer*.

Estos nodos también son los que, tras la simulación de la transacción por parte de los *Endorsers*, recibirán la respuesta de estos, con la transacción validada o no validada.

Es importante clarificar que existen distintos tipos de clientes, como usuarios finales, APIs, Back-End, etc.

4. **Canales**

Según se indicó con anterioridad, los canales son espacios restringidos para determinados participantes, que están formados por dos o más organizaciones, y que utilizan esa red particular para realizar actividades empresariales o negocios que no son visibles para otros miembros de la red.

Este tipo de canales son los denominados “De aplicación”, y estos canales necesitan de los siguientes componentes básicos para la conformación del mismo:

- Dos tipos de *peers*: un *orderer* para el canal y un *anchor* para cada una de las organizaciones que participen en este.
- Un ledger: que será compartido por todos y cada uno de los miembros del canal.
- Por lo menos un *chaincode*: que será el instrumento para que se pueda ejecutar la acción o acciones integradas en el proceso de transacción entre los participantes del canal.

De esta forma, **cada canal contará de su propia *Blockchain***, que no tendrá ninguna relación con las cadenas de bloques que están asociadas a otros canales.

En la imagen siguiente se muestra con claridad este hecho, donde cada canal consta de su propio ***Ledger*** y como mínimo un ***chaincode***.

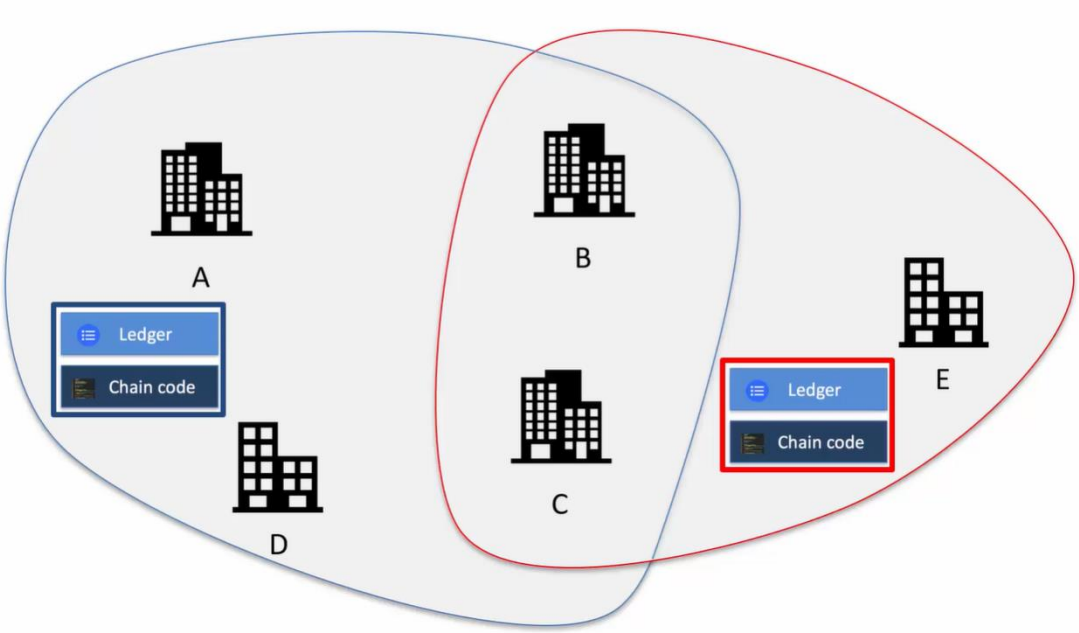


Ilustración 22. Distribución de canales en función de las organizaciones. Fuente: Rajeev Sakhuja (raj) Curso de Hyperledger Fabric Network Design & Setup .

5. **Colecciones**

Es la manera que tiene HLF desde la versión 1.2, para aumentar el grado de privacidad dentro de un mismo canal, y de **utilizar de forma permisionada, determinados datos privados**.

Para conseguir crear una colección, es necesario tener dos elementos fundamentales:

- Los datos: son datos privados que solo pueden ver determinados nodos autorizados dentro del canal.

Como sucede en el caso de la comunicación P2P de los *Anchors*, se utiliza el protocolo *Gossip* para esta acción, y los datos se almacenan en una BBDD privada de los nodos autorizados.

- Los hashes: que, a diferencia de los datos privados, estos sí que son compartidos por todos los *peers* del canal, y sirve de evidencia de que se ha realizado la transacción de los datos privados, evitando que la información que acumulan pueda ser manipulada o dañada.

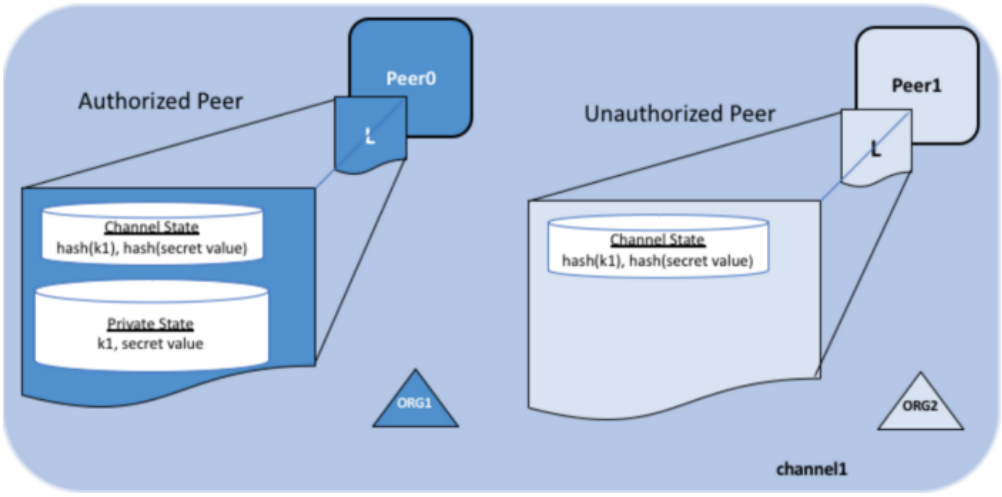


Ilustración 23. Información asociada al peer en función de su nivel de autorización. Fuente: Hyperledger Fabric 1.2 Docs.

Una de las cuestiones claves para el mantenimiento de la privacidad de los datos, es que **esta información no es visible para el *Orderer Peer***, lo que impide que el resto de organizaciones, que no tengan permiso a esas colecciones, puedan visualizar los datos.

El planteamiento del uso de las colecciones está intrínsecamente ligado al RGPD, de tal manera que uno de sus puntos más controvertidos de este reglamento, el **derecho al olvido**, se puede resolver mediante esta funcionalidad de Fabric, a través de las **Políticas de Acceso** que son similares a las Políticas de *Endorsement*, limitando quien tiene acceso a los datos privados; con la definición de determinados bloques con el valor **blocktolive** que serán los que se podrán borrar de forma automática; y finalmente, la eliminación directa de manera **manual** de la base de datos.

6. **Chaincode**

Es la denominación en Fabric para los contratos inteligentes o Smart Contract. Y que según la definición que se establece en el Glosario de hyperledger-fabric.readthedocs.io, es: *"un código, invocado por una aplicación cliente externa a la red Blockchain, que gestiona el acceso y las modificaciones a un conjunto de pares clave-valor en el **World State** a través de una transacción. En Hyperledger Fabric, los contratos inteligentes se*

empaquetan como código de cadena. El encadenamiento se instala en los pares y luego se define y se usa en uno o más canales".

Por lo tanto, la función principal de este código es activar el proceso de transacción y añadir la información al libro mayor.

Diseño de Hyperledger Fabric para el proyecto D.ONE+

Ya conocemos cuales son los componentes principales de HLF, y a partir de este punto podemos implementar el diseño de nuestra red Blockchain en Fabric.

Para ello, tomaremos de referencia el diagrama que se mostrará más adelante, y que servirá para clarificar de qué elementos se compone nuestra red, partiendo de la idea de que tenemos un mínimo de tres organizaciones, concretamente tres AAPP, que están incluidas en un único canal y que para establecer niveles de privacidad dos a dos, se han creado colecciones entre la AAPP A y B, y B y C.

Para mantener la coherencia argumental del proyecto **D.One+**, se considerará que la AAPP_A es el **Ayuntamiento 1 de Gran Canaria**, mientras que las otras dos organizaciones, podrán ser otras instituciones públicas como, por ejemplo, otros dos ayuntamientos de la **Mancomunidad de Medianías**.

En este punto, se quiere hacer un breve inciso, para comentar que la **Isla de Gran Canaria**, donde se sitúa el Ayuntamiento 1 de Gran Canaria, posee **3 mancomunidades** que están agrupadas por los siguientes municipios:

- Mancomunidad del Norte: Agaete, Artenara, Arucas, Firgas, Gáldar, La Aldea de San Nicolás, Moya, Santa María de Guía, Tejeda, Teror y Valleseco.
- Mancomunidad del Sureste: Agüimes, Ingenio y Santa Lucía de Tirajana.
- Mancomunidad de Medianías: San Bartolomé de Tirajana, Valsequillo de Gran Canaria, Tejeda, Vega de San Mateo y Villa de Santa Brígida.

Que para el caso que nos ocupa, estas organizaciones **son corporaciones intermunicipales que aglutinan a una serie de ayuntamientos** de determinadas zonas de la isla.

Volviendo al concepto de las colecciones a crear para el proyecto, es evidente que también se puede plantear otra colección más, fijada para A y C, pero entendemos que, para hacer entender la solución, es suficiente con las dos primeras mencionadas.

Teniendo esta cuestión presente, se arma el diseño de la HLF de **D.One+**, que además de las comentadas particularidades en el uso de datos privados y la generación de colecciones, estará formada por los siguientes componentes que se exponen en la tabla:

Entidad Pública con su MSP y CA				
AAPP_A	AAPP_B	AAPP_C		
Tipos de nodos o peers por organización				
Orderer	Leader	Anchor	Endorser	Regular
Colecciones				
AAPP_A - AAPP_B	AAPP_B - AAPP_C			

Tabla 4. Elementos integrados en el esquema de diseño de **D.One+**.

Descripción de flujo de trabajo asociado al diseño

El diseño de la red del proyecto **D.One+** estará formado por **un único canal** por defecto, y **tres organizaciones**, como ya se ha comentado con anterioridad.

Los **componentes tipo de una organización** son los que se muestran a continuación:

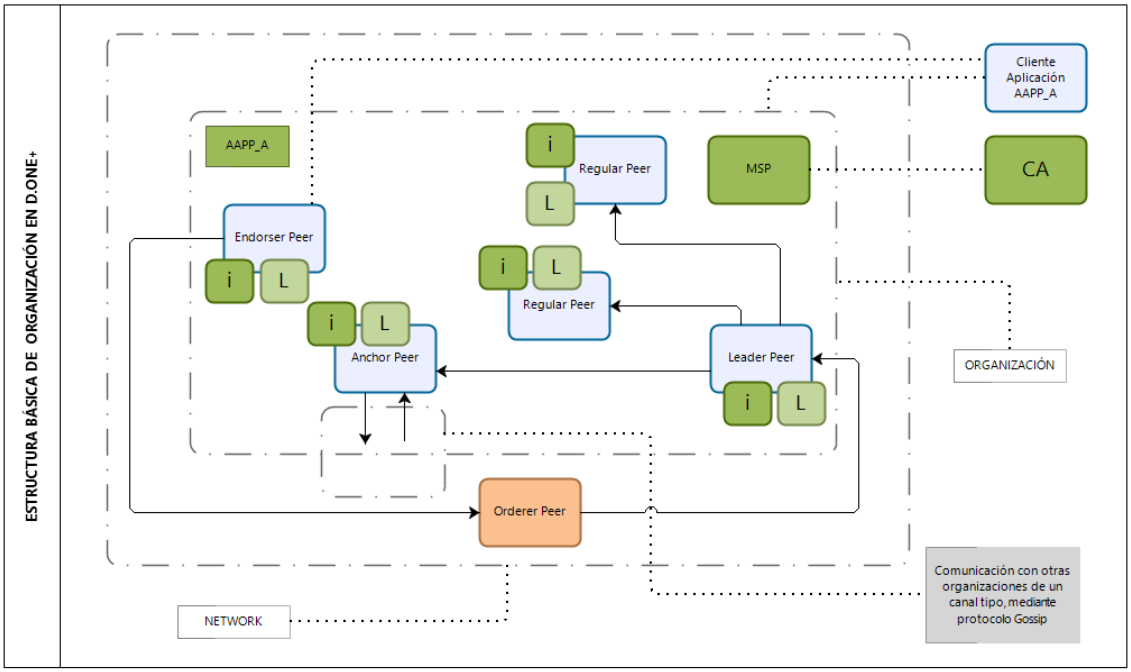


Ilustración 24.. Estructura básica de una organización tipo en **D.One+**. Fuente: Elaboración propia.

Ledger o Libro contable en función de la organización

Identidad en función de la CA de la organización

Cada organización está formada por los *peers* indicados en la tabla anterior, y el flujo de trabajo desde que se produce la solicitud y envío de transferencia o consulta hasta que esta se une a un bloque, y queda registrado en el *Ledger* de la organización es el siguiente:

Paso 1. Solicitud del cliente o aplicación a la AAPP_A, que estará sujeto alguno de los procesos que se detallan en el apartado dedicado a los **procesos establecidos para la implementación de Blockchain en D.One+.**

Paso 2. Se realiza la simulación de la transacción a través del *Endorser Peer* de la organización y que es el que recepciona esta propuesta de transacción. Internamente se produce la invocación de un *chaincode* que tras el análisis del cumplimiento o no de las políticas de *Endorsement*, se valida o no la transacción.

En el caso de que se trate de una consulta por parte de una administración, será necesario ver si esa administración es interna o externa a la *network* y de qué autorizaciones consta, mediante una matriz de permisos. En función de eso, como se verá de manera detallada en el apartado de procesos implicados en nuestra Blockchain, primero podrá saberse si puede realizar la consulta, y en caso afirmativo, a qué tipo de datos puede acceder la organización solicitante dependiendo de los permisos que se le haya conferido.

Paso 3. Durante el proceso de simulación en el *Endorser*, no se realiza la actualización del *Ledger*, y este nodo firma el **paquete *Reads and Writes* o RW**, que son datos de lectura y escritura generados durante la simulación. Y que tras su firma son devueltos al Cliente o la Aplicación, para ser utilizados en los siguientes pasos del flujo de la transacción.

Paso 4. Dado que como se indicado en la descripción del *Endorser Peer*, este también tendrá el rol de *Committer*, se activará el mecanismo de *Ordering Service*, para la validación o invalidación de la transacción.

Paso 5. En caso de que cumpla con las políticas establecidas, se valida la transacción gracias a la función *Committer* del *Endorser Peer*, y se informa al Cliente o Aplicación. Si se produce la invalidez de la transacción también se informará al Cliente/Aplicación.

Paso 6. Posteriormente será el *Orderer Peer*, el que ejecute las acciones de formación del bloque a asociar a la cadena y lo remita al *Leader Peer*.

Paso 7. El nodo *Leader*, realiza la distribución de la cadena de bloques al resto de *Peers*, entre ellos a los dos *Regular Peers* que hemos incluido en nuestra organización, con la idea de mejorar la resiliencia de nuestra red.

El paso que queda pendiente necesita de la implicación de las otras organizaciones que son parte del canal único del proyecto **D.One+**, y que se muestran de manera gráfica en el **Esquema de Diseño de la HLF para D.One+**, que se encuentra al final de la explicación de este último paso:

Paso 8. Dada la existencia de tres organizaciones, la comunicación entre las mismas se realizará por el único nodo conocido por el resto de las organizaciones, el *Anchor Peer*, efectuándose este cruce de comunicaciones mediante el protocolo denominado *Gossip*.

Además, será desde estos nodos de las organizaciones, donde se establezcan las colecciones y se formen las relaciones de cruce de información, en función de lo indicado en la tabla de elementos integrados en el esquema de diseño de **D.One+**.

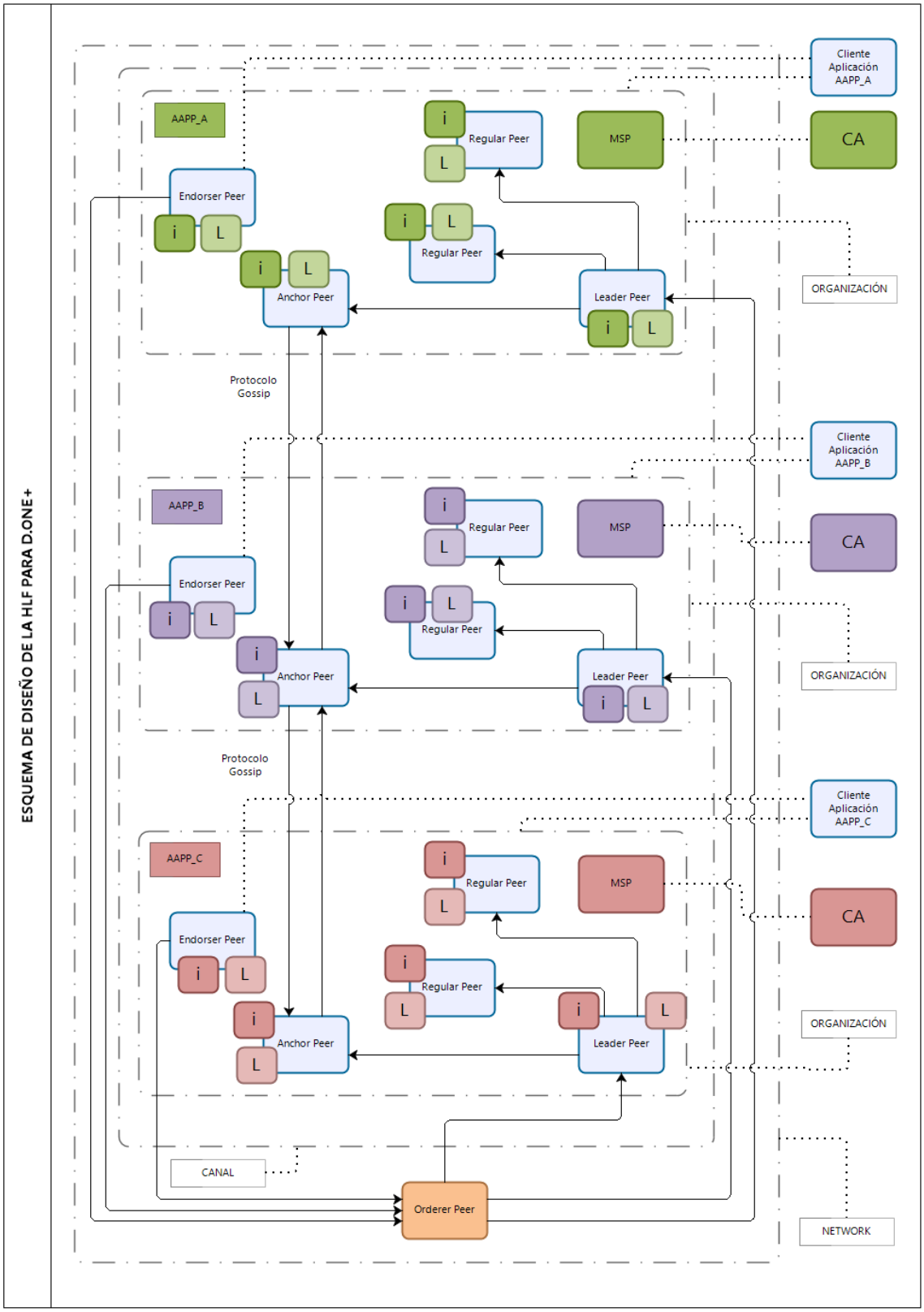


Ilustración 25. Esquema de diseño de la red de HLF de **D.One+**. Fuente: Elaboración propia.

Como se verá en apartados sucesivos, este esquema se complementa con cada uno de los procesos que se han establecido para **D.One+**, y más concretamente, los de "**Consulta por parte del ciudadano**", de "**Carga de registros en MDM**", de "**Consulta de AAPP interna al sistema *Blockchain***", de "**Consulta AAPP externa al sistema *Blockchain***", y de "**Borrado de registros maestros en *Blockchain***", este último proceso, se facilita con la inclusión de las colecciones, según se ha establecido en apartados anteriores.

Procesos establecidos para la implementación de Blockchain en D.One+

A continuación, se detallan los procesos y casuísticas más relevantes de este proyecto, teniendo en cuenta todos los elementos que pueden intervenir o participar en él. Se verá en cada uno de estos ejemplos cómo en la capa de *Blockchain* se ejecutan los *Smart Contract* que llevan a cabo todas las tareas necesarias para obtener, actualizar y borrar los registros maestros, así como la revisión de permisos de accesos a los registros y documentación del ciudadano, cumplimiento de condiciones para seguir con la ejecución de código, etc.

- **Proceso de consulta por parte del ciudadano.**

Para este proyecto se propone dotar al sistema de un servicio como el que ofrece "*carpeta ciudadana"* [https://sede.administracion.gob.es/carpeta], donde el/la ciudadano/a accede autenticado a la infraestructura desarrollada y desde donde puede consultar toda su documentación existente en los sistemas gestores internos de cada administración pública participante en esta red de nodos.

Para ello el ciudadano, mediante una solicitud de parte, inicia la solicitud para acceder a su repositorio de documentos mediante autenticación a través del certificado digital. Con ello nos aseguramos que los datos de acceso al sistema son correctos y seguros.

Primeramente, al ingresar correctamente el/la usuario/a en la sede electrónica, se invoca al *Smart Contract* que comprueba que el usuario y sus datos de identidad existen y están correctos en los registros maestros que están almacenados en cada nodo de la red privada en *Blockchain*.

Tras obtener estos datos, el siguiente *Smart Contract* comprueba que el usuario no esté fallecido para continuar con los procesos. Este *chaincode*, a través del campo de fecha de fallecimiento del registro maestro obtenido, verifica si ha fallecido la persona. Si es así, rechaza la solicitud de consulta y ésta termina sin resultado para el usuario. De esta forma se evita el acceso a la información de la persona fallecida de forma fraudulenta con certificados electrónicos que aún no han sido revocados o anulados. Por el contrario, si el *Smart Contract* comprueba que la persona que representa el certificado electrónico está viva, invoca a otro *Smart Contract* que verifica si los datos almacenados del ciudadano en la *Blockchain*, en concreto el DNI, nombre y apellidos son correctos. Si los datos se deben actualizar, se llama a otro *chaincode* encargado de volcar la información correcta en el registro maestro para tener siempre disponible el dato más fiable de la persona a disposición del interesado y de las administraciones públicas.

Una vez que se ha validado o actualizado al usuario en la *Blockchain*, se procede a leer del registro maestro de la *Blockchain* las posibles variantes del DNI para así aumentar la probabilidad de éxito en la búsqueda de documentación del/la usuario/a en los sistemas gestores internos de las administraciones públicas consultadas. Esto es así porque el DNI en estos sistemas puede aparecer con múltiples formatos.

Cuando finaliza la búsqueda de la información solicitada, se vuelcan todos los documentos encontrados para que el usuario pueda consultarlos.

En la siguiente figura se muestra un diagrama de procesos de consulta por parte del usuario:

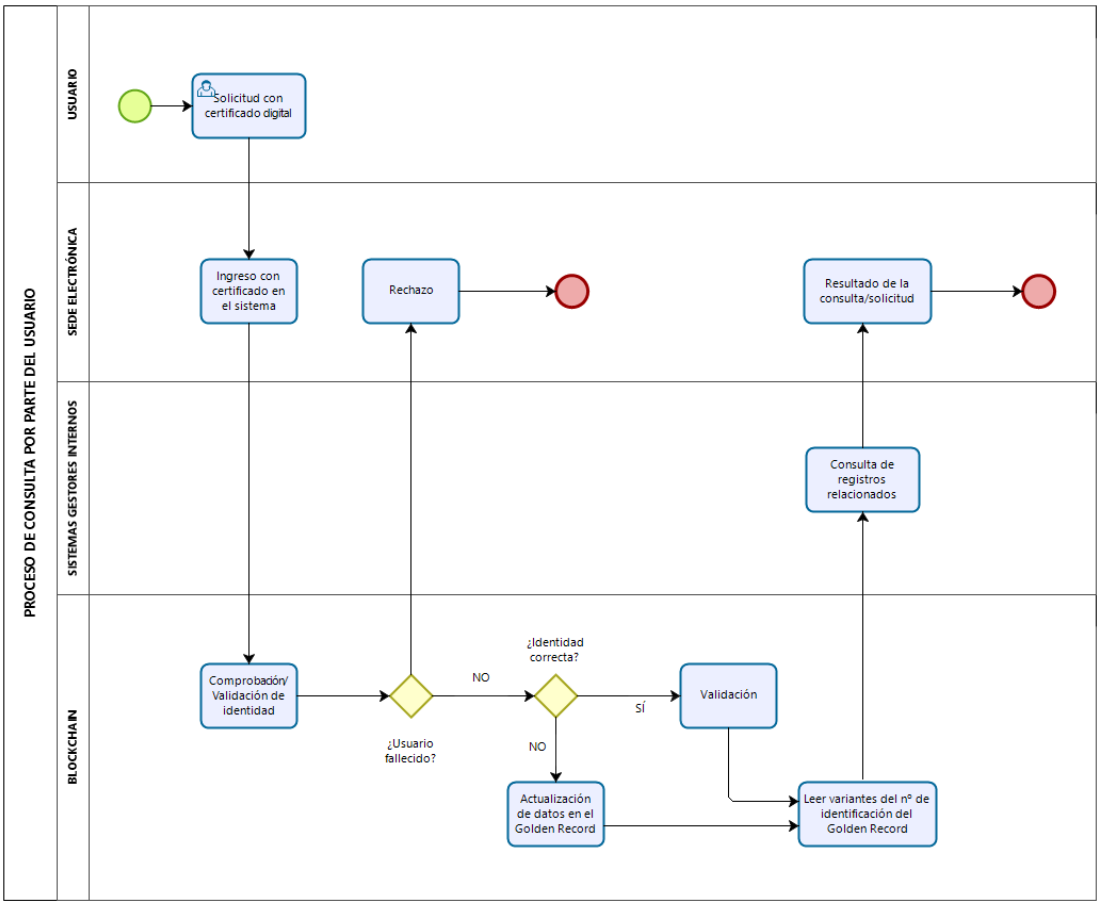


Ilustración 26. Imagen exportada de los procesos planteados para la consulta por parte del usuario.

- **Proceso de carga de registros en MDM.**

Este apartado trata sobre cómo la infraestructura diseñada, primeramente, mediante un proceso inicial y luego a través de un proceso *batch*, trata de obtener el mejor dato del ciudadano almacenado en los sistemas gestores de las administraciones públicas.

Al iniciar el proceso de carga de los datos en el MDM, se invoca la extracción de los datos ya mencionados de interés del ciudadano en los sistemas gestores internos de las administraciones públicas. Estos datos son almacenados en un fichero legible y válido para la capa de Big Data, que será la encargada de transformar los datos para obtener un primer registro del ciudadano.

Para el primer proceso de *deduplicación*, técnica ya comentada para la obtención de la primera versión del registro único del ciudadano, se propone la carga de los datos en *Hadoop*, para que luego, en los procesos de limpieza y *Record Linkage* ya citados en este capítulo, se adquiera una primera base de datos de registros maestros de los usuarios. En esta primera versión de los datos también se añadirán las posibles variantes del DNI del usuario en un campo de cada registro. Estas variantes del DNI nos servirán en un futuro para aumentar el porcentaje de coincidencia del número identificador del ciudadano facilitado con las variantes del DNI que tengan almacenados los sistemas gestores internos de él. De esta manera podríamos obtener toda la información posible del usuario.

Para mejorar los registros maestros del usuario que se almacenarán en los nodos de la *Blockchain*, se presenta la posibilidad de rescatar mejores datos aún del ciudadano al consultar al PID, plataforma que ofrece datos contrastados y actualizados del ciudadano que están almacenados en instituciones y organismos públicos. A continuación, se enumeran las fuentes de información prioritarias para actualizar los campos almacenados en la base de datos de la capa de *Big Data* que se registrarán finalmente en la base de datos almacenadas en la *Blockchain*.

- Al consultar al Servicio de verificación y consulta de datos de identidad (SVDI) del PID, contra las bases de datos de la Dirección General de la Policía, se rescatan los siguientes campos más fiables:
    1. DNI/NIF/CIF/Nº Tarjeta de Residencia/Nº de Pasaporte
    2. Nombre/Razón Social
    3. Apellido 1
    4. Apellido 2
- Al consultar al Servicio de verificación de datos de residencia con fecha de la última variación (SVDR) del PID, encargado de consultar al INE los datos de empadronamiento del ciudadano, se rescatan los campos más fiables:
    1. Dirección postal. Compuesta por los siguientes campos:
        - Código Calle
        - Tipo Vía
        - Literal Vía
        - Número
        - Letra
        - Escalera

- Bloque
- Planta
- Puerta
- Código Postal
- Población
- Provincia
- País

- Al consultar al Ministerio de justicia a través del PID, se rescatan los campos más fiables:
    1. Fecha de fallecimiento
- Al consultar a los sistemas gestores internos de las AAPP, se rescatan los campos más fiables:
    1. Dirección electrónica
    2. Teléfono fijo, teléfono móvil:
- Al crear las variantes del DNI en la plataforma de Big Data, este array de valores es almacenado como un campo en el registro propuesto como registro maestro:
    1. Variantes del Número de identificación

En la última tarea de "Actualización de BBDD con registros deduplicados" de la plataforma *Big Data* desarrollada, los campos que son actualizados por el PID definidos anteriormente son volcados en dicha BBDD para finalmente, actualizar los registros maestros de la *Blockchain* gracias al *Smart Contract* creado para ello. En la figura XX se puede visualizar el diagrama de procesos que representa lo descrito:

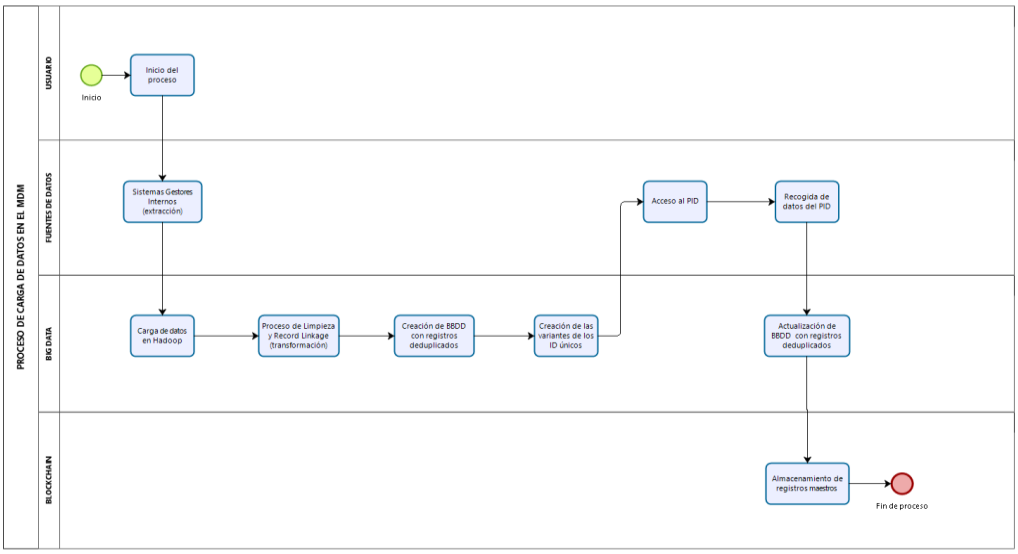


Ilustración 27. Imagen exportada de los procesos planteados para la carga de los datos en el MDM.

- **Proceso de consulta de AAPP interna al sistema Blockchain.**

En este proceso, a través de una solicitud de oficio, es la Administración Pública interna quien solicita información del ciudadano, tanto de identificación y residencia como de documentos de éste. Se considera administración pública interna a aquella que consulta y obtiene los documentos deseados desde sus propios sistemas gestores internos.

Para que la comunicación entre las AAPP y la infraestructura diseñada sea posible, se propone el desarrollo de una API de consulta para administraciones públicas con la cual ésta se puede identificar gracias al sello de órgano y desde donde lanza la solicitud de consulta hacia la red de nodos privada.

Una vez que se logra la autenticación del organismo con éxito, se invoca al *Smart Contract* que inicia el proceso de consulta de los datos y comprueba en la matriz de permisos que dicha administración consultora tiene permisos para acceder a la información requerida. Para ello, el *Smart Contract* accede a una tabla relacional donde cada fila representa a una Administración Pública (Ayuntamiento, por ejemplo) y cada columna representa a cada una de las administraciones públicas presentes en el ecosistema de entidades colaboradoras en el proyecto. Se marcará la celda a Sí si la entidad de la fila tiene acceso a la información de la entidad de su columna. Se pondrá el valor de la celda a No si no la tiene. Cada administración tiene permisos absolutos sobre sus propios sistemas gestores internos.

En la siguiente tabla se muestra un ejemplo de matriz de permisos donde los permisos son simétricos entre tres entidades. La AAPP_A y C se dan permiso entre ellas y la AAPP_B no puede ser accedida por AAPP externas.

Entidad Pública	Permiso AAPP_A	Permiso AAPP_B	Permiso AAPP_C
Permiso AAPP_A	Sí	No	Sí
Permiso AAPP_B	No	Sí	No
Permiso AAPP_C	Sí	No	Sí

Tabla 5. Ejemplo de matriz de permisos del Sistema de la Información.

Seguidamente, si la entidad que solicita acceso a la información no tiene permisos para ello, finalizan los procesos sin resultados de información. En cambio, si tiene habilitado los permisos, se ejecuta el *chaincode* que verifica el tipo de consulta de los datos, interna o externa.

Si la información que se precisa obtener es sobre una consulta interna, es decir, datos de identificación y/o residencia, éstos se consultan a los registros maestros brindando los resultados a la API de consulta desde donde se realiza la llamada a los servicios. Si la consulta es de tipo externo, es decir, se quiere recuperar la *carpeta ciudadana* del ciudadano, con la llamada a otro *Smart Contract* se solicita al registro de la *Blockchain* las variantes del DNI de interés para así

aumentar la probabilidad de éxito en la búsqueda de documentación del/la usuario/a en los sistemas gestores internos de las administraciones públicas consultadas.

Cuando finaliza la búsqueda de la información solicitada, se vuelcan todos los documentos encontrados para que la administración pública solicitante pueda consultarlos.

En la siguiente figura se muestra un diagrama de consulta por parte de la administración pública interna:

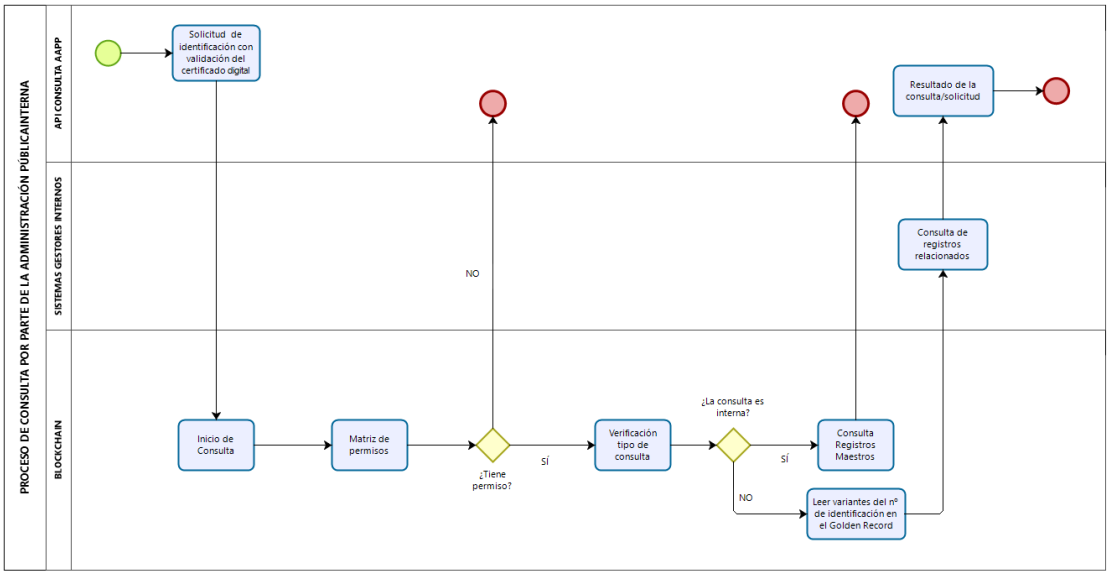


Ilustración 28. Imagen exportada de los procesos planteados para la consulta por parte de la AAPP interna.

- **Proceso de consulta AAPP externa al sistema Blockchain.**

En este proceso, a través de una solicitud de oficio, es la Administración Pública externa quien solicita información del ciudadano, tanto de identificación y residencia como de documentos de éste. Se considera administración pública externa a aquella que consulta y obtiene los documentos deseados de sistemas gestores internos de otra administración pública.

Para este proceso se procede de igual manera que en la consulta de una Administración interna, considerando que la administración pública externa se ha identificado con éxito gracias al sello de órgano.

Se facilita en la siguiente figura el diagrama de procesos de la consulta.

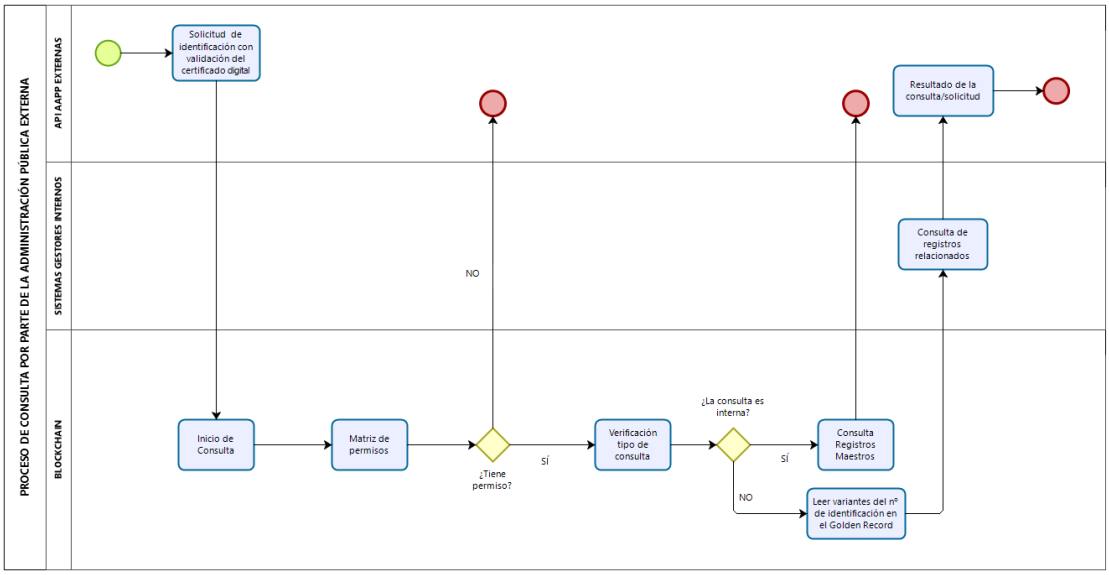


Ilustración 29. Imagen exportada de los procesos planteados para la consulta por parte de la AAPP externa.

- **Proceso de Borrado de registros maestros en Blockchain**

Como último proceso, cabe destacar el proceso de borrado de información en la red de nodos de la *Blockchain*.

Según el Esquema Nacional de Interoperabilidad, en el capítulo X, apartado K, se argumenta la posibilidad del borrado de la información, si así lo establece el resultado del procedimiento de evaluación documental.

En la siguiente figura se muestra el diagrama de procesos para dicho borrado, donde la solicitud de borrado iniciada por el interesado es revisada por el delegado de protección de datos. Si procede el borrado de los datos, se invoca al Smart Contract encargado de ello, marcando un campo a 0 en la base de datos, sin llegar a borrar los datos del registro maestro ubicado en los nodos de la *Blockchain*.

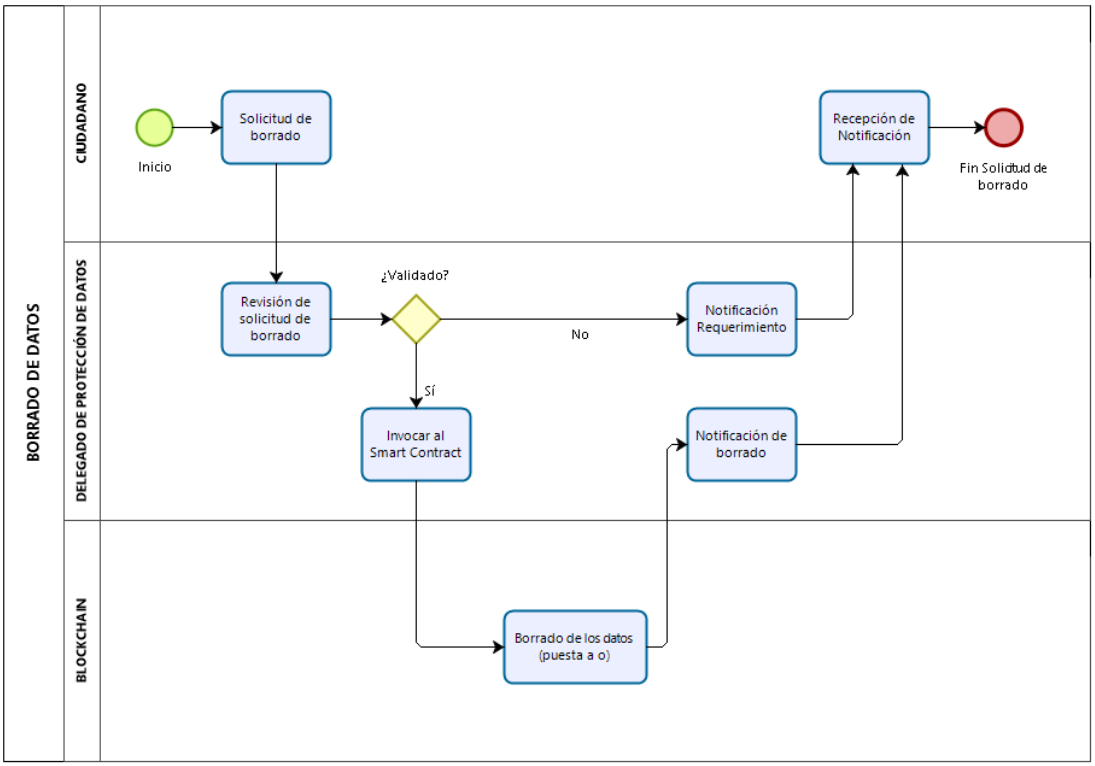


Ilustración 30. Imagen exportada de los procesos planteados para la solicitud de borrado de los datos.

3. Justificación de almacenamiento de los datos en formato JSON

Se ha escogido este formato para los datos puesto que todas las tecnologías escogidas y presentes en la plataforma *Big Data, Smart Contracts*, BBDD para los registros maestros, etc. son completamente compatibles con este formato semiestructurado de los datos. JSON, al carecer de un esquema fijo de los datos, nos ofrece la ventaja de poder evolucionar el esquema de los datos sin riesgos ya que los registros maestros pueden coexistir en la base de datos con distintas propiedades (campos de una tabla relacional) y tamaños.

Veamos cómo compatibilizan las distintas tecnologías del proyecto con el formato escogido.

- Hadoop: este entorno de trabajo para arquitecturas en clúster (múltiples nodos) permite el uso del tipo de datos JSON tanto a su entrada para la lectura de grandes volúmenes de datos como a su salida mediante Apache Avro para la generación de los ficheros JSON que contienen los registros que pasan al proceso de limpieza y *Record Linkage* (https://bit.ly/3kPlxJ7).
- LuceneRDD 2018 (Last Version), librería de Spark compatible con formatos JSON (https://bit.ly/3kQ23UU).
- Bases de datos (consultas en formatos JSON): hay múltiples formas para obtener tanto los datos del usuario que están en formato JSON presentes en la primera versión de la base de

datos con registros deduplicados de la capa de *Big Data* y que serán consultados en la plataforma de intermediación de los datos, así como para actualizar los datos tras esta consulta al PID y que se grabarán en la base de datos de registros maestros alojados en la *Blockchain*. Podremos recurrir a un sistema gestor de bases de datos NoSQL, a conversores de datos no estructurados a estructurados y consultas SQL… En este caso, se propone *MongoDB*, un sistema gestor de bases de datos NoSQL que dota al sistema de bases de datos altamente disponibles, distribuidas y escalables horizontalmente. Además, las consultas que se realizan a ellas no son complejas transaccionalmente hablando (no se realizarán *joins* a priori), ya que sólo se realizan consultas fila a fila (https://bit.ly/3fneNkt).

- Consulta de registros maestros de Blockchain desde HLF: Hyperledger Fabric permite almacenar documentos JSON en el Libro Mayor, sin tener que consultar el histórico de transacciones, para así obtener los valores actuales del objeto de estudio, que en nuestro proyecto son los datos maestros del ciudadano en cuestión.

4. Justificación del lenguaje Python para el desarrollo

El lenguaje de programación Python a día de hoy se ha consolidado como uno de los lenguajes más usados en el ámbito de la ciencia de los datos. Este lenguaje de fácil uso, de fácil integración en los sistemas de información, gran escalabilidad, etc. (https://gtnr.it/35RCkqL) y se puede integrar en los siguientes elementos del proyecto:

- Plataforma *Big Data*: tanto en los procesos ETL de extracción, transformación y carga de los datos, en los procesos de creación de variantes del número de identificación, como en los procesos automatizados de lectura/escritura de esta capa se podrá recurrir al lenguaje Python debido a los amplios desarrollos y librerías *Open Source* existentes y disponibles para ello.
- Fuentes externas: el PID, a través de su plataforma MHAP para el intercambio de información entre el emisor y receptor de la llamada a los servicios, provee librerías para consultar los datos con sentencias Python (https://bit.ly/3nXxmiD).
- Blockchain: Hyperledger Fabric nos permite a través del lenguaje Python configurar la red privada de nodos, escribir los *chaincode* necesarios para los procesos dentro de la Blockchain de la infraestructura, crear las variantes de los Números Identificativos de las personas, etc. (https://bit.ly/36RlLdS).

Soluciones tecnológicas de valor añadido

Cuadro de mando

Como valor añadido al proyecto desarrollado, y partiendo del poder que tiene la analítica de los datos para entender la realidad caótica que nos rodea, se plantea el desarrollo de un cuadro de mandos con el cual las administraciones públicas podrán conocer en tiempo real indicadores de interés que describan los datos almacenados en la Blockchain

y la población de estudio desde una perspectiva cuantificada, así como acceder a información más detallada.

El proyecto **D.One+** no sólo trata de resolver el gran problema del acceso a los datos fiables del ciudadano y su documentación, sino también de conocer de manera objetiva, mediante visualizaciones gráficas de diversas índoles, datos tan importantes como habitantes por códigos postales, por calles, de dónde están llegando los datos más fiables, etc. Estos datos son totalmente fidedignos puesto que son consultados y contrastados con fuentes de datos oficiales y pueden ayudar a la toma de decisiones y a la contextualización de diversos problemas a resolver.

Para desplegar un cuadro de mandos en una aplicación web se pueden utilizar lenguajes de programación front-end (Lindley, Front-end Developer Handbook 2018, 2018) como HTML (Lindley, Front-end Developer Handbook 2018, 2018), CSS (Lindley, Front-end Developer Handbook 2018, 2018) y JavaScript (Lindley, Front-end Developer Handbook 2018, 2018), y al mismo tiempo, lenguajes de programación back-end como PHP (¿Qué es PHP?, 2020) y SQL (Oracle, 2020), entre otros.

Tras un estudio de las diferentes librerías y software *Open Source* disponibles en el mercado para la creación de cuadros de mandos, se ha elegido una herramienta de fácil uso e integración en los sistemas de información del proyecto D.one+, *Google Data Studio* (Google Data Studio, 2020).

Google Data Studio permite aglomerar todos los datos en un mismo lugar y transforma los datos sin procesar en las métricas y dimensiones necesarias para el cuadro de mandos sin necesidad de programar. Esta herramienta también permite crear rápidamente informes dinámicos y visualizaciones atractivas para una rápida comprensión del contexto a estudiar.

Para poder conectar con la base de datos de la *Blockchain* que contiene los registros maestros almacenados en formato JSON, *Google Data Studio* cuenta con múltiples conectores desarrollados por partners que vuelcan la información de fuentes de datos externas al entorno su web; conector Custom JSON/CSV/XML, Ad Data + All Other Sources, entre otros.

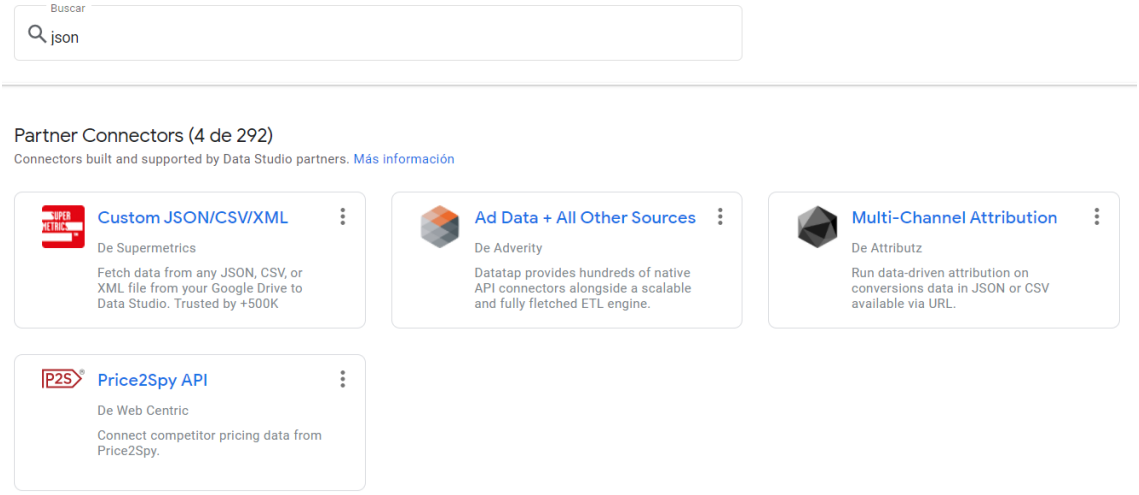


Ilustración 31. Captura de pantalla de la búsqueda de conectores JSON para la herramienta Google Data Studio

Se propone como punto de partida, una interfaz gráfica que cuente con un cuadro de mandos integral (CMI) que persiga la estrategia propuesta de la entidad y pueda ser usado como un sistema de comunicación, de información y de formación.

El CMI es una herramienta muy potente de gestión y consta de múltiples indicadores que facilitan la perspectiva financiera, de procesos internos, de clientes y de aprendizaje y crecimiento. los cuales están ligados a acciones en la administración pública y en la empresa. Los indicadores (KPIs) tienen las siguientes cualidades:

- Específico: especifica una meta concreta.
- Medible: se sabe cómo medirlo; a través de una fórmula, de una conversión, de un dato, etc.
- Alcanzable: los objetivos marcados y que son contrastados con los indicadores escogidos deben ser alcanzables.
- Relevante: hay responsables de su seguimiento y que pueden obtener conclusiones a través él.
- Accesible: está vinculado a un margen de tiempo.

Los diversos tipos de indicadores más comunes en los cuadros de mandos en el mercado son (https://bit.ly/394SQ8M):

- Árboles estratégicos: consta de nodos que representan un objetivo y sus KPI y objetivos de soporte.
- Mapas estratégicos: se muestra cómo se han definido los objetivos para un cuadro de mando y los KPI que miden su progreso se alinean por las perspectivas mencionadas anteriormente.

- Mapas de causa y efecto: permiten ilustrar las relaciones de causa y efecto de un objetivo o KPI del panel Estrategia
- KPIs de *Membership Site:*
    - Miembros activos y crecimiento. Evolución del crecimiento de usuarios en el sistema.
    - Interacción de los usuarios. Uso de la plataforma al mes. Cada cuando tiempo solicitan servicios.
- KPIs de tasa de conversión: cuántos usuarios han sido contactados con éxito tras solicitar los datos maestros del ciudadano.
- KPIs sobre los servicios solicitados en la plataforma creada.

A modo de ejemplo, se muestra un cuadro de mandos generado con la herramienta Google Data Studio escogida donde se aplican filtros de tiempo en la información volcada en el informe y donde se observan indicadores porcentuales, cuantitativos, comparativos y representaciones geográficas que ayudan a tomar decisiones para resolver el problema planteado:

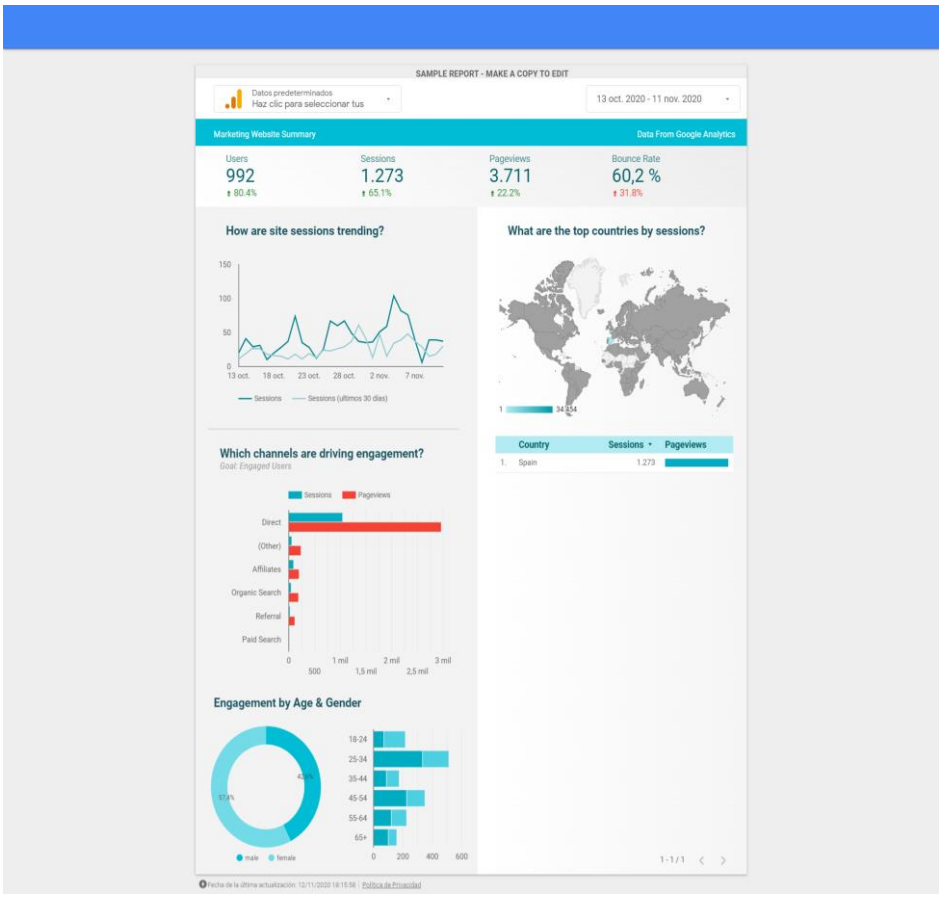


Ilustración 32. Cuadro de Mandos de Ejemplo.

Detección de documentación del usuario a través de datos no estructurados

Como servicio de valor añadido ofrecido en este proyecto, se pretende obtener toda la documentación existente y relacionada con el ciudadano que está vinculada a él de manera no estructurada.

La documentación que está vinculada al ciudadano de manera estructurada es aquella en la cual, en los sistemas gestores internos de las AAPP, está relacionado el identificador único del usuario con la ruta del documento en sí. Este tipo de documentación es fácilmente accesible ya que, una vez identificado al sujeto, sólo se trata de consultar en el sistema de la información al que se ataca los documentos que aparecen vinculados a él.

La documentación que está vinculada al ciudadano de manera no estructurada es aquella en la cual, en los sistemas gestores internos de las AAPP, no está relacionado el identificador único del usuario con la ruta del documento, pero sí dentro del contenido del documento aparece el DNI del usuario. Para poder detectar y rescatar este tipo de documentación, se sugiere, tras identificar de manera segura al usuario, y tras convertir documentos con formato PDF o imágenes a texto gracias a herramientas OCR y librerías Python desarrolladas, localizar todos los documentos del ciudadano con procesos de tokenización, limpieza de datos y matching de identidad con el texto tratado del documento.

Primero, tras iniciarse una solicitud de recuperación de documentación del usuario, una API diseñada verifica que la identidad del usuario es correcta ya que no se puede facilitar información al usuario que no esté relacionada con él.

Si la identidad es correcta, se procede a acceder a datos no estructurados alojados en los servidores de ficheros de los sistemas gestores internos, es decir, documentos que tienen información que no es fácilmente accesible de manera automática. Una vez que ya se accede a dicha información, ésta es tratada por procesos ETL en la plataforma Big Data del proyecto para convertirla a texto para poder detectar el DNI del usuario en el contenido de cada documento. Esto se puede conseguir a través de diversas librerías para el reconocimiento óptico de caracteres, llamadas comúnmente OCR.

Si se comprueba que el DNI del usuario aparece en el texto del documento, llamado también “proceso de matching”, estos documentos serán recuperados de nuevo en los sistemas gestores internos y serán devueltos al usuario en la llamada “carpeta ciudadana” de capítulos anteriores.

Se añade un gráfico que ilustra de manera más intuitiva los procesos que han de realizarse para obtener toda la documentación:

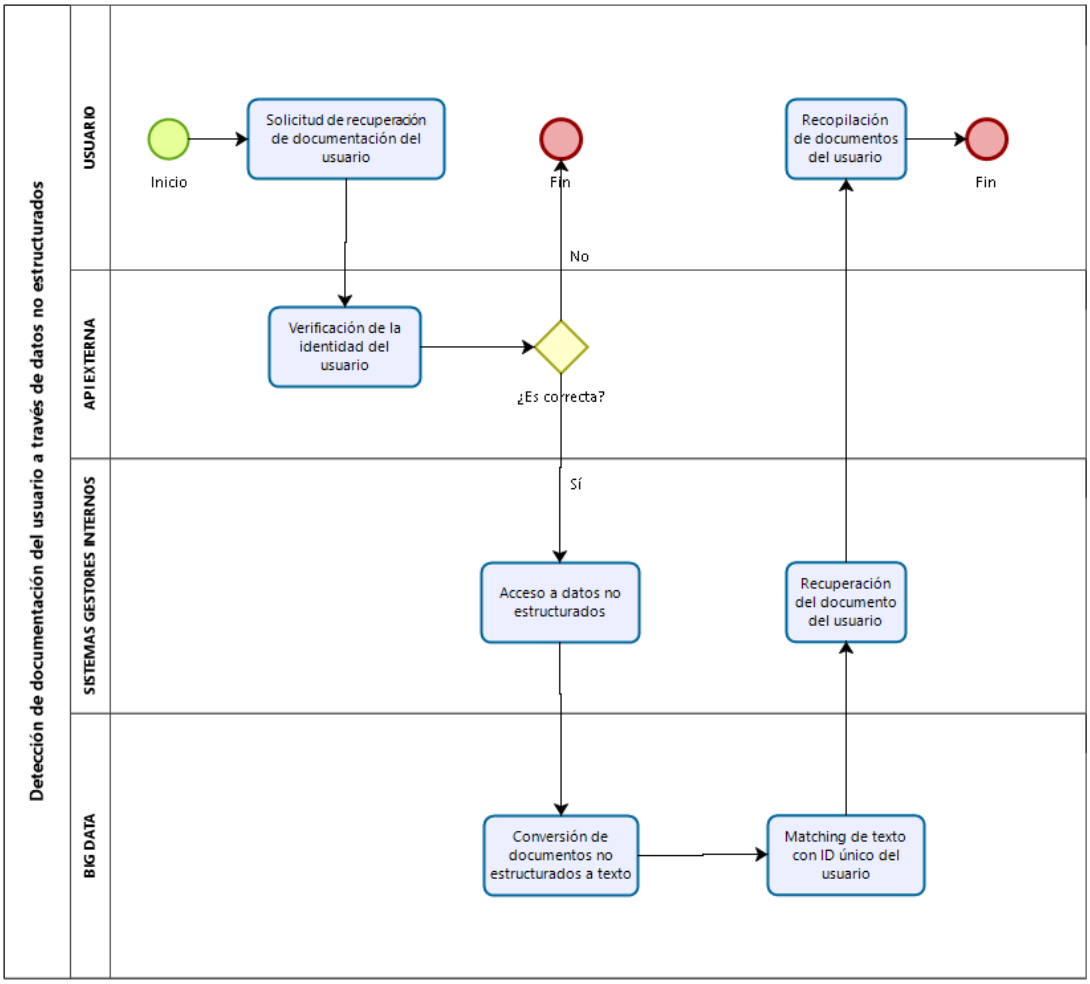


Ilustración 33. Imagen exportada de los procesos planteados para el servicio de detección de documentación del usuario a través de datos no estructurados.

10 OTROS PLANES OPERATIVOS

10.1 PLAN DE MARKETING

Mediante los distintos canales de comunicación que habilitaremos crearemos una potente red publicitaria que mostrará de manera explícita nuestra solución MDM + Big Data + Blockchain y, además, pondremos a disposición de nuestros potenciales clientes un cuestionario específico de registro donde podrán expresar las circunstancias particulares que presenta su organización respecto al problema que nos ocupa y, a través del cual, podremos iniciar nuestra labor de consultoría particular con datos/situaciones reales. De esta manera, agilizaremos sobremanera nuestra capacidad para adecuarnos a las circunstancias específicas de cada proyecto desde el inicio del mismo.

Los canales de distribución que vamos a utilizar son los siguientes:

- La herramienta principal de contacto será nuestra página web (www.d-one.es) donde los clientes podrán visualizar toda la información acerca de los servicios que ofrecemos y acceder a completar el cuestionario al cual hacíamos antes referencia.
- Estaremos presentes en las *redes sociales* Facebook, Twitter y LinkedIn, que, desde nuestro punto de vista, son las más adecuadas para conseguir la atención del público objetivo de nuestro negocio y que, además, tienen gran presencia dentro del mundo empresarial y de la AAPP.
- Otro canal fundamental para nuestro crecimiento será nuestra presencia en *ferias especializadas* del sector donde podremos adquirir un gran número de contactos profesionales para nuestros proyectos. Además, nos dará la oportunidad de observar las novedades de los competidores y así poder realizar mejoras continuas de nuestros servicios ofertados.
- Por último, es conveniente que tanto la startup como sus 5 miembros formen parte de algunos foros vinculados al Blockchain como el Observatorio Internacional del Blockchain, el Instituto Ibérico del Blockchain (https://bit.ly/2IWoJ8A), el foro de Blockchain de Catalunya (https://Blockchaincatalunya.org/) y de la asociación Alastria (https://alastria.io/).

10.2 PLAN DE PROVEEDORES

Nuestro proveedor para la implementación de la plataforma Big Data será Microsoft. Para la implementación de Apache Spark se ha escogido el producto Azure HDInsight (https://bit.ly/337xofz), la implementación que hace Microsoft de Apache Spark en la nube. Además, escogeremos la herramienta Data Lake Storage Gen1, también de Microsoft (https://bit.ly/3706is3) para el almacenamiento de los datos a procesar, entre otros motivos, porque los clústeres de HDInsight procesan fácilmente los datos almacenados en él.

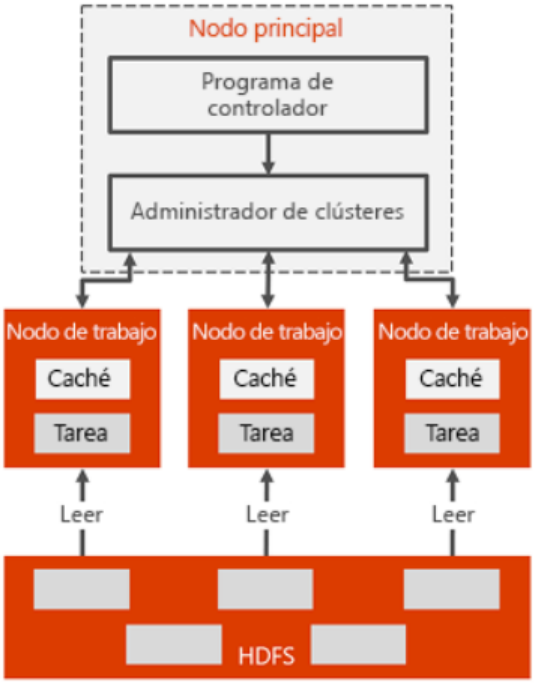


Ilustración 34. Arquitectura de Azure HDInsight.

Se puede acceder a Data Lake Storage Gen1 desde Hadoop (disponible con el clúster de HDInsight) mediante las API REST compatibles con WebHDFS. Data Lake Storage Gen1 está diseñado para habilitar el análisis de los datos almacenados y está optimizado para el rendimiento en escenarios de análisis de datos, incluyendo todas las capacidades de nivel empresarial para la seguridad, manejabilidad, escalabilidad, confiabilidad y disponibilidad de los datos. Además, esta herramienta provee de protección de los datos, esencial para este proyecto, ya que usa autenticación y lista de control de accesos para administrar el acceso a los datos (https://bit.ly/3kTVqkc).

En el siguiente gráfico se observa que Data Lake Storage Gen1 administra tanto datos estructurados, semiestructurados y no estructurados que son facilitados en la siguiente capa de análisis, consultas, etc.

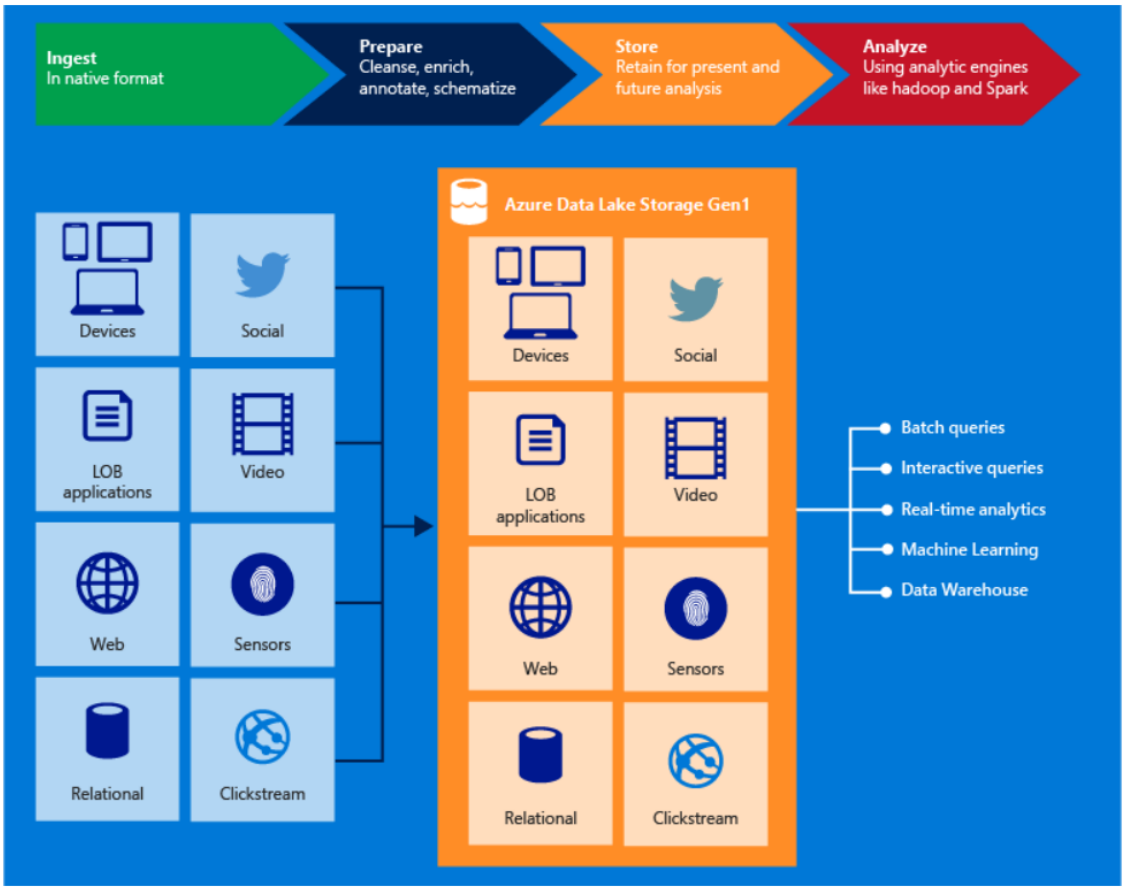


Ilustración 35. Arquitectura de Azure Data Lake Storage Gen1.

En relación a la implementación de la capa Blockchain, la Asociación Alastria será la solución más adecuada para desplegar los nodos de la red.

En cuanto a los equipos informáticos necesarios, internet y telefonía, así como publicidad y marketing se sondeará el mercado y se contratará los servicios que sean más acordes a nuestras **posibilidades crecientes.**

10.3 PLAN DE RECURSOS HUMANOS

Inicialmente no se contratará a personal externo, somos una Start Up. La incorporación de nuevo capital humano estará supeditada a la necesidad motivada por un número de implantaciones de la solución y procesos de mantenimiento que no podamos asumir los actuales integrantes del equipo.

El equipo de D.One será el responsable de gestionar la empresa. Los diferentes perfiles profesionales (administración de empresa, informáticos y analistas de datos) hacen que queden cubiertas las necesidades de las distintas áreas de la misma.

Los perfiles que cubrimos dentro de la empresa D.One son los siguientes (https://ec.europa.eu/esco/portal/occupation):

1. Delegado de Protección de Datos (DPD): este perfil informará y asesorará en materia de protección de datos, revisará el cumplimiento de lo dispuesto en el RGPD y la LOPD, así como tareas de evaluación de impacto a las que se enfrenta la organización.

2. Analista/Científico de datos: detectan e interpretan fuentes de datos ricas, gestionan grandes cantidades de datos, fusionan fuentes de datos, garantizan la coherencia de los conjuntos de datos y crean visualizaciones a fin de ayudar a la comprensión de los datos. Se basan en modelos matemáticos que utilizan datos, presentan y comunican información y conclusiones sobre datos a especialistas y científicos en su equipo y, si fuera necesario, a un público no especializado, y recomiendan formas de aplicar los datos.

3. Gestor/a de seguridad de TIC: proponen y aplican las actualizaciones de seguridad necesarias. Asesoran, apoyan, informan y ofrecen formación y sensibilización en materia de seguridad y toman medidas directas en toda o parte de una red o sistema.

4. Director/a de Sistemas de la Información: orientado más al Data Governance, este perfil define y aplica la estrategia y la gobernanza de las TIC. Determina los recursos necesarios para la aplicación de la estrategia en materia de TIC, anticipa la evolución del mercado de las TIC y las necesidades empresariales. Contribuye al desarrollo del plan estratégico de la organización y se asegura de que la infraestructura de TIC es compatible con las operaciones y prioridades generales de la organización.

5. Experto en Blockchain: este perfil debe tener amplios conocimientos de programación de redes Blockchain con *Hyperledger Fabric* así como lenguajes de programación de Python y formatos de texto estructurados, semiestructurados y no estructurados. Será el desarrollador de los *Smart Contracts* del proyecto que hacen posible las tareas de los procesos en la capa de *Blockchain*.

6. Desarrollador/a de aplicaciones TICs: desarrollan las aplicaciones TIC (software) a partir de los diseños suministrados utilizando lenguas, herramientas, plataformas y experiencias específicas de cada ámbito de aplicación. En nuestro caso, este perfil será fundamental para la creación de APIs presentes en la infraestructura del proyecto, para la interoperabilidad entre las distintas plataformas y como apoyo del experto en Blockchain.

7. Especialista en administración pública: realizan tareas analíticas, conceptuales y prácticas para conocer las administraciones públicas en el ámbito tecnológico, procedimental y legal y poder asesorar en el planteamiento y desarrollo del proyecto **D.One+**.

8. Representante comercial: esta figura se encargará de atraer inversores o dar visibilidad al producto creado en los organismos públicos.

10.4 PLAN DE OPERACIONES

Como se ha mencionado en el plan tecnológico de esta memoria, en el sistema de información del proyecto D.one+ existen procesos de carga de datos, consulta y borrado de los datos necesarios para la creación del Golden Record de registros maestros.

Primeramente, el proceso de carga de datos en el MDM accede a los sistemas gestores internos de las AAPP integradas en la solución propuesta y a la plataforma de intermediación de datos para que, a través de procesamiento Big Data se pueda obtener una base de datos maestra de los ciudadanos que han sido identificados y validados por el sistema a través del servicio de identificación escogido, el certificado digital. Una vez que se obtiene esta base de datos con los datos más fiables del ciudadano, ésta es volcada/actualizada a la base de datos almacenada en los nodos de la Blockchain.

Los procedimientos de consulta de los datos y documentación del ciudadano permite obtener, tanto si la solicitud es de parte como de oficio, los mejores datos del ciudadano almacenados en los registros maestros de la Blockchain, así como la documentación del usuario existente en los sistemas gestores internos de las administraciones públicas, siempre y cuando la matriz de permisos le permita la recuperación de esta información.

Así mismo, en los procesos descritos se presenta la posibilidad de borrar los datos del usuario de los registros maestros de la Blockchain.

A continuación, se muestra el detalle de lo que será nuestro **servicio de mantenimiento estándar para poder realizar los procesos descritos y darle continuidad al proyecto**.

- **Alcance detallado de los trabajos a realizar.**

El contrato cubre la realización de los servicios de mantenimiento y actualización de versiones, así como el apoyo técnico consultor a prestar en el Ayuntamiento, de la solución MDM implantada por la empresa informática **D.One**. En concreto se considerará:

a. MANTENIMIENTO EVOLUTIVO, que incluirá las modificaciones necesarias para incrementar y mejorar la funcionalidad y usabilidad de la solución MDM **D.One+**.

b. MANTENIMIENTO CORRECTIVO, que consistirá en la subsanación de cualquier error o ineficiencia que se detecte en el uso del sistema.

c. MANTENIMIENTO ADAPTATIVO, que consistirá en la modificación del código del sistema según los cambios legislativos que se produzcan durante la vigencia del contrato y que afecten a las funcionalidades legales exigibles a la aplicación. Los cambios legislativos sustantivos que puedan suponer una modificación de más del 50% del código podrá suponer un recargo adicional que será pactado oportunamente si esta circunstancia se produjera.

Estos servicios se concretarán con las siguientes actuaciones:

Actualización y Mantenimiento de la Solución D.One+

- Programa de gestión y seguimiento para mantener el Sistema en operación, subsanando posibles errores y actualizando versiones.
- Asistencia técnica de ingeniería para modificaciones del Sistema y nuevos requerimientos específicos derivados de cambios en la legislación vigente.
- Apoyo para la actualización y mantenimiento de los usuarios.
- Asistencia técnica a los usuarios administradores: elaboración de perfiles, generación de informes y consultas sobre las políticas de seguridad.
- Los trabajos técnicos que requiera la actualización de las versiones de las bases de datos.
- Asistencia técnica para el mantenimiento del Sistema que incluya las instalaciones necesarias cuando se produzcan cambios de equipo y recuperación de versiones y datos.
- Instalación de aplicaciones y personalización de las mismas en casos de implantación de nuevos programas
- Elaboración y mantenimiento de manuales y protocolos derivados de las modificaciones o actualización del Sistema.

Asistencia y Consultoría presencial y remota.

La asistencia consultora se centrará en dos tipos de actuaciones:

a. *Formación en las aplicaciones a los usuarios:*

- Formación a usuarios individuales “in situ”.
- Apoyo directo para la puesta en marcha y mejoras organizativas relacionadas con el funcionamiento operativo de los sistemas prestado por consultores especializados en las materias.
- A iniciativa de la empresa o por requerimiento de varias entidades, se podrán organizar cursos de formación en grupos cuyos destinatarios sean de distintos Ayuntamientos.

b. *Asistencia para la gestión cotidiana de las aplicaciones, que será prestada:*

- Por ingenieros, técnicos informáticos o consultores especializados con conocimiento total del manejo del Sistema.
- Será presencial o remota, en función de la necesidad o demanda del cliente.
- Previa autorización del cliente, se podrán utilizar sistemas telemáticos remotos vía internet para estos fines.

- Compromiso de mantener una línea telefónica para atención directa a los usuarios que estará operativa en horarios de 08.00 a 15.00 horas de lunes a viernes.
- Servicio especial de urgencia 24x7 telefónico para resolver una eventual caída del servicio.

Estimación Horaria

La aplicación estimada mensual de horas a aplicar por asistencia directa técnica y consultora a cada entidad, más la asistencia remota telemática, asistencia telefónica así como la parte proporcional de los cursos de formación que se organicen en el año, las horas de desarrollo que se apliquen a cada aplicación informática por mejoras, cambios legislativos y las actualizaciones de versiones y/o bases de datos se cuantifica en **20 horas mensuales**, produciéndose variaciones en función del tamaño de la organización, que se encuentran promediadas en esta estimación.

10.5 PLAN JURÍDICO - FISCAL

Una vez creada la idea de negocio desarrollada en esta memoria, realizado el estudio de mercado y definidos los elementos claves y diferenciadores en la propuesta de valor, se creará una Start Up con grandes posibilidades de crecimiento y financiación.

D.One SL se creará este año 2020 como Sociedad Limitada (SL, a partir de ahora) formada por 5 socios (que son los componentes actuales de este equipo). Se constituirá con un capital inicial de 10.000 euros, correspondiente a la aportación de 2.000 euros por cada miembro. Al constituirnos como una SL estamos obligados al pago del 15% correspondiente al Impuesto de Sociedades en su tipo reducido para emprendedores. Este tipo se aplica desde el año 2015 para las sociedades constituidas, pero excluyendo las sociedades patrimoniales. Para presentar este tipo se necesita considerar que la sociedad se refiera al inicio de una actividad económica.

Los pasos para registrar debidamente la Startup serán:

- Registrar la marca D.One en la Oficina Española de Patentes y Marcas https://www.oepm.es/es/index.html Al registrar la marca de manera telemática se aplicará un 15% de descuento.
- Registrar el nombre de la empresa en el registro mercantil. Para ello habrá que solicitar el certificado negativo de denominación social para comprobar que no existe una sociedad ya existente con el nombre escogido. Si el nombre de la empresa está disponible, se procederá a la reserva de la denominación de la empresa y la constitución de la sociedad. A continuación, se irá al notario con los estatutos y las normas de la empresa para registrarla y hacer pública la constitución de la empresa. El mismo día de la firma de la escritura se solicitará telemáticamente un NIF provisional para la sociedad, que se convertirá en definitivo cuando se inscriba a la sociedad en el Registro Mercantil.

- Apertura de una cuenta bancaria a nombre de la empresa con un capital de 10.000€ (el capital mínimo exigido por la Ley para la creación de la Startup de 3.000€). Consideramos que la cuenta escogida no genera gastos para la empresa de nueva creación, ni de apertura ni de usos ni de mantenimiento.
- Dar de alta a la sociedad en Hacienda, Seguridad Social y Protección de Datos. A través de servicios locales y regionales de asesoramiento al emprendedor gratuito (PAE): será gratuito para la empresa.
- Publicación de la sociedad en el BORME.
- Alta en el IAE (Impuesto de Actividades Económicas), mediante el modelo 036 o 037 a través de la Agencia Tributaria, donde estableceremos el epígrafe apropiado para la actividad que va a realizarse por D.One, y que dado que no existe un epígrafe que se ajuste a los servicios ofertados por D.One, se opta inicialmente por el epígrafe 999 “Otros servicios no clasificados en otras partes”, tras la búsqueda realizada mediante el Buscador de Actividades de la Agencia Tributaria: https://bit.ly/398pvKH

10.6 PLAN FINANCIERO

En este apartado analizaremos la viabilidad financiera del proyecto, estudiando en profundidad la estructura económica para la puesta en marcha de nuestra empresa y la toma de las decisiones más adecuadas en cada momento. Los estados financieros más importantes y que nos mostrarán la situación tanto del presente, como del futuro son el **balance de situación** y la **cuenta de pérdidas y ganancias**. Aunque su análisis por sí mismo proporciona gran cantidad de información sobre la empresa, se necesita un examen más elaborado de los datos que aportan para emitir un juicio cualitativo sobre si nuestra empresa presenta una situación viable y si está creciendo adecuadamente.

La gestión financiera está destinada a la administración de los recursos financieros, activos y pasivos de la empresa. Con el objetivo de reducir los costes iniciales, los primeros meses no es necesaria la ubicación física del equipo en una oficina porque el trabajo se realizará en remoto, apoyándonos además en una solución construida para su funcionamiento en la nube. El equipo dispondrá de equipos portátiles para poder acceder en todo momento al sistema y a las comunicaciones que se establezcan dando el adecuado servicio a nuestros clientes. Respecto a la parte que concierne al tratamiento contable y fiscal, tal y como comentábamos anteriormente, será realizada completamente por el equipo de D.One que cuenta entre sus integrantes con un experto en el aspecto financiero del proyecto.

Se requerirá una inversión inicial de 10.000 euros para cubrir los costes del primer año de funcionamiento. Los gastos de capital humano, infraestructuras tecnológicas, equipos informáticos, internet, telefonía, publicidad y marketing son los principales gastos que tendrá que soportar la empresa a lo largo de su desarrollo.

- **AÑO 1**: Nuestro primer cliente será el Ayuntamiento 1 de Gran Canaria con el cual se ha firmado un convenio de colaboración para implantar la solución en un entorno de producción con un coste 0 (cero) para la organización. Esta circunstancia nos permitirá desarrollar una herramienta totalmente operativa que podamos trasladar a otras organizaciones a partir del 2º ejercicio.
- **AÑO 2**: Después de implantar con éxito nuestro proyecto en el Ayuntamiento, comenzaremos a buscar clientes realizando visitas periódicas a distintas instituciones y mostrándoles el caso de éxito en San Bartolomé de Tirajana. Nuestra estimación para el 2º año es que podremos contar con 2 nuevas organizaciones a las cuales plantearemos el desarrollo e implantación de un proyecto similar al ya realizado facturando a cada cliente 15.000 Euros (30.000 Euros en total). Para este ejercicio, además, obtendremos el ingreso de 12.000 Euros adicionales correspondientes al servicio de mantenimiento en el Ayuntamiento 1 de Gran Canaria, lo que nos proporcionará un crecimiento óptimo y la obtención de márgenes adecuados que garanticen la viabilidad de nuestra empresa.
- **AÑO 3:** La perspectiva para este ejercicio es la captación de 3 nuevos clientes que proporcionarán a D.One unos ingresos cuantificados de 45.000 Euros (15.000 por implantación), a lo que sumaremos 36.000 Euros por el mantenimiento de los 3 clientes que ya se encuentran en esa fase.
- **AÑO 4:** En este ejercicio los beneficios se verán afectados porque será el primero en el que los integrantes del equipo D.One comenzarán a recibir una retribución por parte de la empresa (concepto de sueldos y salarios). En este punto de la planificación temporal, D.One estará en condiciones de soportar esta circunstancia porque tenemos prevista la captación de 4 nuevos clientes obteniendo de ellos unos ingresos de 60.000 Euros, a lo que sumaremos 72.000 Euros por el mantenimiento de los 6 clientes que ya se encuentran en esa fase.
- **AÑO 5:** Este ejercicio se caracteriza por alcanzar la estabilidad económico-financiera de la empresa. Para el mismo, estimamos unos ingresos de 60.000 Euros por que incrementaremos nuestra cartera de clientes con 4 nuevas organizaciones y, además, obtendremos 120.000 Euros por el mantenimiento de los 10 clientes que ya se encuentran en esa fase. Al final de este año pasaremos a tener un total de 14 clientes.

Nuestro crecimiento como empresa será adecuado si vamos adquiriendo nuevos proyectos/clientes por cualquiera de las vías anteriormente reseñadas. Hay que tener en cuenta que en España hay muchas organizaciones públicas que presentan la misma problemática que tiene nuestro primer “cliente” (las posibilidades dentro de la AAPP no solo se centran en la figura de los Ayuntamientos, sino que tendremos la oportunidad de ofrecer nuestra solución tecnológica con más a Diputaciones, Comunidades Autónomas, etc.)

Para una primera estimación realizada sobre el cálculo de los costes y los beneficios en nuestro proyecto para un horizonte temporal de 5 años, hemos realizado nuestro análisis económico con

el objetivo de nunca perder rentabilidad empresarial al ajustar costes con el objeto de conseguir una estructura económica sostenible durante los 3 primeros años (que son los verdaderamente críticos según nuestros cálculos). La idea es adquirir una cartera de clientes que aumente con una tasa interanual aproximada de 2 a 4 organizaciones para, finalmente, llegar a un total de 14 clientes en el 5ª año. De esta manera, los principales gastos serán soportados por los crecientes beneficios de los 3 primeros años, lo que aportará un crecimiento sostenible para la empresa y una viabilidad adecuada.

A continuación, desarrollaremos el análisis de la viabilidad financiera del proyecto empresarial. Para ello, realizaremos su análisis económico-financiero, mediante el uso de las técnicas más utilizadas para diagnosticar la situación y la perspectiva empresarial 'real' de D.One.

De esta forma, desde una perspectiva interna, la empresa puede ir tomando las decisiones que corrijan los puntos débiles que pueden amenazar su futuro, al mismo tiempo que se saca provecho de los puntos fuertes para alcanzar sus objetivos económicos. Desde una perspectiva externa, estas técnicas también son de gran utilidad para todas aquellas personas interesadas en conocer la situación y la evolución previsible de nuestra empresa.

Costes de Legalización

Tal y como se comentó en el plan jurídico fiscal y laboral, la creación de la sociedad tiene una serie de costes añadidos:

- Coste de registro de la marca/empresa D.One de 122,4€ (https://bit.ly/3lSZVwU).
- Coste del certificado negativo de denominación social de 13,52€+IVA (https://bit.ly/395HrFF).
- Coste de la reserva de la denominación de la empresa y la constitución de la sociedad de alrededor de 19€ (https://bit.ly/2KqO6A5).
- Coste de la escritura ante notario para la constitución de la empresa de 60€.
- Coste de la inscripción de la sociedad en el Registro Mercantil de 100€.
- Saldo de 3.000€ en la cuenta de la empresa de nueva creación.
- Coste de la publicación de la creación de la sociedad en el BORME de 13€ (https://bit.ly/2HqAfsw).
- 15% correspondiente al Impuesto de Sociedades en su tipo reducido para emprendedores (1.500 €).
- **COSTE TOTAL: 4.830,75 €.**

Coste del Servicio de Paralelización

Para la infraestructura de *paralelización* (Apache Spark) se propone el componente Azure HDInsight de Microsoft que se expone en la arquitectura presentada en los capítulos ya descritos. Gracias a la elasticidad de Azure, se pueden probar diversos tamaños de clúster para determinar la combinación de costo y rendimiento que resulte óptima. Suponiendo un caso típico, los clústeres Spark para HDInsight se implementan con tres roles:

- Nodo principal (2 nodos)
- Nodo de trabajo (al menos 1 nodo)
- Nodos Zookeeper (3 nodos) (gratis para nodos Zookeeper A1)

La característica principal de cada nodo es su capacidad de procesamiento general optimizados para memoria para Big Data. Sus especificaciones técnicas son las siguientes:

- 2 Procesadores Virtuales.
- 16 GB de memoria RAM

Los costes por nodo supuesto para contar con este servicio son (https://bit.ly/390BOc1):

Precio base/nodo/hora (€116,96/mes) + €0/núcleo/hora

1. Coste por nodo: 116,96 €/mes
2. Coste 6 nodos: 500,89 €/mes
3. **COSTE ANUAL (consideramos como proyecto tipo una duración de 12 meses): 6.010,77 €/año**

Coste por Almacenamiento de datos

Para la infraestructura de *almacenamiento* se propone el componente Azure Data Lake Storage Gen1 de Microsoft que se expone en la arquitectura presentada en los capítulos ya descritos

El coste de almacenamiento es por uso bajo de demanda, por lo cual se ha estimado un uso mensual inicial por cada nodo de 1TB, que tiene un valor de €29,51 al mes (https://bit.ly/38Z336O) . Lo cual nos da el siguiente detalle:

1. Coste Mensual por nodo: 29,51 €
2. Coste por 3 nodos Big Data Mensual: 88,54 €
3. **COSTE ANUAL 8 NODOS: 1.062,57 €**

Coste Red Blockchain

Según se ha indicado en apartados anteriores, la referencia para la estructura de la capa de Blockchain se enmarca en el acuerdo con la Asociación Alastria.

La asociación con esta entidad implica una **cuota anual** como asociado de: **500€.**

Alastria, en su página web: "HAZTE SOCIO", establece esta cuota para las PYMES con menos de 100 empleados, como en el caso de la Sociedad Limitada que se pretende crear para este proyecto.

Este acuerdo marco nos permite tener acceso a otros socios tecnológicos que nos facilitarán el acceso a la Red H de Alastria, mediante una máquina virtual de las siguientes características físicas:

MÁQUINA VIRTUAL		
4 VCPUs	8 GB RAM	100 GB disco.
NOTA. Consta de un script que limpia los logs de Hyperledger periódicamente, y que es parametrizable.		

En esa infraestructura, por defecto, se despliegan los siguientes elementos:

- 1 Hyperledger CA (integrada con un servicio de desarrollo propio de seguridad, y que incluye una wallet por usuario para el material criptográfico).
- 1 Nodo Peer (Anchor).
- 1 Nodo Orderer (con consenso RAFT).

Que, para el caso en estudio, se le añaden **cuatro nodos más**, con las funciones o roles que se muestran a continuación:

- 1 Nodo Peer (Leader).
- 1 Nodo Peer (Endorser).
- 2 Nodos Peer (Regular).

Partiendo de esta arquitectura se producen los siguientes costes adicionales:

COSTES ADICIONALES MENSUALES	
Licencia de uso del software propio del proveedor	20€
Despliegue, operación y monitorización	20€
Nodos adicionales	80€
TOTAL	**120€**

Tabla 6. Costes adicionales mensuales de la arquitectura. Elaboración propia.

Por lo tanto, los costes mínimos anuales para una organización como AAPP_A serían:

COSTES DE DESPLIEGUE DE LA ARQUITECTURA EN UNA ORGANIZACIÓN	
Costes de la Red H de Alastria	1.440€
TOTAL	**1.440€**

Tabla 7. Costes de despliegue de la arquitectura en la Red H de Alastria (1 organización/nodo). Elaboración propia.

Para el supuesto desarrollado en el apartado **"Diseño de la HLF para el proyecto D.One+"** de este documento, que consta de tres organizaciones habría que tener en cuenta el caso de uso,

y en particular, la parte operativa asociada a los protocolos de comunicación ente nodos de distintas organizaciones.

COSTES DE DESPLIEGUE DE LA ARQUITECTURA EN TRES ORGANIZACIONES	
Costes de la Red H de Alastria	4.320€
Costes adicionales operativos	480€
TOTAL	**4.800€**

Tabla 8. Costes de despliegue de la arquitectura en la Red H de Alastria (3 organizaciones/nodos).

El planteamiento para la distribución de los nodos es diseminarlos por las dos regiones, denominadas **Datacenter Region**, que tiene Azure Microsoft en Europa, concretamente en el North Europe que se encuentra en **IRLANDA** y el West Europe que se sitúa en **PAÍSES BAJOS**, según los datos de web de Microsoft Azure .

Dado que ambas regiones se encuentran en Europa, se cumple un punto muy importante desde el punto de vista de la RGDP, dado que los servidores deben operar en un país que pertenezca a la Unión Europea.

Coste Total Infraestructura Tecnológica

Con esta infraestructura tecnológica estimamos que tanto los clústeres de paralelización, almacenamiento, como los nodos de la red Blockchain, son suficientes para la carga de trabajo provocada por la gestión de todos los registros para los 5 primeros años según la estimación de crecimiento potencial de clientes expuesto al inicio del plan financiero (queremos señalar que nuestros proveedores Azure Microsoft y Alastria, ofrecen un servicio escalable en el tiempo, con lo cual, en caso de crecer la red y/o los requerimientos de clientes se pueden adherir más clústeres para el procesamiento de Big Data o más Nodos para la red Blockchain).

El coste final es la sumatoria de todos los ítems anteriores. Con lo cual, el resultado interanual teniendo en cuenta el crecimiento potencial de clientes expuesto al inicio del plan financiero ascendería a:

AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
-9.133,34 €	-12.013,34 €	-16.333,34 €	-22.093,34 €	-27.853,34 €

Tabla 9. Costes Interanuales Infraestructura Tecnológica. Fuente: Elaboración propia.

Flujos de Caja

El flujo de caja nos muestra las magnitudes a tener en cuenta para conocer la viabilidad de la empresa, presentado la acumulación neta de los activos líquidos para un periodo determinado (en este caso de 5 años) y, por esta razón, constituye un indicador importante para medir la liquidez de una empresa en un periodo determinado. A continuación, vemos el detalle en la siguiente tabla:

FLUJOS DE CAJA					
	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
SALDO INICIAL	10.000,00 €	- €	- €	- €	- €
INGRESOS	- €	42.000,00 €	81.000,00 €	132.000,00 €	180.000,00 €
Nuevos clientes + Mantenimiento	- €	42.000,00 €	81.000,00 €	132.000,00 €	180.000,00 €
INVERSIÓN	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €
Equipos informáticos	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €
GASTOS	- 13.964,09 €	- 19.690,34 €	- 29.860,34 €	- 110.181,94 €	- 123.141,94 €
Personal	- €	- €	- €	- 80.000,00 €	- 80.000,00 €
Plan Comercial + Publicidad	- €	- 3.000,00 €	- 3.000,00 €	- 3.000,00 €	- 3.000,00 €
Infraestructura Tecnológica	- 9.133,34 €	- 12.013,34 €	- 16.333,34 €	- 22.093,34 €	- 27.853,34 €
Costes Legales	- 4.830,75 €	- €	- €	- €	- €
Impuestos	- €	- 4.677,00 €	- 10.527,00 €	- 5.088,60 €	- 12.288,60 €
Saldo Final	- 5.364,09 €	20.909,66 €	49.739,66 €	20.418,06 €	55.458,06 €

Tabla 10. Flujos de Caja. Fuente: Elaboración propia.

Como se puede comprobar, existe una correcta planificación financiera que nos permitirá conseguir un Cash-Flow positivo en los 5 primeros años de la empresa, lo que constituye un pilar fundamental para la viabilidad de nuestro proyecto. Además, permitirá anticiparnos a posibles futuros déficits y establecer una base sólida para la solicitud de créditos a largo plazo a Entidades Financieras.

Análisis de Costes

Nuestra estructura de costes pivotará sobre 2 puntos clave:

1. *Infraestructura Tecnológica*: Basada en *servicios cloud* los cuales definen costes mensuales fácilmente mensurables. Al inicio de cada proyecto, se elevará al cliente una propuesta de servicio para la infraestructura que soportará su solución MDM y que repercutirá directamente en los costes del proyecto.
2. *Consultoría*: forma parte del servicio personalizado que otorgará **D.One** a sus clientes y su coste repercutido vendrá a solventar los gastos asociados al capital humano de nuestra empresa.

La solución tecnológica está compuesta por una serie de elementos hardware que tendrán repercusión en el precio final del servicio para nuestros clientes. A continuación, mostramos una tabla con los baremos comprendidos para distintos **números de nodos** ya que uno de los componentes del precio de la solución dependerá del número de clientes y la estimación realizada está vinculada a lo señalado al inicio del capítulo “Plan Financiero” para cada una de las 5 anualidades propuestas Sobre este importe aplicamos un sobrecoste del 30% para la obtención de un beneficio adecuado.

AÑO	NÚMERO DE NODOS	TOTAL € PROYECTO
1	1	9.133,34 €
2	3	12.013,34 €
3	6	16.333,34 €
4	10	22.093,34 €
5	14	27.853,34 €

Tabla 11. Precio del Proyecto según Número de Nodos. Fuente: Elaboración propia.

Hemos cuantificado un horizonte empresarial ejemplo para 5 años con proyectos de una duración de 12 meses y con costes estimados como reales, tanto para los conceptos de consultoría, como salarios, publicidad, internet/telefonía y equipos informáticos. La aspiración de **D.One** es tener la capacidad de abordar varios proyectos en paralelo, lo cual será la única forma de asegurar la viabilidad económica de nuestra empresa. A continuación, se presenta un modelo ejemplo de costes con las características anteriormente descritas:

COSTES	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
Salarios	- €	- €	- €	- 80.000,00 €	- 80.000,00 €
Infraestructura Tecnológica	- 9.133,34 €	- 12.013,34 €	- 16.333,34 €	- 22.093,34 €	- 27.853,34 €
Plan Comercial + Publicidad	- €	- 3.000,00 €	- 3.000,00 €	- 3.000,00 €	- 3.000,00 €
Equipos Informáticos	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €
TOTAL EUROS	**-10.533,34 €**	**-16.413,34 €**	**-20.733,34 €**	**-106.493,34 €**	**-112.253,34 €**

Tabla 12. Estimación de costes anuales. Fuente: Elaboración propia.

- La aportación económica inicial por parte de los integrantes del equipo ascenderá a un total de 10.000€ (2.000 por persona).
- Durante los 3 primeros años, no existirán gastos imputables a la empresa relacionados con el concepto “Salarios”.
- Los equipos informáticos serán amortizados en los 5 años de referencia.
- La publicidad y el marketing será vital para darnos a conocer en el sector. Para ello, realizaremos desde el 2º año (donde ya habremos iniciado la apertura a nuestros clientes potenciales) una campaña publicitada en revistas digitales y foros especializados con la que nuestra solución MDM ganará visibilidad.
- Los salarios serán distribuidos a partir del 4º año, cuando **D.One** ya tendrá la provisión económica suficiente para su aplicación a todos los socios de la empresa.
- La gestión de todos los asuntos contables y fiscales serán realizados íntegramente por el equipo de **D.One**.

Análisis de Beneficios

En nuestro primer año nuestro único cliente será el Ayuntamiento 1 de Gran Canaria. A partir de la entrada en producción de la solución en este organismo, nuestra misión será la captación de nuevos clientes (tanto en la AAPP, como en la empresa privada) que posibilitarán el crecimiento de la empresa. Por lo tanto, durante el primer año, nuestros ingresos no cubrirán en su totalidad

los costes anuales soportados debido a que únicamente tendremos el cliente ya referido, pero la estimación propuesta a partir del 2º año nos indica que los costes podrán ser amortizados con la obtención de nuevos clientes, centrándonos en la AAPP. A continuación, presentamos los cálculos estimados de los ingresos a percibir y la estimación de los clientes a obtener en un periodo considerado de 5 años y teniendo en cuenta que el precio de la implantación de nuestra solución asciende a 15.000 € (que se corresponde con el importe máximo de licitación de un "Contrato Menor" de servicios para la AAPP) y un mantenimiento anual a 12.000 €.

AÑO 1		
1 Cliente - Contrato Mantenimiento	0,00 €	PROYECTO
1 Cliente	0,00 €	
AÑO 2		
2 Clientes - Consultoria / implantación	30.000,00 €	42.000,00 €
1 Cliente - Contrato Mantenimiento	12.000,00 €	
AÑO 3		
3 Clientes - Consultoria / implantación	45.000,00 €	81.000,00 €
3 Clientes - Contrato Mantenimiento	36.000,00 €	
AÑO 4		
4 Clientes - Consultoria / implantación	60.000,00 €	132.000,00 €
6 Clientes - Contrato Mantenimiento	72.000,00 €	
AÑO 5		
4 Clientes - Consultoria / implantación	60.000,00 €	180.000,00 €
10 Clientes - Contrato Mantenimiento	120.000,00 €	

Tabla 13. Estimación de beneficios anuales en servicio de consultoría y mantenimiento. Fuente: Elaboración propia.

Total Clientes Nuevos por año				
AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
Proyecto	2	3	4	4

Tabla 14. Estimación de clientes nuevos por año. Fuente: Elaboración propia.

Cartera de clientes por año				
AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
1	3	6	10	14

Tabla 15: Estimación de cartera de clientes anual. Fuente: Elaboración propia.

Costes vs Ingresos en los 5 años analizados

A partir de la entrada en producción de la solución en San Bartolomé de Tirajana, nuestra misión será la captación de nuevos clientes que posibiliten el crecimiento de la empresa. Por lo tanto, durante el primer año, nuestros ingresos no cubrirán en su totalidad los costes anuales soportados debido a que únicamente tendremos ese cliente, pero la estimación propuesta a partir del 2º año indica que los costes podrán ser amortizados tal y como se muestra en la siguiente tabla.

	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
COSTES	- 15.364,09 €	- 21.090,34 €	- 31.260,34 €	- 111.581,94 €	- 124.541,94 €
INGRESOS	- €	42.000,00 €	81.000,00 €	132.000,00 €	180.000,00 €
TOTAL EUROS	-15.364,09 €	20.909,66 €	49.739,66 €	20.418,06 €	55.458,06 €

Tabla 16. Costes vs Ingresos. Fuente: Elaboración propia.

Cuenta de Resultados

La cuenta de resultados ofrece información de las partidas de ingresos que ha conseguido la empresa y la partida de los gastos en los que ha incurrido, dando como resultado los beneficios o las pérdidas a lo largo de un periodo. **Beneficio = Ingresos – Gastos**. Por lo tanto, la cuenta de resultados nos muestra una información imprescindible para controlar la evolución de nuestro negocio. Podemos observar los resultados calculados en la siguiente tabla:

Cuenta de Pérdidas y Ganancias	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
Importe Neto de la Cifra de Negocios	- €	**42.000,00 €**	**81.000,00 €**	**132.000,00 €**	**180.000,00 €**
Nuevos clientes + Mantenimiento	- €	42.000,00 €	81.000,00 €	132.000,00 €	180.000,00 €
Aprovisionamientos	**- 9.133,34 €**	**- 15.013,34 €**	**- 19.333,34 €**	**- 25.093,34 €**	**- 30.853,34 €**
Infraestructura Tecnológica	- 9.133,34 €	- 12.013,34 €	- 16.333,34 €	- 22.093,34 €	- 27.853,34 €
Gestión Laboral, Fiscal y Contable	- €	- €	- €	- €	- €
Plan Comercial + Publicidad	- €	- 3.000,00 €	- 3.000,00 €	- 3.000,00 €	- 3.000,00 €
Gastos de Personal	**- €**	**- €**	**- €**	**- 80.000,00 €**	**- 80.000,00 €**
Sueldos, Salarios y Asimilados	- €	- €	- €	- 80.000,00 €	- 80.000,00 €
Amortización de Inmovilizado	**- 1.400,00 €**	**- 1.400,00 €**	**- 1.400,00 €**	**- 1.400,00 €**	**- 1.400,00 €**
Equipos Informáticos	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €	- 1.400,00 €
RESULTADOS DE EXPLOTACIÓN	**- 10.533,34 €**	**25.586,66 €**	**60.266,66 €**	**25.506,66 €**	**67.746,66 €**
RESULTADOS ANTES DE IMPUESTOS	**- 10.533,34 €**	**28.586,66 €**	**63.266,66 €**	**28.506,66 €**	**70.746,66 €**
IMPUESTOS SOBRE BENEFICIOS	- €	**- 4.573,87 €**	**- 10.122,67 €**	**- 4.561,07 €**	**- 11.319,47 €**
RESULTADO DEL PERIODO	**- 10.533,34 €**	**24.012,79 €**	**53.143,99 €**	**23.945,59 €**	**59.427,19 €**

Tabla 17. Cuenta de resultados. Fuente: Elaboración propia.

El proyecto madurará en los 5 años iniciales adquiriendo al final de este periodo unos beneficios ya consolidados. Hemos realizado esta previsión de una manera bastante conservadora y aun así los datos nos proporcionan una situación económica adecuada con la existencia de unos ingresos que nos permiten ser asalariados de la empresa y obtener beneficios a la misma vez ya en el Año 4.

Como hemos comentado anteriormente, para el Año 1 el resultado del periodo es negativo debido a la existencia de 1 sólo proyecto. Los ingresos aparecerán a partir del Año 2 por la existencia de una cartera de clientes que reportará ingresos por los conceptos de mantenimiento e implantación.

Balance de Situación

El Balance de Situación (o balance general) es el resumen de todas las posesiones (activos) y todas las deudas y el capital de una organización en un periodo contable determinado. En sí, el balance es como una fotografía del patrimonio. Junto con la cuenta de pérdidas y ganancias, forma el informe más importante para las cuentas anuales de una organización empresarial. La estructura del balance está dividida en activos y pasivos. El activo recoge todos los bienes y derechos que posee la compañía. Más específicamente, los activos se dividen en inmovilizados (inversiones a largo plazo), existencias, realizables y disponibles. En el pasivo se encuentran los fondos propios, exigibles a largo plazo y exigibles a corto plazo. Cada una de las partidas

mencionadas incluye sus correspondientes subpartidas. A continuación, se muestra la proyección realizada para nuestro periodo de 5 años ya señalado a lo largo de este capítulo:

ACTIVO	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
ACTIVOS NO CORRIENTES	5.600,00 €	4.200,00 €	2.800,00 €	1.400,00 €	- €
Equipos informáticos	7.000,00 €	7.000,00 €	7.000,00 €	7.000,00 €	7.000,00 €
Amortización Acumulada de equipos informáticos	- 1.400,00 €	- 2.800,00 €	- 4.200,00 €	- 5.600,00 €	- 7.000,00 €
ACTIVOS CORRIENTES	2.180,00 €	26.503,00 €	59.653,00 €	31.835,40 €	72.635,40 €
Efectivos y otros activos liquidos equivalentes	2.180,00 €	26.503,00 €	59.653,00 €	31.835,40 €	72.635,40 €
TOTAL ACTIVO	7.780,00 €	30.703,00 €	62.453,00 €	33.235,40 €	72.635,40 €
PATRIMONIO NETO Y PASIVO	**AÑO 1**	**AÑO 2**	**AÑO 3**	**AÑO 4**	**AÑO 5**
PATRIMONIO NETO	7.780,00 €	30.703,00 €	62.453,00 €	33.235,40 €	72.635,40 €
Capital	10.000,00 €	10.000,00 €	10.000,00 €	10.000,00 €	10.000,00 €
Reservas	- €	- 2.220,00 €	22.923,00 €	29.530,00 €	- 6.294,60 €
Resultado del Ejercicio	- 2.220,00 €	22.923,00 €	29.530,00 €	- 6.294,60 €	68.930,00 €
PASIVO NO CORRIENTE	- €	- €	- €	- €	- €
PASIVO CORRIENTE	- €	- €	- €	- €	- €
TOTAL PATRIMONIO NETO Y PASIVO	7.780,00 €	30.703,00 €	62.453,00 €	33.235,40 €	72.635,40 €

Tabla 18. Balance de situación. Fuente: Elaboración propia.

Para el correcto desarrollo de nuestra empresa es necesario que ésta ofrezca un equilibrio entre su estructura económica (activo) y su estructura financiera (neto + pasivo). Debemos tener en cuenta que para una mayor viabilidad de sus inmovilizados y sus activos corrientes necesarios para la continuidad del ciclo de explotación éstos deben estar financiados con recursos propios y/o recursos ajenos a largo plazo en su correcta proporción. Debemos evitar el desequilibrio financiero en nuestro balance de situación y para ello calcularemos el Fondo de Maniobra que, en gestión financiera, se corresponde con la parte del activo circulante que es financiada con recursos de carácter permanente. Es una medida de la capacidad que tiene una empresa para continuar con el normal desarrollo de sus actividades en el corto plazo. Para el caso D.One, en el Año 1 soportaremos un pequeño desequilibrio financiero, el cual es totalmente viable porque en los siguientes 4 años el fondo de maniobra es positivo, porque el activo circulante está siendo financiado con recursos permanentes, lo que indica que D.One mantiene en su conjunto un correcto equilibrio financiero.

	Fondo de Maniobra				
	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5
Activo corriente	2.180,00 €	26.503,00 €	59.653,00 €	31.835,40 €	72.635,40 €
Pasivo corriente	- €	- €	- €	- €	- €
TOTAL	2.180,00 €	26.503,00 €	59.653,00 €	31.835,40 €	72.635,40 €

Tabla 19. Fondo de Maniobra. Fuente: Elaboración propia.

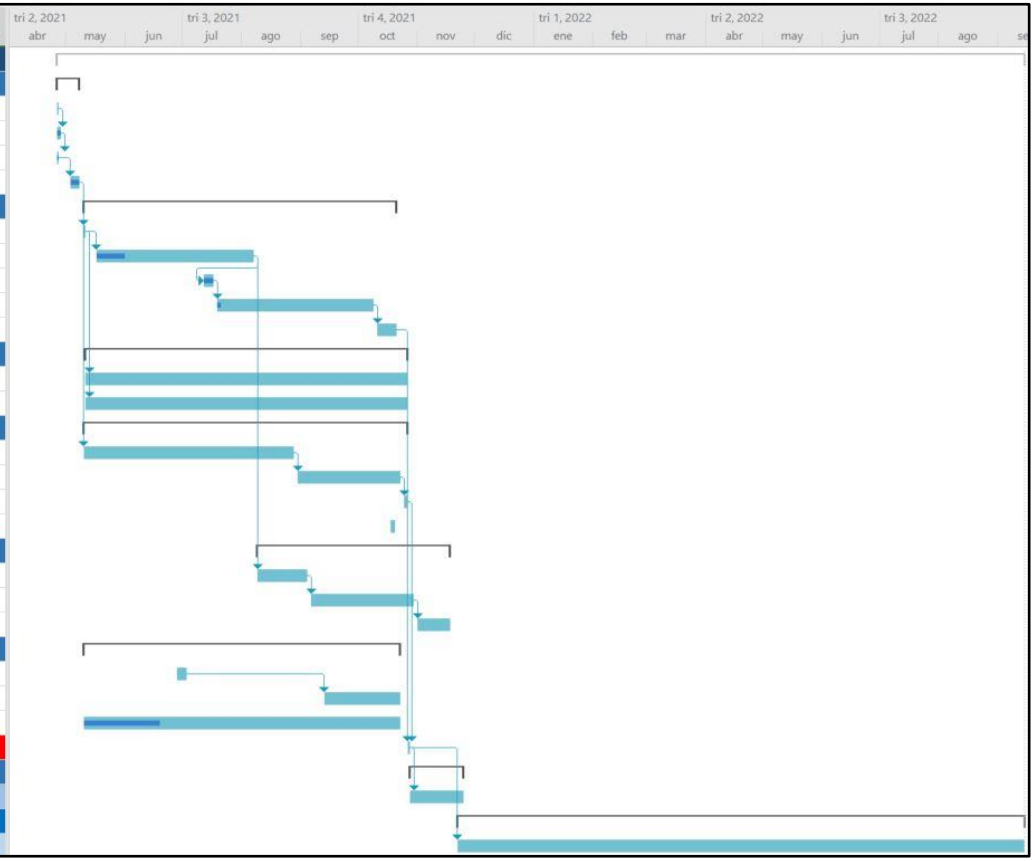
11 CALENDARIO DE EJECUCIÓN

A continuación, procedemos a definir la planificación temporal y de recursos para un proyecto tipo en función de los procesos, así como de las actividades clave anteriormente identificadas. Para su desarrollo, se establece una duración media de 6 meses. Sin embargo, ésta podrá variar

en función de las necesidades específicas de cada cliente y porque contemplamos la posibilidad de desarrollar en paralelo varias actividades en función de los perfiles y las dependencias entre procesos, con el fin de mejorar los tiempos de gestión y entrega dada la interdisciplinaridad de los miembros del equipo. En el cronograma que a continuación mostramos, se identifica para cada tarea los perfiles de los recursos empleados, donde la dedicación a un trabajo concreto (tal y como señalábamos antes) no impedirá el apoyo para otros desarrollados de manera simultánea. Será competencia del jefe del proyecto la gestión adecuada de la asignación de los analistas a las tareas designadas. Aunque los tiempos inicialmente calculados pueden verse comprometidos por circunstancias externas a la empresa, como puede ser la situación que estamos soportando (pandemia COVID-19).

De esta manera, si un proyecto se iniciara el 11 de junio de 2020, la solución MDM **D.One+** estaría en fase de producción el 12 de diciembre de 2021. Posteriormente, dentro del contrato de implantación de la solución, ofrecemos un periodo que dura hasta los 12 meses desde el inicio del proyecto (6 + 6), el cual puede ser prorrogable por el cliente en periodos anuales a través del acuerdo de los subsiguientes contratos de mantenimiento.

Nombre de tarea	Duración
PROJECT IMPLANTACIÓN SOLUCIÓN MDM D.One+	363 días
1 Inicio del Proyecto + Definición del Modelo de Relación	10 días
1.1 Reunión Inicial y Presentación a Usuarios Clave	1 día
1.2 Designación y presentación del equipo de trabajo	2 días
1.3 Acta de Inicio y Constitución del Proyecto	1 día
1.4 Elaboración y Aprobación del Plan de Proyecto	5 días
2 Infraestructuras + Fase de Pruebas	118 días
2.1 Estudio y dimensionamiento de las infraestructuras	1 día
2.2 Esquema de conectividad Sistemas Externos - MDM D.One+	60 días
2.3 Despliegue del SI y disponibilidad del entorno de Preproducción	5 días
2.4 Pruebas y ajuste del sistema de alta disponibilidad	60 días
2.5 Preparación y Pruebas sobre máquina del entorno de Producción	8 días
3 Consultoría RGPD+ENS	121 días
3.1 Marco operacional	121 días
3.2 Marco de protección	121 días
4 Migración +Record Linkage + Smart Contracts + Validación de los Datos	122 días
4.1 Diseñar y desarrollar la extracción, carga y transformación de los datos (ETL).	80 días
4.2 Carga y validación de los datos entorno Preproducción	40 días
4.3 Carga y validación de los datos entorno Producción	2 días
4.4 Entrega de la Documentación Técnica	2 días
5 Personalización + Parametrización del Sistema	73 días
5.1 Análisis previo y validación	20 días
5.2 Personalización e Implementación de Informes, Cuadros de Mandos y Dashboard	40 días
5.3 Asignación de Roles y Perfiles	13 días
6 Formación + Gestión del Cambio	120 días
6.1 Elaboración y Aprobación del Plan de Formación	5 días
6.2 Formación de Usuarios (según Departamento) + Entrega Documentación Manuales de Usuario	30 días
6.3 Acciones de Gestión del Cambio y Reuniones de Seguimiento	120 días
7 PUESTA EN PRODUCCIÓN	1 día
8 Estabilización del Sistema	20 días
8.1 Tareas de corrección y estabilización de datos y procesos	20 días
9 SISTEMA EN PRODUCCIÓN DURANTE LA VIGENCIA DEL CONTRATO	213 días
9.1 Soporte a la explotación	213 días



12 CONCLUSIONES DEL PROYECTO

Podemos señalar que, por todo lo expuesto, planteamos un proyecto rentable, recurrente y totalmente escalable con una utilidad clara en el mercado y sin iniciativas de este tipo de solución en la actualidad por parte de nuestros potenciales competidores. Desarrollaremos tecnología que mejorará operativamente las herramientas que existen en el mercado y, una vez implantada, ofrecemos un servicio de mantenimiento y soporte específico y totalmente personalizado para cada uno de nuestros clientes.

Además:

1. En la actualidad, la existencia de datos duplicados, inconsistentes e incompletos entre los distintos sistemas de gestión de una organización, las incidencias en la gobernabilidad del dato, así como el incumplimiento de la legislación de protección de los registros supone un grave problema para cualquier organización.
2. La predisposición de los responsables de las organizaciones de resolver esta circunstancia a través de la utilización de la tecnología supone que el desarrollo de un Modelo de Datos Maestros se convierta en una línea de negocio acertada para afrontar la problemática derivada de la falta de unicidad y coherencia de los datos más importantes de las organizaciones.
3. La creación de un fichero de datos maestros limpio y confiable es un desafío complejo, pero son muchos los beneficios de crear un repositorio maestro común. El emparejamiento en este proyecto de un MDM + Big Data + Blockchain está repleto de retos que hemos tenido que afrontar, ya que partimos de la combinación de tecnologías sujetas a una evolución continua que ha requerido tiempo, dedicación, implicación e I+D+i para lograr un producto/solución final.
4. Actualmente, no existe en el mercado una solución que cumpla con todas estas características y el desarrollo de D.One posibilitará a las organizaciones la automatización de los procesos de integración de sus datos maestros internos de personas/clientes y, a futuro, también de los datos externos a la misma, cumpliendo además con garantía las exigencias impuestas por el RGPD. La solución supone una auténtica innovación con respecto a las soluciones MDM propietarias existentes en el mercado.
5. Nuestra solución está diseñada para adaptarse a las necesidades particulares de cada cliente con el objetivo de reducir sus costes operativos y mejorar sus procesos de gestión lo que, a medio plazo, se traducirá en inversión para la organización.

13 ANEXO: REGLAMENTO GENERAL DE PROTECCIÓN DE DATOS

Como bien se ha justificado y desarrollado a lo largo de los capítulos, concretamente en el capítulo 9.2 **Reglamento General De Protección De Datos Y Esquema Nacional De Seguridad,** tener claro el marco legal supone una premisa fundamental para la efectiva implantación de la solución **D.One+** en los organismos públicos. Por ello, se presenta un breve y conciso recorrido de los aspectos relevantes de la normativa de protección de datos.

Anteriormente, la Ley Orgánica de Protección de Datos de Carácter Personal 15/1999, LOPD, enunció una serie de obligaciones, ya inexistentes:

- Existencia del fichero o tratamiento, finalidad y destinatarios.
- Carácter obligatorio o no de la respuesta y consecuencias.
- Posibilidad de ejercitar los derechos de: acceso, rectificación, cancelación y oposición.
- Identidad y dato de contacto del responsable de tratamiento.

Con la Ley Orgánica de Protección de Datos y Garantía de Derechos Digital 3/2018, LOPDGDD en adelante, surgen una serie de conceptos de necesario conocimiento para el tratamiento de datos personales:

1. **Responsabilidad proactiva.**
   Este principio requiere que las organizaciones analicen qué datos tratan, con qué finalidades lo hacen y qué tipo de operaciones de tratamiento llevan a cabo. A partir de este conocimiento deben determinar de forma explícita la forma en que aplicarán las medidas que el RGPD prevé.
2. **Consentimiento inequívoco.**
   Es aquel que se ha prestado mediante una manifestación del interesado o mediante una clara acción afirmativa. Debe ser libre, específico, informado e inequívoco. No se admiten formas de consentimiento tácito o por omisión.
   Se contemplan situaciones en las que el consentimiento, además de inequívoco, ha de ser explícito:
   - Tratamiento de datos sensibles.
   - Adopción de decisiones automatizadas.
   - Transferencias internacionales.
3. **Figuras de responsable y encargado de tratamiento, y Delegado de Protección de Datos (DPD).**
   El RGPD, contiene obligaciones expresamente dirigidas a los encargados. En determinadas materias los encargados tienen obligaciones propias que establece el RGPD, que no se circunscriben al ámbito del contrato que los une al responsable, y que pueden ser supervisadas separadamente por las autoridades de protección de datos.

   Por ejemplo:

- Deben mantener un registro de actividades de tratamiento (RAT).
- Deben determinar las medidas de seguridad aplicables a los tratamientos que realizan.
- Deben designar a un Delegado de Protección de Datos en los casos previstos por el RGPD.

Los encargados pueden adherirse a códigos de conducta o certificarse en el marco de los esquemas de certificación previstos por el RGPD.

4. **Derechos de interesados.**

   Hay un considerable cambio de los derechos planteado en la LOPD y la nueva LOPDyGDD. El derecho de cancelación es sustituid por el de supresión y se añaden otros. Los derechos presentes en la nueva normativa de protección de datos son:

   - Acceso.
   - Rectificación.
   - Supresión.
   - Oposición
   - Portabilidad.
   - Oposición.
   - Limitación.

5. **Medidas de responsabilidad activa.**
   - Registro de Actividades de Tratamientos (RAT).
   - Protección de Datos desde el Diseño y por Defecto.
   - Medidas de Seguridad.
   - Notificación de “violaciones de seguridad de los datos”.

Dado el contexto tecnológico en el que se encuadra la solución **D.One+**, cuyo funcionamiento radica en el uso de tecnologías disruptivas y que por su naturaleza realizan un continuo tratamiento de datos, el impacto sobre la privacidad de los mismos y la necesidad de cumplimiento del artículo 25 del RGPD se vuelve una necesidad para asegurar la viabilidad del proyecto. En estos términos, es requisito fundamental la implantación de medidas técnicas y organizativas efectivas que aseguren el respeto a los derechos y libertades de las personas en lo que a su tratamiento de datos personales se refiere.

La privacidad desde el diseño y por defecto implica dar un enfoque orientado a la gestión del riesgo y responsabilidad proactiva que permite establecer los requisitos de privacidad del sistema. Para ello, es necesario recorrer una serie de puntos:

- **Análisis de riesgos**. Que establecen los objetivos de protección de datos y seguridad, desde el punto de vista de la privacidad.
- Estudio de las **estrategias de privacidad** en las que se concretan los requisitos de cada objetivo de privacidad.

- En la **fase de diseño**, integrar los patrones de diseño de la privacidad mediante soluciones ya conocidas.
- En la **fase de desarrollo**, se realiza la implementación de dichos patrones mediante el uso de "Privacy Enhancing Technologies".

La aplicación de medidas de protección de la privacidad debe contemplarse en todos los procesos y prácticas de negocio involucrados en el tratamiento de datos, para conseguir así, una verdadera gobernanza de la gestión de los datos personales. Por ello, la protección de datos no ha sido considerada como una capa añadida a la solución **D.One+**, si no que ha estado presente desde las fases iniciales donde se han definido los procesos del sistema.

Para asegurar el cumplimiento del artículo 25 del RGPD, y por tanto asegurar la privacidad, haremos un necesario recorrido por los siete principios fundacionales del mismo:

1. **Proactivo, no reactivo; preventivo, no correctivo.**
   Anticipación a los eventos que afecten a la privacidad antes de su suceso. Para ello, es necesario identificar los posibles riesgos a los derechos y libertades de los interesados y minimizarlos para que no se materialicen en daños. Por tanto, la privacidad por defecto se aleja de la práctica de "subsanar" y se adelanta a la materialización del evento de riesgo.
2. **La privacidad como configuración predeterminada.**
   La privacidad desde el diseño tiene como objetivo ofrecer al interesado el máximo nivel de privacidad. Este principio se fundamenta en la minimización de datos a lo largo de todas las etapas del tratamiento: recogida, uso, conservación y difusión.
   Es por ello que es necesario:
   - Proponer criterios de recogida limitados a la finalidad que persigue el tratamiento.
   - Limitar el uso de datos personales a las finalidades para las que fueron recogidas y asegurar que existe una base legitimadora del tratamiento.
3. **Privacidad incorporada en la fase de diseño.**
   La privacidad debe estar integrada en el conjunto de requisitos no funcionales desde el mismo momento en el que es concebida y diseñada. Para garantizarla, se debe:
   - Considerar como requisito necesario en el ciclo de vida de sistemas y servicios.
   - Ejecutar un análisis de los riesgos para los derechos y libertades de las personas.
4. **Funcionalidad total.**
   Desmitificando la dicotomía privacidad vs funcionalidad o beneficio empresarial y buscar el balance óptimo con el objetivo de ofrecer una búsqueda "*Win-Win*" en todo el ecosistema. Para ello, las organizaciones deben:
   - Asumir que pueden coexistir intereses diferentes y legítimos.

- Si la solución propuesta plantea amenazas a la privacidad, buscar nuevas soluciones y alternativas para alcanzar las distintas funcionalidades perseguidas.

5. **Aseguramiento de la privacidad en todo el ciclo de vida.**

   Para integrar la privacidad a lo largo de todas las etapas del tratamiento de datos de **D.One+**, se deben analizar rigurosamente las diferentes operaciones implicadas en el proceso e implementar en cada una las medidas más propicias, entre las que destacan:

   - Seudonimización o técnicas de anonimización.
   - Clasificación y organización de los datos y operaciones de tratamiento.
   - El cifrado por defecto de modo de que el estado “natural” de los datos en caso de pérdida o robo sea “ilegible”.

6. **Visibilidad y transparencia.**

   La prueba inequívoca para garantizar la privacidad es poder demostrarla, de manera que sea posible verificar que el tratamiento es acorde a la información dada.

   Para asegurar una efectiva práctica de la transparencia y visibilidad es necesario que **D.One+** establezca una serie de medidas:

   - Publicar las políticas de privacidad y protección de datos.
   - Desarrollar cláusulas de información claras y concisas para el interesado.
   - Difundir la identidad y contacto de la persona responsable en materia de privacidad.
   - Establecer mecanismo de comunicación, compensación y reclamación accesibles y sencillos dirigidos a los titulares de los datos.

7. **Respeto por la privacidad de los usuarios.**

   Para garantizar la privacidad de los sujetos de datos al realizar el diseño de una solución, producto o servicio es necesario:

   - Implementar configuraciones de privacidad por defecto que permita informar a los usuarios sobre las consecuencias a su privacidad.
   - Facilitar información completa y adecuada que conduzca a un consentimiento informado, libre, específico e inequívoco.
   - Proporcionar a los interesados el acceso a sus datos y a la información detallas de las finalidades del tratamiento.
   - Implementación de mecanismos eficientes para el ejercicio de derechos en materia de protección de datos.

Objetivos de privacidad y seguridad

Existen otros factores de riesgo que pueden aparecer durante un procesamiento autorizados de los datos, a parte de aquellos que puedan afectar a los tradicionales objetivos de la seguridad (confidencialidad, integridad y disponibilidad), que han de incluirse en el esquema de análisis. Por ello, aparecen tres **nuevos objetivos de protección**, específicos de la privacidad y cuya garantía se convierte en salvaguarda de los principios de tratamiento establecidos por el RGPD:

- Desvinculación. Incapacidad de vinculación de datos personales.
- Transparencia. Clarificar el tratamiento de los datos.
- Control. Capacidad de poder intervenir en el tratamiento.

Estos objetivos determinan otro tipo de requisitos no funcionales que debe satisfacer el sistema y que se convierten en las entradas de los procesos de diseño de la privacidad. A continuación, se expone en la siguiente tabla, la interrelación de dichos requisitos y los objetivos de protección:

OBJETIVOS DE PROTECCIÓN DE LA PRIVACIDAD		
DESVINCULACIÓN	**TRANSPARENCIA**	**CONTROL**
Minimización de datos **Limitación del plazo de conservación** **Integridad y confidencialidad**	Licitud, lealtad y transparencia Limitación de la finalidad	Limitación de la finalidad Exactitud Integridad y confidencialidad Responsabilidad proactiva

Tabla 20. Objetivos de Protección de la Privacidad. Fuente: Elaboración propia.

Tras la definición en un contexto de protección de datos de la privacidad, así como las funcionalidades más relevantes de la misma, para materializar dichas medidas es necesario conocer el concepto de **ingeniería de la privacidad.** Se trata un proceso sistemático cuya finalidad es traducir en términos prácticos y operativos los principios de la privacidad desde el diseño dentro del ciclo de vida de los sistemas de información encargados del tratamiento de datos personales:

- Especificando las propiedades y funcionalidades de privacidad que debe cumplir el sistema de una manera que sea posible su diseño e implementación.
- Diseñando la arquitectura e implemento los elementos del sistema que den cobertura a los requisitos de privacidad definidos.
- Confirmando que los requisitos de privacidad definidos han sido correctamente implementados y satisfacen las expectativas de las partes interesadas.

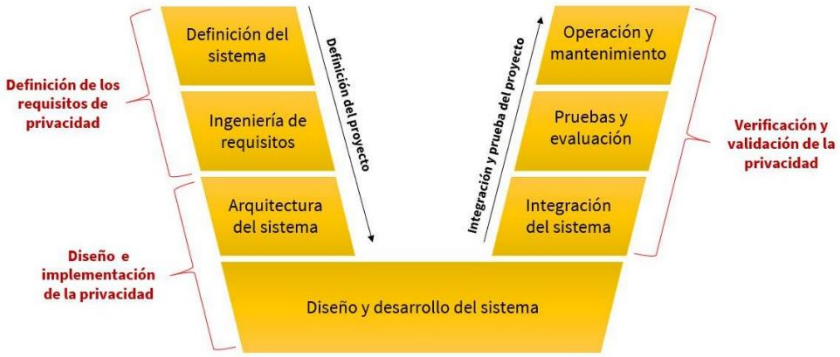


Ilustración 36. Ingeniería de Privacidad. Guía AEPD

14 INDICE DE ILUSTRACIONES

15 INDICE DE TABLAS